SEARCH 4 Protocol
Version 3: March 2, 2017
SUMMARY OF CHANGES

Page 7-8 (Table 3), Updated description of cohort survey-only recruitment targets

Page 8 (Table 3), Echocardiogram recruitment update (CA site not participating)

Page 9, 14, Addition of miRNA collection for storage (cohort participants)

SEARCH 4 Protocol
Version 2: February 11, 2016
SUMMARY OF CHANGES

Page 1, Funding mechanisms identified for Registry Study

Page 2, Specific Aims for Registry Study

Page 2, Funding mechanisms identified for Cohort Study

Page 3, Cohort sampling (subset of SEARCH 3)

Page 3, Summary of new measures in Cohort Study

Page 3, Specific Aims for Cohort Study

Page 5, New unit: Cardiovascular Reading Center

Page 5, Incident year cohorts to be completed and/or initiated in Registry Study

Page 6, Estimated number of registered cases

Page 6, Sampling plan for Registry Study in-person visit

Page 7, Sampling plan for Cohort Study visit (changes in Version 1 and Version 2)

Page 7, Sampling plan for echocardiogram visit and survey-only cohort

Page 8, Period of mortality follow-up

Page 8, Informed consent for surveys obtained online (web-based)

Page 9, Transfer of data and samples to NIDDK Repository

Page 11, Update to laboratory measures obtained in IPV

Page 12-15, New measures to be obtained in Cohort Study (survey instruments, echocardiography, neurocognitive)

Page 16-19, Development & Validation Projects

Page 19-21, Statistical considerations for Registry Study

Page 22- 24, Statistical considerations for Cohort Study

Page 25, Centralized data management system

Page 26, Alert values related to albuminuria, echocardiogram

Page 28-30, Revisions to Appendix A

SEARCH for Diabetes in Youth

Phase 4 Protocol

September 30, 2015 - September 29, 2020

Version 1: October 30, 2015

Version 2: February 11, 2016

Table of Contents

1. Background

Diabetes is the third most prevalent severe chronic disease of childhood (1), and a leading cause of nephropathy, retinopathy, neuropathy, and cardiovascular disease (CVD) later in life. Although there is some evidence that rates of mortality, renal failure, and neuropathy have declined in young adults with youth-onset T1D diagnosed between the 1950s and 1980s (2), data from more contemporary cohorts are scarce. In addition, clinical care for childhood diabetes has evolved, now encompassing new insulin types and delivery systems, and new systems for monitoring glycemic excursions. Concurrently, the epidemiology of diabetes has evolved. The incidence rates of T1D have increased around the world (3, 4) and we have learned from the SEARCH for Diabetes in Youth Study that substantial proportions of adolescent minority youth now have T1D (5). Within the last two decades pediatric T2D has gone from infrequent to 15% of all diagnoses of diabetes in youth (6). Trends in the prevalence and incidence of T1 and T2D in young people are changing. Worldwide, from 1990 to 2008, the incidence of T1D increased by 2.8-4% per year (7), similar to that observed in SEARCH (8). Moreover, SEARCH demonstrated an increased prevalence of T1D between 2001 and 2009 (9). On the other hand, a recent report from Finland, with the world's highest incidence, suggested that the increase in incidence from 2005-2011 has stabilized (10). Regarding T2D, although few longitudinal studies have been conducted, there is evidence that the increase in T2D in youth stems from the increased frequency of obesity in pediatric populations (11). Interestingly, data from SEARCH suggest that prevalence of T2D may not be increasing equally across race/ethnic groups (9). Thus, there is much to be gained in studying the continued trends in incidence and prevalence of T1 and T2D.

2. Objectives

This is the fourth phase of the ongoing SEARCH for Diabetes in Youth Study. SEARCH phases 1-3 were conducted in 2000-2005, 2005-2010, and 2010-2015, respectively, and included both a registry component and a cohort component. Study methods (12) and highlights of SEARCH study findings (13) have been published and protocols are available online (www.searchfordiabetes.org). Unlike SEARCH phases 1-3, SEARCH 4 is supported by two separate grants from different funding agencies, one for the Registry Study (CDC) and one for the Cohort Study (NIH/NIDDK).

2.1 REGISTRY STUDY OBJECTIVES

In response to RFA-DP-15-002, and with funds awarded by the CDC with contribution from the NIH/NIDDK, SEARCH 4 will continue to ascertain newly diagnosed incident diabetes cases throughout the study period and one additional prevalent cohort (index year 2017) for youth age < 20 years across five geographically dispersed study centers that encompass the racial/ethnic diversity of the United States. Surveillance is framed as a tiered approach, starting with the most broad based and cost efficient approach at the highest tier (tier 1) and becoming the most focused in tier 3, optimizing use of electronic health data.

Aim 1: TIER 1 SURVEILLANCE - To ascertain prevalent diabetes cases in calendar year 2017 among youth age < 20 years at diagnosis. Research Question 1.1 What is the prevalence of diabetes in 2017, overall and by age, sex, race/ethnicity, and diabetes type? Research Question 1.2 What are the temporal trends in T1D and T2D prevalence over the three prevalent cohorts (2001, 2009, and 2017) and how do trends differ by race/ethnicity, age, and sex?

Aim 2: TIER 2 SURVEILLANCE - To continue to ascertain newly diagnosed (incident 2013-2020) diabetes cases in youth age < 20 years. Research Question 2.1 What are the temporal trends in T1D and T2D incidence since 2002 in US youth and how do trends differ by race/ethnicity, age, and sex?

Aim 3: TIER 3 SURVEILLANCE - To further determine agreement between the etiological classification of diabetes type using biochemical markers and provider assessment, to describe selected clinical characteristics at diagnosis, and to establish an infrastructure that facilitates the development of more detailed ancillary studies by storing biological samples and preserving contact with potential study participants. Data is extracted from EHRs in all incident years and an in-person visit is planned for incident cohort year 2016, using a strategic sampling plan to minimize cost. Research Question 3.1 Is the proportion of youth with provider diagnosed T1D or T2D who have biochemical evidence of these respective diagnoses consistent over time? Evidence is based on diabetes etiologic types previously established and employed by SEARCH using diabetes autoantibodies (DAA) and the insulin sensitivity (IS) score. Research Question 3.2 Has the prevalence of DKA near the time of diagnosis decreased over time for youth with T1D or T2D?

Aim 4: OPERATIONAL EFFICIENCY - To optimize efficiency of SEARCH surveillance activities through targeted Development and Validation (D&V) Projects designed to utilize electronic health data to operationalize each of the three tiers of surveillance to the extent possible. Methods employ electronic algorithms and text mining/natural language processing with validation, incorporating data from administrative records, medical records including provider notes, pharmacy, and laboratory data. We will then evaluate these approaches with a goal of identifying a model for targeted expansion of the SEARCH Registry to non-SEARCH sites.

2.2 COHORT STUDY OBJECTIVES

In response to RFA DK-14-508, and with funds provided by Special Statuary Funds for T1 Diabetes Research, SEARCH 4 will continue to follow selected incident cohorts from the SEARCH registry. Incident cohorts of youth from 2002-2006, 2008 and 2012 were asked to participate in a baseline research visit where history, demographics, health-care related variables, clinical information and factors essential for the etiologic classification of diabetes type (diabetes related-autoantibodies and markers of insulin sensitivity) (14, 15) were collected near diagnosis. Participants were asked to return at 1, 2, and 5 years from baseline

for repeated measures in SEARCH phase 1 and 2. During SEARCH 3 (2010-2015), individuals who had participated in a baseline visit with at least five years duration were invited to participate in a cohort visit. At the close of SEARCH 3, 2780 individuals participated for a final response rate of 72% among eligible individuals. The current protocol (SEARCH 4 Cohort Study) will follow a subset of this cohort (as well as a subset of participants who completed a 2012 Registry In Person Visit) with another assessment to further assess risk factors, acute and chronic complications, as well as QOL-related outcomes and add measures of cardiac structure and function, neurocognitive outcomes, and social functioning and stress. We will also continue to assess mortality and causes of death.

Aim 1: Establish, compare and contrast the burden (prevalence, incidence, progression and clustering) of acute and chronic complications of diabetes, and explore the responsible risk factors and pathways among youth and young adults with T1D and T2D. We will measure key outcomes, including: retinopathy, nephropathy, cardiac autonomic (CAN) and peripheral neuropathy, arterial stiffness, cardiac damage, neurocognitive outcomes, as well as acute complications (hypoglycemia, diabetic ketoacidosis -DKA). We will explore a variety of risk factors and pathways, including: metabolic; inflammatory; vascular; behavioral; socio- economic; psycho-social and health-care factors. We hypothesize that: **1.1**: Youth with T2D have higher prevalence, incidence, faster rate of progression and different patterns of clustering of chronic complications, but lower burden of acute complications than youth with T1D, independent of age, sex, diabetes duration and race/ethnicity; **1.2:** The risk factor patterns associated with these outcomes are different in T2D vs. T1D.

Aim 2: Explore, compare and contrast processes of care (including barriers to care and quality of care- QOC) and their influence on QOL among youth with T1D and T2D, as they transition from pediatric to adult care. Measures to assess barriers include: consistent health insurance, out of pocket costs, continuity of care, employment, completion of education, finances, stressors from independence (school, work, marriage, children), social support, depression and neurocognitive factors. QOC variables include: frequency of visits with diabetes provider and receipt of screening for retinopathy, nephrology, neuropathy, foot exams, blood pressure and A1c. We hypothesize that**: 2.1:** Compared to youth with T1D, youth with T2D will a) have more and different barriers to care; b) benefit less from emerging treatment technologies; c) have worsening QOC and QOL as they transition from pediatric to adult care.

Aim 3: Conduct surveillance of mortality including cause of death in the SEARCH cohort. We hypothesize that**: 3 . 1 :** The frequency and causes of mortality in patients with youth-onset diabetes are different than among non-diabetic, age, sex and race/ethnicity comparable persons; **3.2**: Youth and young adults with T2D have higher mortality and different causes of death than youth with T1D, independent of age, sex, diabetes duration and race/ethnicity.

Aim 4: Maintain, supplement and promote access to the SEARCH Cohort repository for biological specimens to conduct scientifically and logistically appropriate ancillary studies.

3. Study Population

3.1. STUDY SITES

The five clinical centers that participated in SEARCH 3 will continue their participation in SEARCH 4. These sites are based in Ohio, Colorado, Washington, South Carolina, and California. Four SEARCH centers (Ohio, Colorado, Washington, and South Carolina) are geographically based - that is, newly diagnosed diabetes cases are identified from a geographically defined population. One SEARCH center (California) is membership-based - that is, newly diagnosed diabetes cases are identified from the membership of the participating health plan. Each of the five centers participates in both the Registry and the Cohort Studies.

1. Ohio - Cases ascertained from Cincinnati and the 8 surrounding counties; oversight, recruitment and clinic visits provided by Children's Hospital Medical Center.
2. Colorado - Cases ascertained from the state of Colorado and members of the Navajo Indian tribe in AZ, UT, or NM residing on the Navajo Nation reservation; oversight, recruitment and clinic visits provided by University of Colorado, Denver.
3. Washington - Cases ascertained from Seattle and Tacoma and the 5 surrounding counties; oversight, recruitment and clinic visits provided by Seattle Children's Research Institute.
4. Carolinas - Cases ascertained from the state of South Carolina with oversight provided by the University of North Carolina at Chapel Hill. Sub-centers are located at three locations in SC (Charleston, Greenville, Columbia) to assist with recruitment and clinic visits.
5. California - Cases ascertained from Kaiser Permanente Southern California Health Care Plan membership (other than San Diego) with oversight, recruitment and clinical visits provided by the same.

3.2. OTHER SITES

The Coordinating Center (CC) is located at the Wake Forest School of Medicine in Winston-Salem, NC, and has served as the CC for all phases of SEARCH. The laboratories and reading centers, listed below, are supervised by and operate as subcontracts to the CC.

1. Central Laboratory- Northwest Lipid Metabolism and Diabetes Research Laboratories, University of Washington
2. Neuropathy Reading Center, University of Michigan
3. Ocular Epidemiology Reading Center, University of Wisconsin-Madison

4. Cardiovascular Reading Center, Cincinnati Children's Hospital Medical Center

3.3. STUDY POPULATION AND ELIGIBILITY

3.3.1. *The SEARCH Registry Study*

Over the three phases of SEARCH, investigators have registered more than 25,000 cases of youth with diabetes, including completed incident cohorts from 2002-2012, prevalent cohorts in 2001 and 2009, and ongoing efforts for registration of incident 2013-2015 cohorts. During SEARCH Phase 4, incident 2013-2017 cohorts will be completed, and incident 2018-2020 will be initiated but not completed (Table 1).

Table 1. Surveillance activities during SEARCH Phase 4

Phase 4 Period	Case ascertainment of youth diagnosed in:	Incident year to be closed (30 months after the end of the incident year):
Yr1: Oct 2015-Sept 2016	2013, 2014, 2015, 2016*	2013
Yr2: Oct 2016-Sept 2017	2014, 2015, 2016*, 2017	2014
Yr3: Oct 2017-Sept 2018	2015, 2016*, 2017, 2018+	2015
Yr4: Oct 2018-Sept 2019	2016*, 2017, 2018+, 2019+	2016
Yr5: Oct 2019-Sept 2020	2017, 2018+, 2019+, 2020+	2017

+ Registration for incident years 2018, 2019, and 2020 will not be completed during SEARCH 4. We will begin registering these cases in anticipation of future funding to fully register these incident years. Note that incident years 2013- 2015 initially began registration during SEARCH 3.
*In-person-visits (IPV) will be conducted on 2016 incident cases. Yr=Year.

Registry Aims 1 and 2. Centers in SEARCH Phase 4 will continue to conduct population-based ascertainment of cases of diabetes in youth less than 20 years of age for incident years 2013 through 2020, using methods consistent with those employed in SEARCH Phases 1-3. Prevalent cases will be obtained in index year 2017. Briefly, cases are ascertained primarily though networks of pediatric endocrinologists, with pediatric diabetes databases, electronic health records from participating inpatient and outpatient settings, hospitals, and other health care organizations being queried to identify the remainder of the cases. Cases will be validated based on physician reports, medical records reviews or self-reports of a physician diagnosis of diabetes based on an established set of criteria. Further eligibility is defined by the following: 1) children/youth who, in addition to having an onset of physician-diagnosed diabetes in the index year, are also are < 20 years of age on December 31 of the index year; 2) are resident of the population defined for geographically-based centers at any time during the index year, or a member of the participating health plan for the membership-based center at diagnosis, and 3) are not active duty military personnel or institutionalized. Young women who develop gestational diabetes mellitus (GDM) but who are not diagnosed with diabetes when not pregnant are not eligible. Sites are provided a 30 month window after the close of the incident year to identify all potential cases. For example, for incident

year 2016, the window closes 6/30/2019 (Table 1). A total of 13,440 incident cases are expected to be registered during SEARCH Phase 4 (Table 2).

The prevalence study for 2017 will attempt to identify and validate all unique, eligible cases of diabetes in youth less than 20 years who are residing in or are members of the SEARCH geographic areas and health plans in 2017. Previous prevalence studies have been conducted in SEARCH in 2001 and 2009. A total of 1004 new prevalent cases not previously identified through the incidence study are expected to be registered (Table 2). Completeness of case ascertainment will continue to be monitored via capture-recapture analyses, as described in detail on the SEARCH website (16).

Table 2: Estimated Number of Registered Cases (Incident and Prevalent) and IPV, Overall and By Site, SEARCH 4

	Carolinas	Ohio	Colorado	California	Washington	Total
2013 Incident*	378	179	422	265	284	1511
2014 Incident*	389	184	435	273	292	1557
2015 Incident*	401	189	448	281	301	1603
2016 Incident	413	195	461	290	310	1652
2017 Incident	425	201	475	299	320	1701
2018 Incident	438	207	489	307	329	1752
2019 Incident	451	213	504	317	339	1805
2020 Incident	465	220	519	326	349	1859
Total Incident	**3360**	**1588**	**3753**	**2358**	**2524**	**13,440**
2017 Prevalent**	57	112	86	508	241	1004
Total Cases (I + P)	**3417**	**1700**	**3839**	**2866**	**2765**	**14,444**
Total IPV* **	**207**	**81**	**220**	**171**	**153**	**832**

* Total number of cases we expect to register for these incident years, including those registered under the SEARCH 3 protocol. **Excludes incident 2017 cases, and all previously registered incident and prevalent cases included in the 2017 prevalent sample. *** IPV for 2016 incident cases.

The calculation of incidence and prevalence rates require information on the population at risk. Race-bridged post-censal estimates of the July 1 resident US population, released yearly by the National Center for Health Statistics, are used as the denominators for the geographic sites. Each file contains population estimates for each US county by single year of age, bridged-race, sex, and Hispanic origin. Active duty military are excluded. The membership site (California) uses July 1 health plan enrollment data by single year of age and sex as the denominator. Addresses for each of the members are geocoded and census block level data are used as a source of race/ethnicity (17). The Indian Health Service user population for eligible service units on the Navajo Nation, defined as

persons age < 20 years with one or more visits in the past 3 years (including the index year) is used to estimate denominators for this Colorado sub-site.

Registry Aim 3. A sample of cases diagnosed in 2016 will be invited for an in-person visit (IPV). Cases eligible for the IPV will include all cases diagnosed during 2016 who are of minority race/ethnicity, those with a provider diagnosis of T2D, and 25% of non-Hispanic white youth with a provider diagnosis of T1D, randomly selected for invitation. This sampling plan will yield approximately 832 IPVs (Table 2).

3.3.2. *The SEARCH Cohort Study*

Cohort Study Follow-Up (Cohort Aims 1 and 2). A subset of SEARCH 3 Cohort (C_1) and SEARCH 3 Registry (R_1) participants will be invited for a SEARCH 4 in-person visit. The eligible group will include all SEARCH 3 (C_1 and R_1) participants with T2D, all minority youth with T1D, and a random sample of NHW youth with T1D. The Coordinating Center (CC) will provide a list of randomly selected NHW youth with T1D to be invited for participation such that all participants will be 10 years or older, have at least 3 years of time elapsed since their SEARCH 3 (C_1 or R_1) visit and have at least 5 years of duration of diabetes at the time of their planned SEARCH 4 IPV. NHW sampling is performed since based on the limited available budget it was determined that there was minimal gain in statistical power to invite all T1 NHW youth for a return visit and that all proposed analyses could be addressed with the random sample of NHW T1. Table 3 shows the total number of participants expected to complete a SEARCH 4 Cohort visit (N~1,846) based on the proposed sampling. These estimates are based on an expected 75% response rate. In addition, the SEARCH 4 IPV will include a sample of 500 participants to be identified by the CC to have cardiac echocardiogram measurements taken. This sample will include 250 T1D and 250 T2D with representation from all five clinical sites and have racial/ethnic diversity.

The remainder of SEARCH 3 (C_1) participants will form the survey-only group with no IPV in SEARCH 4. The survey-only option will also be offered to individuals who are eligible but refuse participation in the IPV. The survey-only group will be asked to complete questionnaires by mail, phone or internet. Survey data will be combined from the IPV and survey-only participants (at least2,546) to address Aim 2.

Table 3. Number of Expected participants for each component of the SEARCH 4 Cohort Study, by Clinical Site			
Site	In-Person Visit	Echocardiogram	(Estimated) Survey Only
Carolinas	434	140	150
Ohio	298	130	150
Colorado	504	130	225
California	352	0	25
Washington	258	50	150
Total	1846	450	700

Mortality follow up (Cohort Aim 3): All incident cases identified by the Registry study during calendar years 2002-2015 will be included in the mortality follow-up through 12/31/17 using the National Death Index (NDI) (18). This is the second mortality assessment, with the initial one done for incident cases identified during 2002-2008 and followed through 12/31/10. Mortality status will be obtained by matching with the NDI as soon as the NDI has complete data for 2017, usually ~18 months from the close of the time period. Conservatively, we estimate that there will be 89 additional deaths for a total of 130, using mortality rates from the prior period. This will allow us to examine cause-specific deaths in selected subgroups.

3.4. INFORMED CONSENT

Consent for the SEARCH Study is handled through three important mechanisms under the supervision of local IRB(s). Individual differences exist based on requirements of local IRBs. In general: 1) initial data for the Registry Study is collected (without participant contact) under a HIPAA waiver; 2) completion of surveys (either by mail or web-based/online) is covered by a waiver of documentation of consent [aka, implied consent] according to local IRB requirements in the Registry and Cohort Studies; and 3) written informed consent is obtained prior to all IPVs in the Registry and Cohort Studies.

As in previous phases of SEARCH, the initial data collection in the Registry Study (case ascertainment) is covered by a HIPAA waiver. That is, identification of all new cases of diabetes in a defined geographic area or health plan does not require that registered cases provide written or implied informed consent; HIPAA requirements are waived.

Mailed and/or web-based online surveys are utilized in both the Registry Study and Cohort Study. In this case, consent is implied with completion of the surveys. In the Registry Study,

potentially eligible cases are mailed (or emailed with internet link) an introductory letter that gives a brief description of the research study along with the Initial Participant Survey (IPS). For individuals who are less than 18 years of age, the introductory letter is mailed to a parent or guardian. If the completed IPS is not returned and the participant does not refuse after receiving the introductory letter, a member of the local research team may call the individual or the parent to complete the IPS. Again, consent requirements for completion of the IPS are governed by the local IRB. In the Cohort Study, a subset will be asked to participate in the survey-only group, for which surveys will be mailed (or emailed with internet link) for completion at home.

Written informed consent is obtained for all individuals/parents who agree to participate in the Registry Study IPV as well as for the Cohort Study IPV in accordance with local IRB requirements. If the participant is less than 18 years of age, the parent or guardian must give written informed consent prior to the initiation of any study procedures or data collection, according to the requirements of the local IRB. Written assent of participants who are less than 18 years of age is also governed by the requirements of the local IRB. If the participant is 18 years of age or older, the participant must give written informed consent. Copies of completed consent forms are maintained in the participant's local research record.

There are three optional components to the written informed consent: storage of serum/plasma/urine and DNA/miRNA; transfer of data and samples to the NIDDK Repository; and sharing of data and genetic information with dbGaP (**d**ata**b**ase of **G**enotypes **a**nd **P**henotypes). In each case, participants or their parent must indicate in writing whether or not they are providing consent for these optional components. The NIDDK Central Repository is a research resource supported by the National Institutes of Health. At the end of the SEARCH, de-identified research data and samples of blood and urine will be provided to the Repository for participants who have consented to this component. For all optional components, participants may choose to participate in SEARCH but not provide consent to participate in these components.

3.5. RECRUITMENT & RETENTION

The SEARCH Registry Study sites continue to employ a wide variety of methods shown to be highly effective at recruiting study participants for the IPS and IPV. Recruitment strategies have included meeting the family at a medical appointment to introduce the study; mailing study brochures and other informational letters; posting study materials in clinics; enlisting the encouragement by diabetes care providers; emailing, texting, or using social media such as Facebook to contact potential participants; phoning participants to complete the surveys and/or schedule a visit; offering online surveys; and one or more reminder calls prior to the scheduled visit. Participation in the IPV is facilitated by flexible weekday appointments, as well as Saturdays, satellite clinics, and home visits. Sites offer to pair research visits with clinical appointments when possible; provide transportation and/or lodging; and generally assist participants with removing barriers to study participation.

Study participants are offered remuneration that is appropriate for the length and burden of the study visit. Participants and their providers receive the clinically-relevant research laboratory test results, which may assist with their clinical care. To retain Registry Study participants for future studies and to share study progress, we utilize traditional, proven, retention strategies including: birthday cards, study newsletters, updating contact information annually, and utilizing internet-based search systems to locate individuals lost to follow-up.

Similarly, the SEARCH Cohort Study has maintained outstanding participant retention throughout its history. We continue to employ traditional, proven, retention strategies as described above. We also offer flexible study date appointments including home visits, offer assistance with transportation, mail pre-visit instructions, one or more reminder calls prior to the scheduled visit, provide acknowledgement of participation, and provide participant remunerations that are appropriate for the length and the respondent burden of the proposed study visit. Investigators and study personnel also continue to solicit the support of diabetes providers to encourage on-going study participation. Communications with providers include letters, e-mail messages, telephone calls, newsletters, individual discussions, and group presentations of study goals and preliminary results.

4. Study Measurements

For SEARCH phases 1-3, all clinical sites have operated under a common protocol. This approach is followed in SEARCH 4 Registry and Cohort Studies as well. That is, data from each site is obtained, managed, and protected according to a standard study protocol that has been developed and vetted by the Steering Committee and approved by all participating IRBs and by the NIDDK Observational Studies Monitoring Board (OSMB). Clinic sites use a standard informed consent template, modified as needed by local IRB requirements. All clinic staff are trained and certified, operate under a single Manual of Procedures (MOP), and follow a standard set of data collection procedures. Clinic staff participate in both central and local training as needed. Clinical Center investigators and staff participate in ongoing working groups and established study committees to ensure that identical procedures are followed at each site for the purpose of recruitment, retention, and ensuring the highest quality of study data.

4.1. MEASUREMENTS - REGISTRY STUDY

Centers in the SEARCH Registry Study continue to conduct population-based ascertainment of cases of diabetes in youth < 20 years of age using methods consistent with those employed in SEARCH 1-3. This involves identification, case validation, confirmation of eligibility, deduplication, and registration of cases centrally with the SEARCH Coordinating Center. There are three aspects of data collection in the Registry Study: 1) data obtained from all potential registered cases; 2) the Initial Participant Survey (IPS); and 3) data obtained during an in-person visit (IPV) on a subset of the 2016 registered cases.

Collection of Data on all Registered Cases: A minimum amount of demographic and clinical information is needed for all cases in order to calculate population-based incidence

rates and prevalence of diabetes mellitus by age, sex, diabetes type and race/ethnicity. The primary source of this information is the medical record except for race and Hispanic ethnicity, which, when obtained by self-report using the IPS, supersedes the report via medical record. Study staff abstract information from the medical record for the period from diabetes diagnosis to six months after this date to obtain the following information: 1) date of birth, 2) sex, 3) race/ethnicity, 4) diagnosis date, 5) zip code at diagnosis, 6) county and state of residence at diagnosis, 7) diabetes type at the time of diagnosis and the diabetes type reported closest to 6 months, 8) whether diabetes autoantibodies were measured up to 6 months after diagnosis [GAD/GAA, IA2/ICA512, ICA, IAA, and ZnT8], 9) height, 10) weight (closest to diagnosis), 11) whether the participant ever used insulin, 12) whether insulin was discontinued, 13) presence of acanthosis nigricans, and 14) whether DKA was noted (with dates, bicarbonate, pH, and glucose values). For potential cases not eligible for registration, minimal demographic data are maintained in order to facilitate validation and de-duplication of local cases.

Initial Participant Survey (IPS): All registered cases are invited to complete the IPS. The IPS is used to: a) verify case eligibility (e.g., residence in the year of diagnosis); b) obtain self-reported race/ethnicity and selected clinical and demographic information; and c) introduce participants to SEARCH to facilitate future studies. The IPS queries symptoms at presentation, potential secondary causes of diabetes, use of insulin and other medications, diabetes treatment history, height and weight, family structure, usual language spoken, type of health insurance, usual provider for diabetes care, highest parental education, household income, nativity of person with diabetes and their parents, and contact information. All registered cases are eligible to complete the IPS online, by mail, or by interviewer administration by telephone or in person.

In-Person Visit (IPV): A sample of registered cases diagnosed in 2016 will be invited to an IPV. The IPV enables an analysis comparing agreement between provider assigned diabetes type compared to SEARCH etiologic type in order to interpret the potential meaning of trends over time according to provider type, and to enable statistical adjustment for differences in agreement over time. The IPV, lasting approximately 60 minutes, includes collection of fasting blood (3 TBSP) and urine samples, a brief physical examination, and a medication inventory, all conducted under SEARCH standardized protocols and described below for the Cohort Study. Measurements made to inform dimensions of diabetes type include diabetes autoimmunity (GAD65, IA-2, and ZnT8 antibodies), and the SEARCH validated insulin sensitivity index (waist circumference, HbA1c, and triglycerides) (15). We will also measure markers of kidney function (albumin and creatinine from a first morning void, cystatin-c and serum creatinine), the latter two measures pending availability of funds. To facilitate work that requires additional funding in the future, we will store plasma, serum, DNA, and urine. Specimens are processed locally and shipped within 24 hours to the central

laboratory. Diabetes autoantibodies are measured by standardized protocol and a common serum calibrator developed by an NIDDK- sponsored standardization group.

4.2. MEASUREMENTS - COHORT STUDY

The study visit for the cohort study participants is expected to take approximately four hours and includes physical measures and questionnaires. A parent/guardian is required to attend if the individual with diabetes is < 18 years old. Most of the measures obtained during the SEARCH 4 visit are the same as to those obtained in previous visits. New measures, particularly cardiac echocardiography and neurocognitive testing, are noted in Table 4, along with data obtained at C_0 (baseline visit), intermediate visits (12, 24, 60 months), and the SEARCH 3 Cohort visit (C_1).

Surveys. SEARCH has included surveys in multiple domains over time. Surveys in SEARCH 4 include:

a) Health history including pregnancy history of women;

b) Treatment including all prescribed medications, insulin regimens and glucose monitoring devices (19,20);

c) Behavioral factors including diet (21), physical activity (22), TV and computer use (23), smoking (24), and substance use (25);

d) Psychosocial factors using CES-D scales (26-28), the PROMIS and PHQ9 depression and anxiety screening tools (29), Hypoglycemia Fear Survey (HFS-C,P) (30-33), the updated Diabetes Responsibility and Conflict Scale to assess diabetes-specific family conflict (34); Stigma and Discrimination. Diabetes self-care is assessed with the Diabetes Self-Management Questionnaire (DSMQ) (35);

e) Socio-cultural factors including household and per capita income, family structure, preferred language, migration status, parental and participant attained education, participant employment status, household food security;

f) Processes of care including type and frequency of utilizing health care providers, processes of diabetes self-management training, and recent hospitalizations (36);

g) Quality of care based on ADA guidelines for pediatric diabetes care in terms of testing frequency for HbA1c, blood pressure, lipids, urine albumin, retinopathy screening, and foot checks (37). Receipt of services is measured by self-report by parents (participant age <18) or adult participants (age ≥ 18 yrs.);

h) Quality of life using the Pediatric Quality of Life Inventory (PedsQL) (38-40) with age-specific and parental scales for participants < 18 years and validated scales for young adults 18-25 and over 26;

i) Barriers to care via items from the Consumer Assessment of Healthcare Providers and Systems survey (CAHPS 3.0) Supplemental Item Set for Children with Chronic

Conditions. Additional information about continuity of health insurance, continuity of care, cost-related non-adherence and financial burden is collected using the following surveys, adapted for youth and young adults: Medical Expenditure Panel Survey (MEPS) [Agency for Healthcare Research and Quality (AHRQ)]; Perceived Financial Burden of Diabetes and Cost-related Medication Non-adherence (41);

j) Transition to adult care: Specific questions about processes of care, motivations, satisfaction with, and preparation for transition from pediatric to adult care, adapted from validated measures that have been developed to assess patients' perceptions of other kinds of care transition, such as the Care Transition Measure (42). We will also measure care transition planning by adapting items from the National Survey of Children with Special Health Care Needs.

Physical exam. Standardized anthropometry methods include height, weight, waist circumference (using NHANES and WHO protocols); systolic and diastolic blood pressure; and evaluation for acanthosis nigricans.

Laboratory parameters. Fasting blood (3 TBSP) and first morning urine are collected following standard protocols. Blood and urine laboratory parameters continue to be measured using established protocols at the Northwest Lipid Research Laboratory. Samples are shipped from clinical centers to the central laboratory. Results are sent from the laboratory to the CC through established secure protocols.

Cardiac echocardiography. Measures of cardiac structure and function are obtained using cardiac echocardiography in a subgroup of Cohort Study participants. Measures include two dimensional (2-D) directed M-mode echo images to determine left ventricular mass (LVM), left atrial size and relative wall thickness, as well as shortening fraction, LV strain and diastolic function. The primary outcome measure is LVM determined by 2-D guided M-mode echo at end diastole (43, 44) using the autopsy corrected equation of Devereux (45). Echocardiograms are read on a Digiview instrument and strain is read on a Tomtec instrument. Digital images recorded on CDs identified only by participant ID number are sent to Cardiovascular Reading Center.

Retinal Photography. We will continue obtaining retinal images using Canon CR-1 Mark II fundus cameras. Consistent with NHANES protocol (46), two 45-degree images are taken of each eye: one centered on the optic nerve and the other on the fovea. The Ocular Epidemiology Reading Center at the University of Wisconsin-Madison (47) will grade the images for presence and severity of diabetic retinopathy (DR), macular edema and will make measurements of retinal vessel calibers. After grading the retinal images from the 2nd retinal visit, a separate longitudinal review will be conducted to confirm progression/regression status of diabetic retinopathy or macular edema severity.

Table 4. Data Collected on Cohort Study Participants				
Variables	**Baseline Visit (C_0)**	**12, 24, 60 months**	**SEARCH 3 Visit (C_1)**	**SEARCH 4 Visit (C_2)**
Surveys:				
Demographics: Sex, Race/ethnicity, Parental age	X			
Employment, education: parent or youth > 18 years		X	X	X
Medical Record: Diabetes type, date of diagnosis	X			
Health History: Birth date &weight, age at onset	X			
Pubertal status, co-morbidities; family history	X	X	X	X
Pregnancy outcomes in females				X
Medication: Diabetes & related conditions	X	X	X	X
Behavioral: Diet, physical activity, alcohol use	X	X	X	X
Marijuana, other substance use				X
Processes of care/quality of care			X	X
Health care costs			X	X
Psychosocial: Depression (CES-D)	X	X	X	X
Family conflict; fear of hypoglycemia			X	X
Transitions of care			X	X
Food security and assistance				X
Stressors; work ability index; stigma/discrimination				X
Physical exam: BMI, waist , blood pressure, acanthosis	X	X	X	X
Laboratory measures (blood):				
Autoantibodies	X	X	X	
Fasting glucose, cystatin C, serum creatinine, fasting C-peptide, lipid profile, inflammatory markers (CRP, IL6), A1c, AGE (CML), DNA, miRNA extraction	X	X	X	X
URINE: albumin, creatinine (spot)	X	X	X	
URINE: albumin, creatinine (first morning void)			X	X
Stored Samples: DNA, miRNA, serum, plasma, urine	X	X	X	X
Outcome(s):				
Cardiovascular: Arterial stiffness (PWV, AiX)	X		X	X
Cardiac echocardiography: LV mass, systolic & diastolic function				X
Neuropathy: heart rate variability; peripheral neuropathy		X (pilot)	X	X
Retinopathy Retinal photos, vessel caliber		X (pilot)	X	X
Nephropathy: Albuminuria	X	X	X	X
Cystatin C		X	X	X
Neurocognitive tests: NIH Toolbox.				X
Acute complications: DKA, hypoglycemia	X	X	X	X
Quality of life (Peds QL3.2 Diabetes Module)	X	X	X	X
Mortality surveillance (NDI)	X	X	X	X

Measures of Kidney Function

Urine albumin:creatinine ratio: We will collect first morning void (FMV) urine samples for storage and calculation of urine albumin:creatinine ratio (UACR). We will request a second FMV sample in the case of a positive urine for microalbuminuria (UACR $\geq$ 30μg/mg), or if a urine sample is rejected due to possible contamination (positive leukocyte esterase and/or nitrite). In previous phases we had collected random urine samples (with the addition of the FMV at SEARCH 3).

Estimated glomerular filtration rate: Equations with the most accurate and precise estimation of glomerular filtration rate (GFR), utilize both serum creatinine and cystatin C (48, 49). Both tests have been measured in SEARCH 1-3 and continue to be measured in SEARCH 4. Different equations are currently used in children versus adults, and on the expected range of GFR (hyperfiltration versus normal GFR versus low GFR) (48-51). The natural history of eGFR in diabetic kidney disease can be heterogeneous and so we will also investigate the optimal equations for use in children versus adults and at different spectrums of GFR.

Measures of Neuropathy

Peripheral Neuropathy: The Michigan Neuropathy Screening Instrument (MNSI) will be used to screen for the presence of diabetic neuropathy. It consists of 15 self-administered questions on foot sensation including pain, numbness and temperature sensitivity. The second part of the MNSI is a brief physical examination involving 1) inspection of the feet for deformities, dry skin, hair or nail abnormalities, callous or infection, 2) semi-quantitative assessment of vibration sensation at the dorsum of the great toe, 3) grading of ankle reflexes and 4) monofilament testing. Patients screening positive on the clinical portion of the MNSI (greater than 2 points on a 10 point scale) are considered neuropathic.

Cardiac Autonomic Neuropathy: Heart rate variability (HRV) analysis allows us to assess the autonomic nervous system by examining sympathetic balance, which raises heart rate and blood pressure and causes vasoconstriction, and the parasympathetic balance which has opposite effects (52,53). Assessments use a SphygmoCor SCOR-CPV device (AtCor Medical, Australia) as performed previously (54).

Arterial stiffness. Pulse wave velocity (PWV) is measured using the SphygmoCor (55). The average of 3 ECG R-wave gated arterial waveforms are recorded from the carotid and then the femoral arteries. Augmentation index (Aix*)* is measured with the same device (56).

Neurocognitive tests. Neurocognitive measures are computer administered utilizing the NIH Toolbox (57); domains include attention, verbal skill, working memory, mental flexibility, episodic memory, speed of processing, and response inhibition. These areas were chosen to reflect both more generalized (depressed psychomotor speed) and distinct areas of deficit (memory, attention, and mental flexibility). Receptive language vocabulary is used as a proxy for educational attainment/premorbid functioning.

Acute complications. Acute complications studied are severe hypoglycemia and diabetic ketoacidosis (DKA). Severe hypoglycemia is defined as a hypoglycemia event requiring assistance of another person (58). For DKA, occurrence is recorded as emergency department visit or hospitalization. This aligns well with prior publications on acute complications, and data frequently recorded in patient surveys and medical records (59).

Mortality Surveillance. All centers will systematically identify deaths that occur between the date of diagnosis and December 31, 2017 among youth in the 2002-2015 incident cohorts. The National Death Index will serve as the primary data source, plus individual case reports of deaths made to the study team during the course of the study.

5. Development and Validation Projects

The recent implementation of robust EHR systems throughout the US provides opportunities to substantially enhance the efficiency of surveillance and to pilot expansion of the SEARCH Registry beyond the currently funded sites. SEARCH 4 will attempt to optimize efficiency of SEARCH surveillance activities through targeted Development and Validation (D&V) Projects designed to utilize electronic health data to operationalize each of the three tiers of surveillance. Methods will employ electronic algorithms and text mining/natural language processing with validation, incorporating data from administrative records, medical records including provider notes, pharmacy, and laboratory data. We will then evaluate these approaches with a goal of identifying a model for targeted expansion of the SEARCH Registry to non-SEARCH sites.

Four of the five SEARCH centers are part of networks funded by Patient-Centered Outcomes Research Institute (PCORI)'s multi-institutional clinical data research networks (CDRN) formed in 2014. Three of the centers, OH, WA, and CO, are part of "A National Pediatric Learning Health System Network" (PEDSnet). The fourth clinical center, CA, is part of "Kaiser Permanente & Strategic Partners Patient Outcomes Research to Advance Learning" (PORTAL) Network. In South Carolina, a new entity, Health Sciences South Carolina (HSSC), has been establishing a data warehouse to bring together EHR data from at least four of the six major provider systems from which cases are ascertained for the Carolina site.

This work will follow a three-step process to include development, validation and implementation. First, new approaches will be developed and initially validated through the D&V Projects in limited locations. For each approach, the established SEARCH processes for case ascertainment, case validation, and determination of diagnosis date, diabetes type, and key clinical and demographic data will be considered the "gold standard" against which new approaches will be compared. Second, approaches that meet appropriate initial validation criteria will be further refined and validated at additional SEARCH centers. Third, implementation as part of ongoing SEARCH Registry work will occur only after pre-determined metrics (e.g., sensitivity, specificity, PPV) are demonstrated for each EHR system in which the approach is to be implemented.

5.1. PROJECT #1: CASE ASCERTAINMENT BY DIABETES TYPE

The goal of D&V Project 1 is to maximize the automation of ascertainment of diabetes cases, overall and by diabetes type, by applying case identification algorithms and text analytics/natural language processing (NLP). This work is critical to Tier 1 (Prevalence) efforts. Two approaches will be employed building on previously described SEARCH work (17, 60) one based on algorithms using EHR and administrative data, the second using natural language programming (NLP) to extract and analyze text. We will attempt to replicate our previous work using EHR-based algorithms as developed in the Carolinas site to determine if these algorithms perform in a similar manner in an integrated health care system in the California site. Additionally, we will apply case identification algorithms to the PEDSnet data for the three SEARCH sites, and compare results to those using the SEARCH gold standard methods using metrics described above. Regarding the text analytics approaches, we propose to apply work as developed in the Carolinas site, including re-training of the machine learning models, to clinical notes from at least one provider from each of the five SEARCH centers, to include the three SEARCH centers that are part of PEDSnet, the CA center utilizing the KPSC data systems and Carolina working with HSSC.

5.2. PROJECT #2: DETERMINATION OF DIAGNOSIS DATE

The critical information element that distinguishes Tier 1 (Prevalence) surveillance from Tier 2 (Incidence) surveillance is date of diagnosis, which generally is not available as a structured data element that can be easily extracted from the EHR. Thus, the current literature that describes various EHR-based algorithms for case identification is generally applicable only to prevalence. The goal of D&V Project # 2 is to use electronic ascertainment methods to determine diabetes diagnosis date with an expectation that at least 95% of the estimated dates will identify the correct calendar year of diagnosis. As in D&V Project #1, two approaches will be employed: 1) use of EHR-based algorithms applied to structured data; 2) use of text analytics applied to unstructured data.

All five SEARCH centers will participate in the EHR-based algorithm work. Regarding use of text analytics, the Carolinas site will continue to refine the machine learning model with the goal of attaining at least 90% accuracy for year of diagnosis. Once optimal algorithms are established for T1 and T2D, we will expand the effort to build on the text analytic work being done for D&V Project # 1 at each of the five SEARCH centers.

5.3. PROJECT #3: AUTOMATION OF CARE AND CLINICAL DATA COLLECTION

The third project will focus on whether the collection of core and selected demographic and clinical information can be automated by directly importing information from the EHR and other clinical and administrative data systems. In addition to data obtained as part of case ascertainment by diabetes type (Project #1, Tier 1) and date of diagnosis (Project # 2, Tier 2), additional information of importance includes race and ethnicity, measurement of diabetes autoantibodies, clinical information including laboratory values related to diabetic

ketoacidosis, diabetes medications, etc. The evaluation of the data capture procedures will consider both completeness as well as the accuracy of the information extracted compared to manual extraction of core data.

5.4. PROJECT #4: EXPANSION OF SURVEILLANCE TO ADDITIONAL AMERICAN INDIAN (AI) TRIBES

Since its inception in 2000, SEARCH has been conducting surveillance of youth onset diabetes in AI tribes under the direction of the Colorado site. These results indicate that AI youth have the highest incidence and prevalence of T2D of any major race/ethnic group (61, 6, 9). Unfortunately, the AI population under surveillance is the smallest of the major race/ethnic groups (~95,000 youth), and results in less than ~40 incident cases per year across all sites. The goal of this project is to develop and validate an algorithm that may be used to identify AI and possibly Alaskan Native (AN) youth with diabetes using data extracts from *existing* electronic health records (EHR) of the Indian Health Service (IHS). For the proposed pilot project, the SEARCH Colorado site will partner with the Center for AI AN Health (CAIANH), both located within the Colorado School of Public Health.

An algorithm will be developed to identify AI youth aged < 20 years with diabetes using IHS data that includes diagnostic codes, provider and service type information and dispensed medications. The algorithm will be developed from the IHS data for the Chinle and Tuba City Service Units on the Navajo Reservation and validated by comparison to the Navajo SEARCH registry (gold standard) for the same service units and will result in metrics for sensitivity, specificity, positive and negative predictive value. Next, and coordinated through CAIANH, the best algorithm developed in Phase 1 will be used to identify AI youth with diabetes in another IHS Service Unit, from a different tribe, using the same IHS National Data Warehouse.

5.5. COST OF THE REGISTRY

Efforts to enhance efficiency are driven, in part, by potential cost savings. In SEARCH 4, we will conduct a prospective assessment in the two parallel aspects of SEARCH: conventional case ascertainment (Aims 1-3) and the D&V projects (Aim 4) in order to estimate the cost of case registration. The primary means of data collection will be the time diary in which staff members will be asked to record all SEARCH activities over a typical work week, periodically over time, with attention to infrequent tasks (e.g., those conducted monthly). The types of activities to be tracked include: managing people, clerical (mailing, logging, filing), training of staff, IT support, meetings, locating/reviewing/entering data, identification/validation/deduplication/registration, analyzing/generating reports, and local travel. For the conventional case ascertainment, diaries will be completed one week each quarter, over a period of one year. For the D&V projects, diaries will be completed more frequently, depending upon the length of the project. Actual salary and benefit rates will be applied to the time elements. A count of the number and type of cases registered during the

time period will be obtained from the SEARCH registration database. Specifically, we will evaluate time and cost for the Registry as it is currently conducted, then will systematically model the incremental differences that can be attributed to approaches determined to be valid from the D&V Projects.

6. Statistical Considerations

6.1. REGISTRY STUDY - STATISTICAL CONSIDERATIONS

6.1.1. *Aim 1: Detectable Differences in Prevalence*

The third assessment of the prevalence of diabetes in youth is scheduled for 2017. Similar to previous work, prevalence will be expressed as the number cases with T1 or T2D per 1,000 youth pooled across all SEARCH sites. Prevalence estimates will be derived by sex, age and by race/ethnicity groups within each diabetes type. Trends in prevalence will be assessed by comparing the 2017 estimates to those observed in 2009 and 2001. Poisson regression models will be fitted to incorporate results from all 3 surveys. Standard errors associated with the estimated change in prevalence rates between any 2 time points will be computed using a 2 sided skew-corrected inverted score tests for binomial distribution. Standard error for the trends in prevalence estimates will be derived from the Poisson regression model. This model will also be used to generate adjusted prevalence where adjustment will be made for race/ethnicity, age and sex. Our power calculation suggests that we are well-powered to detect changes in prevalence by diabetes type, and across race/ethnic group within each diabetes type. For example, we will have at least 90% power to detect a rate of change of 4.1% in NHW youth with T1D, and a rate of change of 19.1% in NHB youths with T2D for the period between 2009 and 2017.

6.1.2. *Aim 2: Detectable Differences in Incidence*

A similar approach will be taken to estimate the incidence rates of diabetes by type, race/ethnicity, sex and age. Incidence rates will be estimated as the number of diagnosed cases across all sites divided by the total number of individuals who are at risk across these sites. The incidence rates will be expressed in terms of the number of cases diagnosed per year per 100,000 individuals. Adjusted incidence rates will also be provided by race/ethnicity, sex and age. SEARCH 4 will add 5 additional years of incidence data taking the current time series from 12 to 17 years of data, thereby providing improved power to detect changes in incidence rate during this period for various subgroups. Based on our power calculations (see Table 5), SEARCH 4 will have 90% power to detect changes as small as 1.04% in NHW females with T1D and 2.1% in NHB females with T2D. However, we will have limited power to detect changes in API and AI youth. This is the rationale for D&V Project # 4 in which we propose an approach to develop a model for extension of the SEARCH Registry to increase inclusion of population subgroups for which our sample size is limited.

Table 5.: Detectable rate of change in incidence rate by diabetes type, sex and race/ethnicity, and power

Race	Sex	Type 1		Type 2	
		90%	80%	90%	80%
All	All	0.6	0.5	1.1	0.9
	F	0.9	0.8	1.4	1.2
	M	0.8	0.7	1.8	1.5
NHW	All	0.7	0.6	2.5	2.2
	F	1.0	0.9	3.2	2.8
	M	1.0	0.9	4.0	3.4
Hispanic	All	1.5	1.3	2.0	1.7
	F	2.1	1.8	2.6	2.2
	M	2.1	1.8	3.1	2.7
NHB	All	1.9	1.6	1.8	1.5
	F	2.7	2.3	2.1	1.8
	M	2.7	2.3	3.4	2.9
API	All	4.7	4.1	4.7	4.1
	F	7.4	6.4	6.9	6.0
	M	6.1	5.3	6.5	5.6
AI	All	8.0	6.9	4.2	3.6
	F	11.4	9.8	6.6	5.7
	M	11.2	9.6	5.3	4.6

Detecting a "leveling off" of T1D incidence in NHW Youth. The first 8 years of incidence data collected during the 2002-2009 period suggests a linear trend with a constant rate of increase of about 3% per year. With the accumulation of 5 more years of data SEARCH could be in a position to detect potential changes in incidence trends, and estimate retrospectively the incident year when the change happened. Simulation studies were performed to assess the power to correctly identify the year corresponding to the change point. The simulation study started with the data that is already available in SEARCH, which was used to fit Poisson regression models, which was then used to predict yearly incidence rates until 2018 assuming a linear trend. The model is then perturbed to mimic the effect of a change point that could occur respectively in 2012, 2013, 2014, 2015 and 2016. The perturbed model assumes that the reduction in incidence rate happened at the selected year and remained constant at the new rate in future years. This simulation process indicated that we will have ~70% power to detect a reduction of 5% in the incidence rate after 2016. It should be noted that the Finnish T1D registry study needed more than 30 years of data to be able to retrospectively identify 1988 and 2002 as the years where changes in the incidence rate happened, with only the change point observed in 1988 being statistically significant (62). Our proposed analysis will be conducted in the second half of 2019 – after the close of the 30 months window for the incident 2016 cases.

Adjusting Results for Potential Differences in Agreement between Provider Type and Etiologic Type Over Time. Estimation of incidence trends can be affected by potential temporal changes in provider assessment of diabetes type. Such changes can lead to biased estimation of the trend. We will test for homogeneity of association between diabetes type as assessed by the provider and SEARCH etiologic type over the time period, and adjust for the difference in agreement over time as needed.

6.1.3. *Aim 3: Detectable Differences in Prevalence of DKA*

Assuming a significance level of 0.05, we have 80% power to detect an absolute change of 0.22% (from 30.3% to 30.1% for example) in the prevalence of DKA among T1D cases, and an absolute change of 0.25 (from 7.2% to 6.95%) in T2D cases. This analysis will be conducted after the completion of ascertainment efforts for the 2017 incident cohort.

6.1.4. *Aim 4: Adjusting Results Using Capture-Recapture Analysis*

The completeness of ascertainment for each site will be estimated by dividing the number of identified cases by the estimated total number obtained from the capture-recapture analysis. The capture-recapture corrected estimate will be computed by dividing the observed incidence rate by the estimated capture-recapture rate. This corrected estimate can be seen as a ratio of 2 random variables. Pooled estimates that borrow information across site, sex and age groups will be used to guarantee that the capture-recapture rate and its associated standard error can be computed for all combinations of the variables considered in the analysis. Stratification by site, diabetes type, race/ethnicity, sex and age group can sometimes lead to small cell count causing convergence failures in the maximum likelihood estimation routines. Pooled estimation performed assuming a log-linear model makes it possible to obtain the maximum likelihood estimates in these cases and simplifies the derivation of the standard error associated with the estimated percentage completeness.

6.2. COHORT STUDY - STATISTICAL CONSIDERATIONS

6.2.1. *Aim 1: Burden of Complications*

Three main analytic approaches will be employed to examine the prevalence, incidence, progression and clustering of complications by diabetes type and responsible risk factors and pathways: a) estimating incidence and prevalence using multiple logistic regression methods; b) estimating the progression of complications using longitudinal mixed models; and c) estimating the clustering of risk factors and outcomes using longitudinal mixed models. For each of these approaches we will incorporate participant level characteristics, measured at multiple time points, to examine potential mediators and moderators of outcomes.

Incidence Rate Estimation: Participants have had at least two previous in-person visits (C_0, C_1); however, for many outcomes (retinopathy, neuropathy, etc.) participants will have had only one previous assessment (C_1 visit). For these outcomes we will be able to define a group of participants who were free from the event of interest (e.g. no retinopathy) at C_1. Multiple logistic regression methods will be employed to examine the incidence rates of binary measures, with categorical (e.g.T1D vs. T2D) or continuous (e.g. A1c) predictors. We will evaluate potential confounding and/or effect modification based on our extensive databases.

Prevalence Rate Estimation: Some of the outcomes of interest have not been measured previously (e.g. echocardiography); therefore, we will estimate the prevalence of these outcomes. Associations of risk factors and diabetes type with prevalent outcomes will be examined using logistic regression models.

Statistical Power: For each of the primary dichotomous outcomes of interest (incidence or prevalence) we estimated proportions that will have the specific outcomes of interest, based on data from the C_1 visit. Table 6 shows the expected sample sizes available for comparing T1D and T2D, the corresponding detectable differences in rates, and the power for each outcome. We also provide the expected detectable differences in prevalence rates of LV hypertrophy between T1D and T2D in the sample of patients receiving echocardiography. These calculations are performed using Fisher's exact tests with α =0.05 (2-sided).

Table 6. Power for detectable differences for primary outcomes

Outcome	T1D/T2D available	T1D rate	T2D rate	Power
Incidence Comparisons				
Retinopathy	1215/230	10%	18%	88%
Neuropathy	1344/283	5%	10%	81%
Nephropathy	1172/268	20%	29%	85%
Prevalence Comparisons (Cardiac echocardiography)				
LV hypertrophy	250/250	5%	13%	86%

Longitudinal Models: We will use a longitudinal mixed effects analysis of covariance approach to make comparisons among groups which includes duration of diabetes as a time-varying covariate and participant as a random effect. This approach models the varying durations of disease prior to the initial visit, and the varying durations of time allowed by the data collection windows between visits. These mixed effects models are flexible to allow for non-linear relationships to be modeled over time, and permit random rates of progression.

To estimate the sample size needed to detect a significant difference with sufficient power, calculations were based on comparing measurements after adjusting for C_0 or C_1 data. If the correlation between measurements is moderate (0.50) then we have 80% power to detect a difference of 0.139 standard deviations (SD) for each outcome of interest. For example, based on data collected on a subset of SEARCH T1D participants, the standard deviation for PWV carotid-femoral was estimated to be 0.7 m/s, thus we would have 80% power to detect a difference of 0.10 m/s in progression of PWV between youth with T2D vs T1D.

Clustering of Outcomes: In addition to examining each endpoint separately, we have the opportunity to look simultaneously at several outcomes in the same analysis. We

will create variables that describe the co-occurrence (clustering) of outcomes for each participant and examine whether there are differences in the patterns of these clusters between T1D and T2D youth. Approaches will utilize ordinal logistic regression methods or longitudinal mixed models depending on whether the clustering outcome is a count or categorical. More sophisticated statistical methods may also be used such as principal components analyses to determine which risk factors may form different components.

6.2.2. *Aim 2: Processes of Care*

Analytically, the approach for addressing the questions related to processes of care, their influence on quality of life (QOL) during transition from pediatric to adult care, by diabetes type, will follow closely the approach described above for longitudinal models. For some analyses we need to assess the potential effects of mediators on the examined relationships. Potential mediator data has been measured in at least 3 time points. Furthermore, since the Affordable Care Act was implemented during the time frame when data has been collected, we can examine changes in outcomes that occur before or after that period.

For power calculations we will conservatively estimate a total of 2000 T1D and 382 T2D participants. If we assume that the correlation from the initial assessment of the outcome and the final assessment of the outcome is 0.5 then we can detect an effect size of 0.135 SD with 80% power (alpha=0.05, 2-sided). Thus, for an instrument such as the QOL scale where estimates of the standard deviation range from 13 to 17 this would correspond to having sufficient power to detect a difference between groups of 1.8 to 2.3 units, which is a clinically meaningful difference.

6.2.3. *Aim 3: Mortality*

We will perform both direct and indirect standardization to compare the death rates observed in SEARCH to the age-, race- and sex- matched US population and calculate the Standardized Mortality Ratio (SMR) for each subgroup. Statistical inference will be based on confidence interval estimation and Wald tests. SEARCH data will be used to compare mortality by diabetes type. Time to all-cause mortality will serve as the primary outcome for these analyses, and will be modeled using Cox proportional hazards.

The death rate estimated in the SEARCH 2002-2008 incident cohorts was 91.3 per 100,000 person-years. Based on projections, we expect to observe ~130 deaths in 142,000 person-years by 12/31/2017 (~82 among T1D and ~48 among T2D cases). Assuming alpha=0.05, we will have over 80% power to detect a hazard ratio (HR) of 1.9 or higher. With an observed HR for time to all-cause death of 2.7 in T2D relative to T1D currently, the proposed study will be well-powered to identify differential effect of diabetes type on mortality.

7. Study Organization

The organizational structure of SEARCH 4 is patterned after the successful structure of the previous phases of the study. The Steering Committee is the main governing body, and includes the Principal Investigators from each study site, the central laboratory, and the Coordinating Center; the chair of the Project Managers Committee; and the Project Scientists from the Center for Disease Control (CDC)/ National Center for Chronic Disease Prevention and Health Promotion (NCCDPHP) and the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)**.** When voting is required, each site has one vote (five total), the CC has one vote, the funding agencies have one vote (combined), and the Project Manager Chair has one vote. Two co-chairs are selected from the non-federal Steering Committee members. The Steering Committee has primary responsibility to develop a common research protocol and manual of operations, facilitating the conduct and monitoring of the study, and reporting study results. The Steering Committee also oversees policies for access to participant data and specimens and ancillary studies. A Study Group is comprised of all Steering Committee members, plus additional investigators as well as consultants and project managers from the clinical sites and the CC. Key operational committees report directly to the Steering Committee.

An Observational Study Monitoring Board (OSMB) consisting of appropriately qualified independent experts provides review of data on study progress and participant safety. The purpose of the board is to assure independent review as to whether study participants are exposed to unreasonable risk because of study participation, and to monitor study progress and integrity. Board members are chosen by NIDDK, and typically convene twice a year (every 6 months) unless a need arises. The CC produces a report for review by the OSMB according to pre-determined format, contents, and reporting frequency. The reports present information regarding (1) adverse events and safety violations experienced by study patients as a result of undergoing the study procedures and (2) conduct of the study, including withdrawals and visit attendance.

8. Quality Control

The SEARCH Coordinating Center is responsible for developing and implementing quality control (QC) procedures. QC techniques are incorporated into each phase of the study from case ascertainment, recruitment and registration of persons with diabetes through data acquisition, reading and/or interpretation of the results and their analyses and publication. The Coordinating Center continues to work with the QC Committee reporting to the Steering Committee and to the OSMB. The QC Committee works in concert with the Coordinating Center to oversee the standardized measurement protocols for collecting data during clinic visits and interviews. The committee oversees and recommends any revisions to, or further development of, study data collection forms; develops guidelines for and oversees the central laboratory and reading centers; reviews and monitors quality control related to study measures; and reports on quality control to the study group. This committee also reviews the certification of clinic staff and assists with training and certification/ recertification of study staff on measurement protocols. Any problems

identified with laboratory and reading centers or clinic performance are addressed with remediation plans.

9. Centralized Data Management System

The SEARCH study features an integrated web-based system for managing operations and capturing data as developed by the CC. Once entered, data are immediately validated against sets of rules. Some of these rules identify errors that must be corrected immediately; others present validation warnings for review which are saved to the database for later reconciliation. Data are immediately available in alert/tracking systems and dynamic reports based on relational databases. No records are ever deleted, all changes produce audit trails, and back-ups are created hourly. This provides a high degree of integrity, detail, and flexibility in responding to unexpected study needs related to report generation, auditing, and monitoring. A comprehensive security program is in place that integrates policy and practice (see Appendix A).

The system allows authorized users to access clinic and participant information for the purpose of entering and editing study data. Only authorized users may access and enter/update information regarding participants' study data. Only local site staff and investigators and authorized Coordinating Center staff have access to data from individual sites. A correct username and password is required to gain access to the system and role-based security is employed to restrict user access to only authorized areas and data. All data are stored in a secured Microsoft SQL Server (2008) database system at Wake Forest School of Medicine. The system employs audit logs that capture and store each version of every record that is saved on the system. Users who access the system, once authenticated, establish a secure SSL encrypted session and all transmissions are encrypted until they logout or close the browser. The system is backed up nightly onto dedicated backup storage equipment.

10. Confidentiality

As in previous phases of SEARCH, every precaution is taken to maintain the confidentiality of all study participants. For both the Cohort and Registry Studies, confidentiality of data is maintained by using research identification (ID) numbers that uniquely identify each individual. Hardcopies of individual participants' research records will be retained and secured by each SEARCH Clinical Center. The file that links participants' names and demographic information with their research ID numbers is retained separately from the study data, using an approach consistent with local IRB requirements. After the study is completed, local data are stored with that of other completed studies in a secure storage area following all applicable local regulations for the storage, maintenance, and destruction of research data.

As in previous phases of the SEARCH Study, an NIH Certificate of Confidentiality is maintained at the CC to offer further protection of privacy.

11. Safety Management

The potential risks to individuals participating in the SEARCH 4 Cohort and Registry study components are very few. Participant safety is monitored through center specific guidelines. Study-related adverse events are documented on the Event Reporting Form and submitted to the Coordinating Center. An external reviewer reviews all events reported in this manner and reports findings to the SEARCH Quality Control Committee. The risks are described below along with strategies that are used to minimize these risks.

Blood samples

To minimize the possibility of risks associated with phlebotomy experienced medical staff obtain the blood samples in accordance with local guidelines. A numbing medicine may be placed on the skin before the blood is drawn to decrease any pain. Participants who have a history of fainting or who develop symptoms of light-headedness may be placed in the supine position and blood sugar levels are checked with a blood glucose meter.

Results reporting

Participants (or their parent/guardian if <18 years of age) are given all clinically relevant test results based on measurements and samples collected during their study visits. Transmission of results is based on the age of the participant at the time that the results become available. If the participant's parent agreed to have the samples drawn but the participant is at least 18 years of age when the results become available, then the participant is notified of the results.

Participants (or their parent/guardian if <18 years) are asked whether or not they wish their diabetes and/or primary care provider(s) to receive their clinically relevant test results such as HbA1c, glucose, lipid profile, C-peptide, diabetes autoantibodies, and urine albumin and creatinine. Receipt of these results is viewed as a possible but not definite benefit to the participant as such information may or may not affect subsequent diabetes (or complication) management. If critical laboratory values do occur, the central laboratory contacts the local Principal Investigator and/or his/her designee, and the information is shared with the participant or his/her parent/guardian if <18 years of age, as well as the provider if permission was given at the time of the study visit. Participants with abnormalities needing medical management are referred to their primary care provider (PCP).

Information from interviews is not to be shared with parents or guardians with the exception of the Centers for the Epidemiologic Studies of Depression (CES-D) scale results that are at or above the alert value.

Identification of Alert Values

The following components of the Registry Study IPV and the Cohort Study exam have identified alert levels and a detailed action plan in the Manual of Procedures:

- serum glucose level < 45 mg/dl or > 300 mg/dl;

- triglyceride levels >1000 mg/dl;
- blood pressure > the 95th percentile;
- two first morning urine samples that are positive for leukocytes and nitrites, or blood;
- urine albumin:creatinine ratio ≥ 30µg/mg on repeat first morning void;
- untreated ulcer or infection of feet;
- pathology identified on retinal photography;
- pathology identified on cardiac echocardiography;
- elevated CES-D total score: > 24 for participants < 18 yrs. of age and ≥ 16 for participants ≥ 18 yrs.

APPENDIX A

Information System Security Plan for Wake Forest/Public Health Sciences

General System Description of Data Management System

The SEARCH data management system allows only authorized users to access participant information and enter/update information regarding participants' study data. The application maintains audit logs which identify the activity of each user at all times while logged into the system. This system is built as a web-based application which is accessed via the Internet. A correct username and password is required to gain access to the system and role based security is employed to restrict user access to only authorized areas and data. The application is built using HTML forms and Macromedia's ColdFusion middleware product for database interactions. Javascript and a ColdFusion based rules engine provides data validation and integrity checking on all submitted data. All data is stored in a secured Microsoft SQL Server (2008) database system. The system employs audit logs that capture and store each version of every record that is saved on the system. Users who access the system, once authenticated, establish a secure SSL encrypted session and all transmissions are encrypted until they logout or close the browser.

System Environment

The system is comprised of a Microsoft-based web server which runs Adobe's ColdFusion application server for integration of the database information with the web site. All data resides in a Microsoft SQL Server database with the appropriate role-based security maintained on the data. The application itself also implements role-based security to prevent unauthorized access to or manipulation of confidential information. The system is backed up nightly onto dedicated backup storage equipment. The application is hosted on a virtual server using VMWare. The server is in a secure DMZ zone. The server is maintained as all other servers in a secure data center and updated monthly with patches to the operating system and to the VMWare software. The server is backed up nightly and is on a UPS in the event of a power failure.

Backups

Nightly backups, moved offsite regularly, are made of all data and stored in secure fireproof cabinets. The backup schedule consists of full monthly backups and nightly incremental backups. Backup tapes are handled by two system administrators. Tapes are transported by one of two identified tape custodians. The tapes are moved from the data center to the offsite storage facility and are stored in fireproof cabinets. At all times during the transport, one of the tape custodians is present with the tapes. Tapes are identified by unique bar code labels accessible only by the systems administrators. This is the only information on the tape label. The backup system stores the information for each bar code with details of directories/files backed up that includes the date and time of backup. The backup system, when needing to restore files, will identify which tape is needed based on the bar code label. Only designated system administrators can restore the backup tapes.

Server Management and Data Center

The servers involved in this project are contained within a secure Data Center with environmental controls which detect abnormal conditions such as power outages, high heat or humidity, and loud sound. In the event of an abnormal condition, the system contacts three (3) individuals to notify them of the alerts. The Data Center has several secure access points that

are accessible only by a badge reader. Only authorized staff will have accessible badges to these areas. The building is surrounded by a 10 foot fence with a gate access through badge control. The outside building door is accessed through badge control. The data center room is housed in a locked computer room that is accessed through badge control. Each of these access controls is in place 24 hours a day and seven days a week. All servers have uninterruptible power supplies (UPS). The building has a backup generator that will automatically initiate in the event of a power failure. The computer room is equipped with fire suppression equipment. This equipment is tested on a scheduled timetable by the institution. The entire Data Center is fire-protected by a clean agent system which is backed up by a dry-pipe pre-action sprinkler system. The Data Center room is located on the second floor of the building in an area with no windows and has a raised floor to protect against flooding. The system is protected by a Cisco firewall and is located in a secure DMZ. Servers are protected by institution supported and maintained intrusion detection software as well as by SecureIIS which monitors incoming server traffic.

Password Security

Minimum password requirements must meet Wake Forest Health Sciences Security Policy requirements of:

- Must be changed every 90 days
- Administrative level passwords must change every 30 days
- Must be at least six characters long
- Must include any three of the following items
- English uppercase characters (A through Z)
- English lowercase characters (a through z)
- Numerals (0 - 9)
- Special characters (!, $, #, %, @, etc.)
- The same password cannot be reused in less than 4 previous passwords.

Code Scanning/Testing

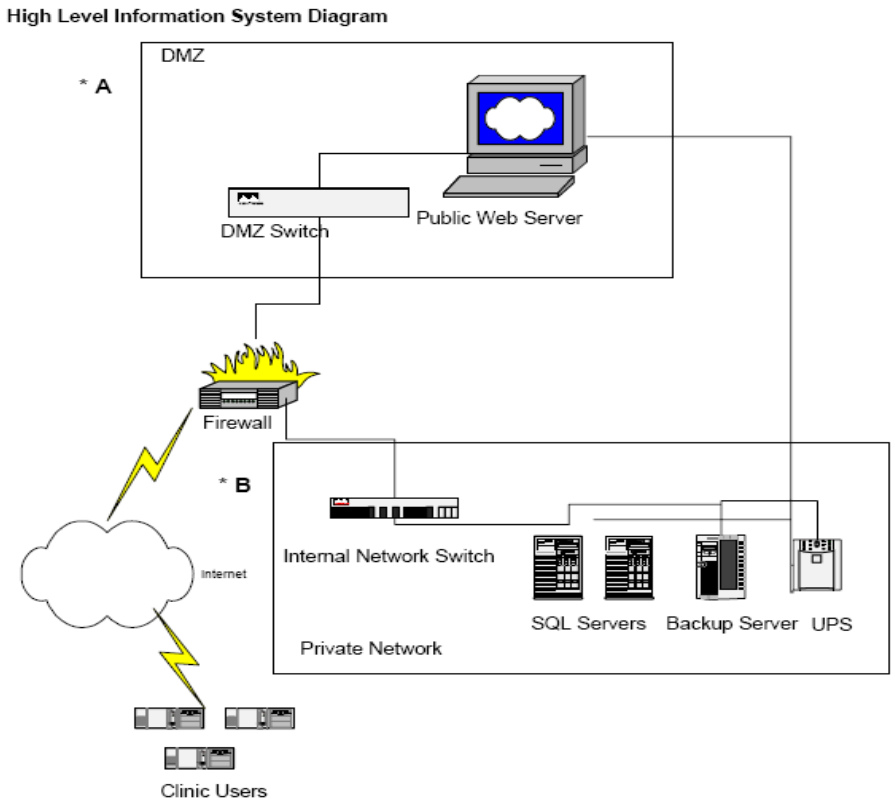
Prior to the release of the web site for public access, the Security Office scans the site for vulnerabilities such as, but not limited to, cross-site scripting, SQL injection attacks, and unsecured logins. The vulnerabilities are classified into five categories of Critical, High, Medium, Low and Best Practices. All Critical and High vulnerabilities must be resolved. Each medium and low vulnerability is reviewed and after discussion with the Security Office, decisions are made to remediate the issue or that the issue is not a security risk to the organization. The Security Office uses the WebInspect product from HP. The tool is automatically updated at each scan for new vulnerabilities. The web site is scanned at the initial release and at least annually thereafter. If significant changes have been made to the site, the site is required to undergo additional scans prior to the annual scan.

Disaster and Contingency Planning

Hurricanes Katrina (ACCORD) and Sandy (SPRINT) have made clear the need for careful disaster planning. While our CCs were not directly impacted by these acts of nature, each forced a clinical site to close (at least temporarily). The Department of Biostatistical Sciences

has a disaster plan as part of our NHLBI-approved information security plan. This plan identifies key personnel that need to be notified in times of disaster as well as which critical systems need to be brought online first. The plan describes how we would continue business operations should a disaster happen by identifying alternative human and computational resources that we could leverage should a disaster strike.

System/Network Diagram



Reference List

1. Primary Care of the Child with a Chronic Condition. 4 ed. St. Louis: Mosby-Yearbook; 2004.

2. Pambianco G, Costacou T, Ellis D, Becker DJ, Klein R, Orchard TJ. The 30-year natural history of type 1 diabetes complications: the Pittsburgh Epidemiology of Diabetes Complications Study experience. *Diabetes* 2006;55(5):1463-1469.

3. Patterson C, Guariguata L, Dahlquist G, Soltesz 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.

4. Patterson CC, Gyurus E, Rosenbauer J, Cinek O, Neu A, Schober E, Parslow RC, Joner G, Svensson J, Castell C, Bingley PJ, Schoenle E, Jarosz-Chobot P, Urbonaite B, Rothe U, Krzisnik C, Ionescu- Tirgoviste C, Weets I, Kocova M, Stipancic G, Samardzic M, de Beaufort CE, Green A, Dahlquist GG, Soltesz G. Trends in childhood type 1 diabetes incidence in Europe during 1989-2008: evidence of non- uniformity over time in rates of increase. *Diabetologia* 2012;55(8):2141-2147.

5. The Writing Group for the SEARCH for Diabetes in Youth Study Group. Incidence of Diabetes in Youth in the United States. *JAMA: The Journal of the American Medical Association* 2007;297(24):2716-2724.

6. Pettitt DJ, Talton J, Dabelea D, Divers J, Imperatore G, Lawrence JM, Liese AD, Linder B, Mayer-Davis EJ, Pihoker C, Saydah SH, Standiford DA, Hamman RF, for the SEARCH for Diabetes in Youth Study Group. Prevalence of Diabetes Mellitus in U.S. Youth in 2009: The SEARCH for Diabetes in Youth Study. *Diab Care* 2014;37(1):402-408.

7. Patterson CC, Gyurus E, Rosenbauer J et al. Trends in childhood type 1 diabetes incidence in Europe during 1989-2008: evidence of non-uniformity over time in rates of increase. *Diabetologia* 2012 August;55(8):2142-7.

8. Lawrence JM, Imperatore G, Dabelea D et al. Trends in incidence of type 1 diabetes among non-Hispanic white youth in the U.S., 2002-2009. *Diabetes* 2014 November;63(11):3938-45.

9. Dabelea D, Mayer-Davis EJ, Saydah S et al. Prevalence of type 1 and type 2 diabetes among children and adolescents from 2001 to 2009. *JAMA* 2014 May 7;311(17):1778-86.

10. Harjutsalo V, Sund R, Knip M, Groop PH. Incidence of type 1 diabetes in Finland. *JAMA* 2013 July 24;310(4):427-8.

11. Pinhas-Hamiel O, Zeitler P. "Who is the wise man?--The one who foresees consequences:". Childhood obesity, new associated comorbidity and prevention. *Prev Med* 2000 December;31(6):702-5.

12. SEARCH Study Group. SEARCH for Diabetes in Youth: a multicenter study of the prevalence, incidence and classification of diabetes mellitus in youth. *Control Clin Trials* 2004 October;25(5):458-71.

13. Hamman RF, Bell RA, Dabelea D et al. The SEARCH for Diabetes in Youth study: rationale, findings, and future directions. *Diabetes Care* 2014 December;37(12):3336-44.

14 Mayer-Davis EJ, Bell RA, Dabelea D, D'Agostino R, Jr., Imperatore G, Lawrence JM, Liu L, Marcovina S, SEARCH for Diabetes in Youth Study Group. The many faces of diabetes in American youth: type 1 and type 2 diabetes in five race and ethnic populations: the SEARCH for Diabetes in Youth Study. Diab Care 2009;32 Suppl 2:S99-101.

15. Dabelea D, Pihoker C, Talton JW, D'Agostino RB, Jr., Fujimoto W, Klingensmith GJ, Lawrence JM, Linder B, Marcovina SM, Mayer-Davis EJ, Imperatore G, Dolan LM. Etiological Approach to Characterization of Diabetes Type: The SEARCH for Diabetes in Youth Study. *Diab Care* 2011;34(7):1628-1633.

16. Hamman, R.F. et al (2013). Estimation of completeness of case ascertainment using capture-recapture. Unpublished report. SEARCH For Diabetes in Youth: a multi-center study. https://www.searchfordiabetes.org/public/dsphome.cfm

17. Lawrence JM; Black MH; Zhang JL; Slezak JM; Takhar HS; Koebnick C, Mayer-Davis EJ; Zhong VW; Dabelea D; Hamman RF; Reynolds K. Validation of Pediatric Diabetes Case Identification Approaches for Diagnosed Cases Using Information in the Electronic Health Records of a Large Integrated Managed Health Care Organization. Am J Epidemiol. 2014 Jan 1;179(1):27-38. doi: 10.1093/aje/kwt230. Epub 2013 Oct 7.

18. Doody MM, Hayes HM, Bilgrad R. Comparability of National Death Index Plus and Standard Procedures for Determining Causes of Death in Epidemiologic Studies. *Ann Epidemiol* 2001;11(1):46- 50.

19. Tamborlane WV, Beck RW, Bode BW, Buckingham B, Chase HP, Clemons R, Fiallo-Scharer R, Fox LA, Gilliam LK, Hirsch IB, Huang ES, Kollman C, Kowalski AJ, Laffel L, Lawrence JM, Lee J, Mauras N, O'Grady M, Ruedy KJ, Tansey M, Tsalikian E, Weinzimer S, Wilson DM, Wolpert H, Wysocki T, Xing D. Continuous glucose monitoring and intensive treatment of type 1 diabetes. *N Engl J Med* 2008;359(14):1464-1476.

20. Beck RW, Buckingham B, Miller K, Wolpert H, Xing D, Block JM, Chase HP, Hirsch I, Kollman C, Laffel L, Lawrence JM, Milaszewski K, Ruedy KJ, Tamborlane WV. Factors predictive of use and of benefit from continuous glucose monitoring in type 1 diabetes. *Diab Care* 2009;32(11):1947-1953.

21. Mayer-Davis EJ, Nichols M, Liese AD, Bell RA, Dabelea DM, Johansen JM, Pihoker C, Rodriguez BL, Thomas J, Williams D. Dietary intake among youth with diabetes: the SEARCH for Diabetes in Youth Study. *J Am Diet Assoc* 2006;106(5):689-697.

22. O'Neill JR, Liese AD, McKeown RE, Cai B, Cuffe SP, Mayer-Davis EJ, Hamman RF, Dabelea D. Physical Activity and Self-Concept: The SEARCH for Diabetes in Youth Case Control Study. *Pediatr Exerc Sci* 2012;24(4):577-588.

23. Li C, Beech B, Crume T, D'Agostino RB, Jr., Dabelea D, Kaar JL, Liese AD, Mayer-Davis EJ, Pate R, Pettitt DJ, Taplin C, Rodriguez B, Merchant AT. Longitudinal association between television watching and computer use and risk markers in diabetes in the SEARCH for Diabetes in Youth Study. *Pediatr Diabetes* 2014;10.1111/pedi.12163.

24. Reynolds K, Liese AD, Anderson AM, Dabelea D, Standiford D, Daniels SR, Waitzfelder B, Case D, Loots B, Imperatore G, Lawrence JM. Prevalence of tobacco use and association between cardiometabolic risk factors and cigarette smoking in youth with type 1 or type 2 diabetes mellitus. *J Pediatrics* 2011;158(4):594-601.

25. Winters KC. Development of an adolescent alcohol and other drug abuse screening scale: Personal experience screening questionnaire. *Addictive Behaviors* 1992;17(5):479-490.

26. Lawrence JM, Standiford DA, Loots B, Klingensmith GJ, Williams DE, Ruggiero A, Liese AD, Bell RA, Waitzfelder BE, McKeown RE. Prevalence and correlates of depressed mood among youth with diabetes: the SEARCH for Diabetes in Youth study. *Pediatrics* 2006;117(4):1348-1358.

27. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *App Psych Measure* 1977;1:385-401.

28. Hood KK, Lawrence JM, Anderson A, Bell R, Dabelea D, Daniels S, Rodriguez B, Dolan LM. Metabolic and inflammatory links to depression in youth with diabetes. *Diab Care* 2012;35(12):2443-2446.

29. Olino TM, Yu L, McMakin DL, Forbes EE, Seeley JR, Lewinsohn PM, Pilkonis PA. Comparisons across depression assessment instruments in adolescence and young adulthood: an item response theory study using two linking methods. *J Abnorm Child Psychol* 2013;41(8):1267-1277.

30. Cox DJ, Irvine A, Gonder-Frederick L, Nowacek G, Butterfield J. Fear of hypoglycemia: quantification, validation, and utilization. *Diab Care* 1987;10(5):617-621.

31. Green LB, Wysocki T, Reineck BM. Fear of hypoglycemia in children and adolescents with diabetes. J *Pediatr Psychol* 1990;15(5):633-641.

32. Clarke WL, Gonder-Frederick A, Snyder AL, Cox DJ. Maternal fear of hypoglycemia in their children with insulin dependent diabetes mellitus. *J Pediatr Endocrinol Metab* 1998;11 Suppl 1:189-194.

33. Gonder-Frederick L, Nyer M, Shepard JA, Vajda K, Clarke W. Assessing fear of hypoglycemia in children with Type 1 diabetes and their parents. *Diabetes Manag* (Lond) 2011;1(6):627-639.

34. Hood KK, Butler DA, Anderson BJ, Laffel LMB. Updated and Revised Diabetes Family Conflict Scale. *Diab Care* 2007;30(7):1764-1769.

35. Schmitt A, Gahr A, Hermanns N, Kulzer B, Huber J, Haak T. The Diabetes Self-Management Questionnaire (DSMQ): development and evaluation of an instrument to assess diabetes self-care activities associated with glycemic control. *Health Qual Life Outcomes* 2013;11:138.

36. Agency for Health Care Research and Quality. CAHPS Item Set for Children with Chronic Conditions. https://cahps ahrq gov/surveys-guidance/item-sets/children-chronic/index html, 2014 (Accessed December 17, 2014 https://cahps.ahrq.gov/surveys-guidance/item-sets/children-chronic/index.html).

37. American Diabetes Association. Diagnosis and Classification of Diabetes Mellitus. *Diab Care* 2014;37(Supplement 1):S81-S90.

38. Naughton MJ, Ruggiero AM, Lawrence JM, Imperatore G, Klingensmith GJ, Waitzfelder B, McKeown RE, Standiford DA, Liese AD, Loots B, SEARCH for Diabetes in Youth Study Group. Health-related quality of life of children and adolescents with type 1 or type 2 diabetes mellitus: SEARCH for Diabetes in Youth Study. *Arch Pediatr Adolesc Med* 2008;162(7):649-657.

39. Lawrence JM, Yi-Frazier JP, Black MH, Anderson A, Hood K, Imperatore G, Klingensmith GJ, Naughton M, Mayer-Davis EJ, Seid M. Demographic and Clinical Correlates of Diabetes-Related Quality of Life among Youth with Type 1 Diabetes. *J Pediatrics* 2012;161(2):201-207.

40. Naughton MJ, Yi-Frazier JP, Morgan TM, Seid M, Lawrence JM, Klingensmith GJ, Waitzfelder B, Standiford DA, Loots B. Longitudinal associations between sex, diabetes self-care, and health-related quality of life among youth with type 1 or type 2 diabetes mellitus. *J Pediatrics* 2014;164(6):1376-1383.

41. Ngo-Metzger Q, Sorkin DH, Billimek J, Greenfield S, Kaplan SH. The effects of financial pressures on adherence and glucose control among racial/ethnically diverse patients with diabetes. *J Gen Intern Med* 2012;27(4):432-437.

42. Coleman EA, Mahoney E, Parry C. Assessing the quality of preparation for posthospital care from the patient's perspective: the care transitions measure. *Med Care* 2005;43(3):246-255.

43. Lang RM, Bierig M, Devereux RB, Flachskampf FA, Foster E, Pellikka PA, Picard MH, Roman MJ, Seward J, Shanewise JS, Solomon SD, Spencer KT, Sutton MS, Stewart WJ. Recommendations for chamber quantification: a report from the American Society of Echocardiography's Guidelines and Standards Committee and the Chamber Quantification Writing Group, developed in conjunction with the European Association of Echocardiography, a branch of the European Society of Cardiology. *J Am Soc Echocardiogr* 2005;18(12):1440-1463.

44. Gottdiener JS, Bednarz J, Devereux R, Gardin J, Klein A, Manning WJ, Morehead A, Kitzman D, Oh J, Quinones M, Schiller NB, Stein JH, Weissman NJ. American Society of Echocardiography recommendations for use of echocardiography in clinical trials. *J Am Soc Echocardiogr* 2004;17(10):1086-1119.

45. Devereux RB, Alonso DR, Lutas EM, Gottlieb GJ, Campo E, Sachs I, Reichek N. Echocardiographic assessment of left ventricular hypertrophy: comparison to necropsy findings. *Am J Cardiol* 1986;57(6):450-458.

46. National Center for Health Statistics. National Health and Nutrition and Examination Survey (NHANES) Ophthalmology Procedures Manual. 2005 Sep; www.cdc.gov/nchs/data/nhanes/nhanes_05_06/OP.pdf<http://www.cdc.gov/nchs/data/nhanes/nhanes_05_06/OP.pdf>.

47. Klein BE, Knudtson MD, Tsai MY, Klein R. The relation of markers of inflammation and endothelial dysfunction to the prevalence and progression of diabetic retinopathy: Wisconsin epidemiologic study of diabetic retinopathy. *Arch Ophthalmol* 2009;127(9):1175-1182.

48. Schwartz GJ, Schneider MF, Maier PS, Moxey-Mims M, Dharnidharka VR, Warady BA, Furth SL, Munoz A. Improved equations estimating GFR in children with chronic kidney disease using an immunonephelometric determination of cystatin C. *Kidney Int* 2012;82(4):445-453.

49. Inker LA, Shaffi K, Levey AS. Estimating glomerular filtration rate using the chronic kidney disease- epidemiology collaboration creatinine equation: better risk predictions. *Circ Heart Fail* 2012;5(3):303- 306.

50. Sharma AP, Yasin A, Garg AX, Filler G. Diagnostic Accuracy of Cystatin C–Based eGFR Equations at Different GFR Levels in Children. *Clin J Am Soc Nephro* 2011; 6: 1599–1608.
51. Bouvet Y, Bouissou F, Coulais Y, Seronie-Vivien S, Tafani M, Decramer S, et al: GFR is better estimated by considering both serum cystatin C and creatinine levels. *Pediatr Nephrol* 2006; 21: 1299–1306

52. Feldman EL, Stevens MJ, Thomas PK, Brown MB, Canal N, Greene DA. A practical two-step quantitative clinical and electrophysiological assessment for the diagnosis and staging of diabetic neuropathy. *Diab Care* 1994;17(11):1281-1289.

53. Moghtaderi A, Bakhshipour A, Rashidi H. Validation of Michigan neuropathy screening instrument for diabetic peripheral neuropathy. *Clin Neurol Neurosurg* 2006;108(5):477-481.

54. Herman WH, Pop-Busui R, Braffett BH, Martin CL, Cleary PA, Albers JW, Feldman EL, The DCCT/EDIC Research Group. Use of the Michigan Neuropathy Screening Instrument as a measure of distal symmetrical peripheral neuropathy in Type-á1 diabetes: results from the Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications. *Diabetic Med* 2012;29(7):937-944.

55. Wadwa RP, Urbina EM, Anderson AM, Hamman RF, Dolan LM, Rodriguez BL, Daniels SR, Dabelea D. Measures of arterial stiffness in youth with type 1 and type 2 diabetes: the SEARCH for Diabetes in Youth study. *Diab Care* 2010;33(4):881-886.

56. Lang RM, Bierig M, Devereux RB, Flachskampf FA, Foster E, Pellikka PA, Picard MH, Roman MJ, Seward J, Shanewise JS, Solomon SD, Spencer KT, Sutton MS, Stewart WJ. Recommendations for chamber quantification: a report from the American Society of Echocardiography's Guidelines and Standards Committee and the Chamber Quantification Writing Group, developed in conjunction with the European Association of Echocardiography, a branch of the European Society of Cardiology. *J Am Soc Echocardiogr* 2005;18(12):1440-1463.

57. NIH Toolbox for the assessment of neurological and behavioral function. http://www.nihtoolbox.org/Pages/default aspx, 2014 (Accessed December 8, 2015).

58. The DCCT research group. Epidemiology of severe hypoglycemia in the diabetes control and complications trial. *Am J Med* 1991;90(4):450-459.

59. Svoren BM, Volkening LK, Butler DA, Moreland EC, Anderson BJ, Laffel LM. Temporal trends in the treatment of pediatric type 1 diabetes and impact on acute outcomes. *J Pediatr* 2007;150(3):279-285.

60. Zhong VW, Pfaff ER, Beavers DP et al. Use of administrative and electronic health record data for development of automated algorithms for childhood diabetes case ascertainment and type classification: the SEARCH for Diabetes in Youth Study. *Pediatr Diabetes* 2014 December;15(8):573-84.

61. Dabelea D, DeGroat J, Sorrelman C et al. Diabetes in Navajo youth: prevalence, incidence, and clinical characteristics: the SEARCH for Diabetes in Youth Study. *Diabetes Care* 2009 March;32 Suppl 2:S141-S147.

62. Elding Larsson H, Vehik K, Bell R, Dabelea D, Dolan L, Pihoker C, Knip M, Veijola R, Lindblad B,Samuelsson U, Holl R, Haller MJ, on behalf of the TEDDY Study Group, SEARCH Study Group, Swediabkids Study Group, DPV Study Group, Finnish Diabetes Registry Study Group. Reduced Prevalence of Diabetic Ketoacidosis at Diagnosis of Type 1 Diabetes in Young Children Participating in Longitudinal Follow-Up. Diab Care 2011;34(11):2347-2352.