Available online at www.sciencedirect.com

SCIENCE DIRECT®

Controlled Clinical Trials 25 (2004) 458–471

Controlled Clinical Trials

www.elsevier.com/locate/conclintrial

SEARCH for Diabetes in Youth: a multicenter study of the prevalence, incidence and classification of diabetes mellitus in youth

The SEARCH Study Group[1]

Received 7 November 2003; accepted 17 August 2004

Abstract

SEARCH for Diabetes in Youth is an observational, multicenter study focusing on physician-diagnosed diabetes in individuals <20 years old. The study will estimate the population prevalence and incidence of diabetes by type, age, gender, and ethnicity and develop practical approaches to diabetes classification in 5 million children (~6% of the <20 U.S. population) with wide ethnic and socioeconomic representation from four geographically

Corresponding author. Ronny Bell. Wake Forest University School of Medicine, Department of Public Health Sciences, Medical Center Blvd., Winston-Salem, NC 27157-1063, USA. Tel.: +1 336 716 9736; fax: +1 336 713 4300.

E-mail address: rbell@wfubmc.edu.

[1] Writing group: Dr. David J. Pettitt (Chair), Sansum Medical Research Institute, Santa Barbara, CA; Dr. Ronny Bell, Wake Forest University School of Medicine; Dr. Dana Dabelea, University of Colorado Health Sciences Center, Denver, CO; Dr. Lawrence Dolan, Cincinnati Children's Hospital, Cincinnati, OH; Dr. Giuseppina Imperatore, Centers for Disease Control and Prevention, Atlanta, GA; Dr. Jean M. Lawrence, Kaiser Permanente Southern California, Pasadena, CA; Dr. Angela D. Liese, University of South Carolina, Columbia, SC; Dr. Lenna L. Liu, Children's Hospital and Regional Medical Center, Seattle, WA; Ms. Beth Waitzfelder, Pacific Health Research Institute, Honolulu, HI. Search steering committee: Pacific Health Research Institute, Honolulu: Dr. Beatriz L. Rodriguez, Dr. Teresa Hillier, Ms. Beth Waitzfelder; Children's Hospital and Regional Medical Center, Seattle: Dr. Catherine Pihoker, Dr. Irl Hirsch, Dr. Carla Greenbaum, Dr. Lenna Liu; Kaiser Permanente Southern California, Pasadena: Dr. Diana B Petitti, Dr. Jean M Lawrence, Dr. Ann Kershnar, Dr. David J. Pettitt; University of Colorado Health Sciences Center, Denver: Dr. Richard F. Hamman, Dr. Dana Dabelea, Dr. Georgeanna J. Klingensmith, Dr. Marian Rewers, Dr. Jonathon Krakoff, Ms. Patricia V. Nash, Ms. Carissa M. Smith; Children's Hospital Medical Center, Cincinnati: Dr. Lawrence Dolan, Ms. Debra A. Standiford, Dr. Stephen R. Daniels; University of South Carolina School of Public Health, Columbia: Dr. Elizabeth J. Mayer-Davis, Dr. Angela Liese, Dr. John Oeltmann; Northwest Lipid Research Laboratories, Seattle: Dr. Santica Marcovina, Mr. Alan Aldrich; Wake Forest University School of Medicine: Dr. Timothy Morgan, Dr. Lyn Hardy, Ms. Susan Vestal, Dr. Ronny Bell; National Institute of Diabetes and Digestive and Kidney Diseases: Dr. Barbara Linder. Centers for Disease Control and Prevention: Dr. Giuseppina Imperatore, Dr. Michael Engelgau, Dr. Henry Kahn, Dr. Venkat Narayan, Dr. Jinan Saaddine, Dr. Rodolfo Valdez, Dr. Desmond Williams.

0197-2456/$ - see front matter
doi:10.1016/j.cct.2004.08.002

defined populations and two health plans. An estimated 6000 prevalent and 800 incident diabetes cases per year will be identified with annual follow-up. Cases will be ascertained through clinical and nonclinical resources or partnerships at each site. Data collection involves patient interviews, physical examinations, laboratory measurements (diabetes autoantibodies, fasting/stimulating C-peptide, hemoglobin A1c, blood glucose, lipids, urine albumin, creatinine), medical records reviews, and documentation of risk factors for complications and processes of care.

Keywords: Type 1 diabetes; Type 2 diabetes; Children; Adolescents; Ethnic groups

Diabetes is the third most common chronic disease of childhood [1]. However, major gaps exist in knowledge of the types, frequency, pathophysiology, natural history, and processes of care. Until the past decade, types of diabetes other than type 1 were rarely diagnosed in children and adolescents. Recently, several reports describe type 2 diabetes as a pediatric disease [2,3]. However, outside of some American-Indian groups and limited data in African-American and Hispanic populations [4], there are virtually no population-based studies of childhood type 2 diabetes, and the population prevalence of type 2 diabetes is not known. Also, it is unclear whether the frequency of type 1 diabetes is increasing. Between 1989 and 1994, there was an increase in the average annual incidence of type 1 diabetes in Europe of 3.4% overall and 6.3% in children aged 0 to 4 years [5]. However, this varied from country to country. In the United States, the incidence of type 1 diabetes appears to be stable in Colorado and Chicago but has been reported to be increasing in Allegheny County, Pennsylvania, and Hawaii [6,7]. The reasons for these discrepancies are not clear.

To date, there are no gold-standard definitions of different types of diabetes presenting in youth. Clinical phenotypes at onset frequently overlap. Obesity and diabetic ketoacidosis can be found in both type 1 and type 2 diabetes [2,8,9], and age at diagnosis poorly differentiates between types. Therefore, a diagnosis based on pathogenesis is a more effective tool for identifying diabetes type [10,11]. Finally, there is a paucity of data concerning not only the frequency of complications but also processes of care by diabetes type and by race/ethnicity and gender among youth with diabetes.

To address these issues, SEARCH for Diabetes in Youth (SEARCH) was initiated in 2000. The primary aims of SEARCH are (1) to estimate the population prevalence and incidence of type 1, type 2, and other types (or hybrids, i.e., having evidence of autoimmunity as well as continuing insulin secretion) of diabetes overall and by age, gender, and race/ethnicity; (2) to develop efficient and practical approaches to the classification of diabetes type for prevalent and incident cases; and (3) to describe and compare clinical presentation and course of type 1, type 2, and other types (or hybrids) of diabetes. The secondary aims of the study are to describe by type (1) the distribution of risk factors for selected microvascular and macrovascular disease complications; (2) the distribution of selected acute and chronic complications; and (3) the health care utilization, processes of care, and quality of life. Much work has been done previously in type 1 diabetes, hence SEARCH will be confirming and updating findings in type 1 while breaking new ground for type 2 and other types of diabetes. Thus, SEARCH will bridge existing gaps in the knowledge of diabetes types, prevalence, pathogenesis, frequency of the complications, quality of care, and health utilization for diabetes in youth in the United States. These data are essential for the development of public health strategies to prevent diabetes, effectively treat

diabetes and its complications, and limit the personal and financial burden on the patient and the financial burden on health care resources.

1. Methods

SEARCH for Diabetes in Youth is an observational, multicenter, population-based study focusing on cases of physician-diagnosed diabetes in individuals <20 years old. The study is attempting to identify and enroll all eligible cases of diabetes that are (a) prevalent in the year 2001 and (b) newly diagnosed (incident) on and after January 1, 2002 through 2004. A nationally standardized data collection effort therefore builds on the local case ascertainment. SEARCH will then estimate the population prevalence and incidence of diabetes by type, age, gender, and race/ethnicity. In addition, the project has a prospective cohort component that involves annual follow-up.

1.1. Study populations and denominator definitions

Six clinical centers are participating in the SEARCH study (Table 1): Cincinnati Children's Hospital Medical Center, Cincinnati, OH; University of Colorado Health Sciences Center, Department of Preventive Medicine and Biometrics, Denver, CO; Seattle Children's Hospital and Regional Medical Center, Division of Endocrinology, Seattle, WA; University of South Carolina, Department of Epidemiology and Biostatistics, Columbia, SC; Kaiser Permanente Southern California, Pasadena, CA; and Pacific Health Research Institute, Honolulu, HI. Four sites (Cincinnati, Colorado, Seattle, South Carolina) are identifying cases of diabetes in geographically defined populations. Two sites (Hawaii and Southern California) are identifying cases of diabetes in membership-based health plans. The ethnic distribution of the Southern California site is very close to that of the population of all of Southern California, and because Kaiser Permanente subsidizes the dues for families that cannot otherwise afford health insurance and accepts children whose health insurance is paid by Medicaid, the full socioeconomic range is represented. In Hawaii, approximately 95% of the population of the state is enrolled in one of the health plans that is participating in this study. At some sites, surveillance populations are larger for the incidence than for the prevalence component to assure an adequate number of incident cases in the study.

SEARCH includes a wide range of racial and ethnic groups. Within the base population of eligible SEARCH participants, the estimated racial and ethnic distribution is 65% non-Hispanic White, 11% Hispanic, 13% African-American, 6% Asian, 2% Pacific Islander, and 3% Native American.

To estimate the total number of persons in the denominator in 2001, the four geographically based sites (Cincinnati, Colorado, Seattle, South Carolina) plan to use the nonmilitary, noninstitutionalized 2000 census denominators if these are available at the level of detail required. For years 2002 and beyond, the geographically based centers are using projections of population changes based on the 2000 Census to estimate denominators for incidence. At the membership-based sites, address data geocoded to the 2000 census data at the block level will be used to estimate the ethnicity distribution of the member population for the 2001 prevalence year and for incidence years over a period of time during which the ethnicity distribution remains stable as determined based on consultation with local and census experts.

Table 1
Description of base population and summary of source of estimated cases for prevelance component of SEARCH study populations under surveillance

Study Site	Prevelance Component		Incidence Component	
	Base Population	Size	Base Population	Size
Cincinnati	Cincinnati and 8 surrounding urban counties (Hamilton, Butler, Warren, Clermont OH; Boone, Kenton, Campbell KY; Dearborn IN)	550,430	Cincinnati and 8 surrounding urban counties (Hamilton, Butler, Warren, Clermont OH; Boone, Kenton, Campbell KY; Dearborn IN)	550,430
Colorado	Denver and 4 surrounding counties (Adams, Douglas, Jefferson, Boulder), Six rural counties in South-Central Colorado (Conejos, Costilla, Alamosa, Sauguache, Mineral, Rio Grande), Mesa county in Western Colorado, Native American reservations in Arizona and New Mexico	808,503	All 63 counties in Colorado, Native American reservations in Arizona and New Mexico	1,420,839
Hawaii	Members of Hawaii Medical Service Association, Kaiser Foundation Health Plan-Hawaii and the Hawaii State Department of Health Services Med-Quest in Oahu county	240,260	Members of Hawaii Medical Service Association, Kaiser Foundation Health Plan-Hawaii and the Hawaii State Department of Health Services Med-Quest in all counties in Hawaii	300,327
Seattle	King, Pierce, Snohomish, Kitsap, Thurston counties of Washington State	982,920	King, Pierce, Snohomish, Kitsap, Thurston counties	982,920
South Carolina	4 counties (Richard, Lexington, Orangeburg, Calhoun) surrounding Columbia, South Carolina	179,238	All 46 counties in South Carolina	1,118,022
Southern California	Members of the Kaiser Permanente Medical Care Program in Southern California except San Diego	700,450	Members of the Kaiser Permanente Medical Care Program Health Plan Southern California except San Diego	700,450
All sites		3,444,039		5,072,988

For all centers, race/ethnicity data will be grouped in a uniform manner and collapsed into groups (non-Hispanic White, Hispanic, African-American, Asian, Pacific Islander, Native American, Other and Unknown) using rules and conventions developed by the Census [12]. Other potential racial groupings of scientific interest will be considered.

1.2. Case definition and validation

SEARCH cases are considered valid (or true cases) if (a) there is a physician diagnosis of diabetes; or, (b) the parent or the youth self-reports a physician diagnosis of diabetes. A physician-diagnosed case of diabetes is established if any of the following criteria are met: medical record review indicates a physician diagnosis of diabetes; the diagnosis of diabetes is directly verified by a physician; the diabetes

case is "referred" to the study by a physician; diabetes is listed as the underlying or contributing cause of death on a death certificate; or the case is included in a clinical database that has a requirement for verification of diagnosis of diabetes by a physician. This study will exclude cases of gestational diabetes.

1.3. Eligibility and exclusion criteria

To be eligible to participate in SEARCH (Table 2), participants must be in the index year (prevalent 2001 or incident 2002, 2003, 2004): (a) <20 years old and (b) resident of the population defined in the index year for geographically based centers or member of the participating health plan in the index year for membership-based centers. Because the denominators are derived from Census data for the four geographically based centers, individuals who are active duty military or institutionalized are excluded.

1.4. Case ascertainment

The approaches for identification of prevalent and incident cases vary by site based on availability of existing diabetes databases and access to clinics, physicians, and computer-stored data resources. Despite different approaches, each geographic or membership-based site is attempting to identify every case of diabetes within its purview. Geographic-based sites (Cincinnati, Colorado, Seattle, South Carolina) have established active surveillance systems de novo for SEARCH, based on networks of pediatric and adult endocrinologists, existing pediatric diabetes databases, hospitals, health plan databases, and other health care providers. Membership-based sites (Hawaii and Southern California) are using existing diabetes databases in addition to their administrative databases as the source for case identification. Death certificate searches are conducted to identify missed fatal cases (and to assess mortality from diabetes). In the population under surveillance, all deaths, regardless of cause, for persons with diabetes <20 years old are being identified and recorded.

1.5. Assessment of completeness

In SEARCH, evaluation of completeness of case ascertainment is crucial and is being conducted using both statistical methods and additional, targeted data collection efforts. Capture–recapture [13–16],

Table 2
Eligibility criteria for SEARCH

Prevalence	*Incidence*
Physician diagnosed cases of diabetes mellitus	Physician diagnosed cases of diabetes mellitus
Prevalent in 2001	First clinical diagnosis of diabetes in a non-pregnant state January 1 through December 31 in the incidence year
Age less than 20 years on December 31, 2001; Born between 1/1/82–12/31/2001	Age less than 20 years at diagnosis
Resident of the population at any time in 2001 or member of the participating health plan at any time in 2001	Resident of the population at diagnosis or member of the participating health plan at diagnosis
Not active–duty military	Not active–duty military
Not living in an institution (Census definition)	Not living in an institution (Census definition)
Not Gestational Diabetes	Not Gestational Diabetes

a statistical approach that attempts to estimate completeness from incomplete samples, is used in the geographically based sites with multiple independent sources of cases (Cincinnati, Colorado, Seattle, South Carolina). These sources may include hospital discharge, laboratory, pharmacy, ambulatory billing, and pediatric endocrinology case lists. This method will not be applied in the two membership-based centers, because sources of case identification are highly dependent. In addition, intensive case-finding based on a mailed survey to a defined sample of providers in specialties likely to see youth with diabetes, who are not included in primary case ascertainment, will be conducted in geographically based centers. The cases identified will be compared with cases that have already been identified by SEARCH as permitted by local Institutional Review Boards (IRBs).

1.6. Protection of human subjects

The SEARCH protocol has been reviewed and approved by the Institutional Review Boards (IRBs) of each of the participating institutions. Written informed consent is obtained from participants age >18 years or their parents or legal guardians if <18 years. Assent is obtained in accordance with local IRB requirements. Subjects may agree to participate in the study at a number of levels, which is reflected in staged consent and assent forms. In order to further protect the privacy of participants, SEARCH has obtained a certificate of confidentiality for each of the sites.

Because subjects will be asked to appear fasting and some laboratory tests that are not generally a part of standard diabetes care will be performed, alert values have been defined for clinically relevant data measurements and processes established to provide appropriate medical care. Adverse events will be reported to Wake Forest University School of Medicine (coordinating center), each clinical center, and the local IRBs. The study protocol will be modified as required to maintain participant safety.

1.7. Data collection and measures

Data collection is organized into a series of sections that consist of one or more data collection instruments or measurements. All staff have been trained and certified on the standardized protocol and manual of procedures.

All data collection forms are available in English and Spanish. Staff is either bilingual in English and the subject's preferred language or arrangements for an on-site translator are made.

1.7.1. Initial patient survey

The Initial Patient Survey (IPS) is a questionnaire used to collect information to assist with case validation and to confirm eligibility. It also collects information about race/ethnicity, diabetes type, and preliminary treatment information. The IPS may be completed either as a self-administered mailed questionnaire (by the parent, or depending on age, by the participant), during a telephone interview, or at the time of the in-person visit. It is expected that some youth with diabetes will be willing to complete the IPS by mail or telephone even if they are not willing to participate in other parts of the study.

1.7.2. In-person visit

The in-person visit consists of a physical examination (anthropometry, blood pressure, acanthosis), laboratory work (fasting blood and urine collections), and the administration of several questionnaires to

collect information on medical history, family history, quality of life, depression, and health behaviors. This visit is designed so it can be conducted in clinical research settings, health clinics, or the participants' homes. For incident case participants, the in-person visit is conducted as soon as possible after the subject becomes clinically stable (approximately one month after initial diagnosis).

Questionnaires are interviewer-administered (health, family history, quality of life—child report, health behaviors) or self-administered after staff instruction (quality of life—parent report, depression, diet). The primary respondents for the health and family history questionnaires are the parent or legal guardian. Participants >10 years old are asked to complete additional questionnaires, either interviewer-administered or self-administered after staff instruction, focusing on health behaviors (diet, smoking, sleeping patterns) and depression. The parents or legal guardians of participants <18 years old are asked to waive their right to review their children's responses to these questionnaires prior to completion.

Blood and urine specimens are collected from all prevalent and incident cases, and the physical examination is conducted on those aged 3 years or older. The exceptions are cases with other specific types of diabetes, e.g., diabetes known to be secondary to another illness, such as cystic fibrosis or medications, such as steroids, in whom data collection is limited to brief, annual mailed surveys.

Laboratory samples (with the exception of diabetes autoantibodies{DAA}) are obtained under conditions of metabolic stability, defined as no episode of diabetic ketoacidosis during the previous month. The DAA that will be obtained include glutamic acid decaboxylase, IA-2, and insulin autoantibodies. Specimens are processed locally at the sites and shipped within 24 h to the central laboratory (Northwest Lipid Laboratory, University of Washington, Seattle, WA) where they are analyzed. Serum or plasma samples are analyzed for GAD65 antibodies in a radioligand-binding assay. Details of the assay have been published previously [17,18]. In short, human recombinant GAD65 is produced by in vitro transcription and translation system (Promega, Madison, WI) and labeled with ^{35}S-methionine [19]. Reduction in the background radioactivity is achieved by filteration of the assay buffers through a 0.22-μm filter (Millipore, Bedford, MA). The labeled GAD65 is incubated with serum in duplicate determinations. Protein-A-Sepharose (Zymed, San Francisco, CA) is used to separate the GAD-tracer–GAD-antibody complexes from the free ^{35}S-GAD65. Following repeated washings in Millipore Multiscreen micron filteration plates, plates are dried, added with microscint fluid, and radioactivity in wells is determined using a Top Count (96-well plate Beta counter, Packard). The laboratory uses a set of positive and negative standard in each plate. Concentrations of GAD65 Ab are expressed as a GAD65AB index to correct for the interassay variation according to the following formula: GAD65Ab index=[counts per minute (cpm) of the unknown sample-average cpm of two negative standards]/(cpm of the positive standard-average of two negative standards). The positive and negative controls are run in duplicate in each assay. Standards and quality controls: the assay uses a positive control which is the WHO standard for islet cell antibodies. A negative control sample used for the assay was prepared from a pool of normal sera, and therefore the pool can be reproduced as necessary. A signal-to-noise ratio of 10 or above is a requirement before an assay is considered acceptable.

The IA2 antibody assay is identical to the GAD65Ab assay, except using ^{35}S-labelled IA2 as a tracer. IA2Ab index=(cpm of the unknown sample-average cpm of two negative standards)/(cpm of the positive standard-average of two negative standards). The positive and negative controls are run in duplicate in each assay. Standards and quality controls: the assay uses a positive control which is the WHO standard for islet cell antibodies. A negative control sample used for the assay was prepared from a pool of normal sera, and therefore the pool can be reproduced as necessary. A signal-to-noise ratio of 10 or above is a requirement before an assay is considered acceptable.

The assay for insulin autoantibody (IAA) utilizes a competition of insulin antibodies in a serum or plasma sample for unlabeled and labeled insulin for quantitative determination of the antibody levels [20]. 125 I-Insulin (Amersham) of 20,000 cpm is incubated with 5 μl of serum with and without cold insulin diluted in buffer A (20 mM Tris, 150 mM NaCl, 1% BSA, 0.15% Tween-20, and 0.1% sodium azide). Following a 3-day incubation at 4 °C, 50 μl of 50% protein-A/8% protein-G-Sepharose (Pharmacia) mixture is added to the incubation in a Multiscreen-NOB 96-well filteration plate (Millipore plate), which has been precoated with buffer A overnight at room temperature. The plate is shaken for 45 min at 4 °C followed by extensive washing in buffer B (buffer A containing 0.1% BSA) using a vacuum-operated plate washer. After washing, 40 μl of scintillation fluid (Microscint-20, Packard) is added to each well, and radioactivity is determined using a Top Count (96-well plate Beta counter, Packard). The results are calculated based on the difference in counts per minute (Δcpm) between the well without cold insulin and the well with insulin and expressed as an index (IAA index).

1.7.3. C-peptide testing

C-peptide is measured fasting and following a mixed meal challenge with Boost® (Mead Johnson and Company, Evansville, IN) with blood samples drawn at 30, 60, and 90 min after the liquid meal. The subgroups of participants who undergo this test and the frequency of this test are described under link to Classification of Diabetes Type. There are three objectives to stimulated C-peptide testing: (1) to develop better ways to distinguish type 1 and type 2 diabetes; (2) to better define diabetes type in SEARCH participants; and (3) to understand the evolution of insulin secretion in children with diabetes. C-peptide in serum or plasma is measured in the central laboratory by an in-house-developed radioimmunoassay [21]. Plasma or serum samples are incubated with limiting concentrations of an in-house-produced guinea pig antibody specific to human C-peptide and ^{125}I-labeled purified C-peptide. The C-peptide in patient samples and the ^{125}I-labeled C-peptide compete for the antibody binding sites. The decrease in the amount of ^{125}I-labeled C-peptide is proportional to the concentration of C-peptide in the samples. The antigen–antibody complex is precipitated by the addition of a goat antiguinea pig IgG and polyethylene glycol. Supernatants are decanted and the specific radioactivity in the immunocomplexes counted using an auto gamma counter (Packard Instruments). The concentration of C-peptide in the samples is determined using a standard curve generated with known concentrations of purified human C-peptide. The assay is linear up to a 5000 pg/ml concentration. To monitor the long-term consistency of the C-peptide results and to avoid assay drift, the laboratory has prepared in 1995 two lyophilized plasma pools representing high and low C-peptide reference ranges. Six replicates of these two samples are included in an assay at the end of each month, and obtained values are plotted against the expected values. Over 300 vials of each pool are stored at −80 °C and should last for an extended period of years. Assay precision is excellent, with a CV of 6.6 for the high quality control and 10.7 for the low quality control. Sensitivity limit for this assay is 0.15 ng/ml. The values are reported in nanograms per milliliter. These can be converted to nanomole per liter by multiplying with a factor of 0.333.

DNA and biologic samples are being stored for future studies designed to meet the primary and secondary aims of SEARCH only for participants for whom written informed consent and assent has been obtained according to local IRB guidelines.

1.7.4. Medical record review

For incident cases, medical records are reviewed to collect information on clinical presentation, initial clinical course, and utilization of health care services. Specific information regarding tests for DAA, C-

peptide, diabetes-related genes, diabetes complications and comorbidities, and prescription medications are recorded. Information is collected from all provider visits (inpatient and outpatient) that occur within 2 months prior to and up to 6 months following the diabetes diagnosis. For prevalent cases, an abbreviated medical record review is conducted if information is needed to establish diabetes type in untypeable cases (see Classification of Diabetes Type).

1.7.5. Annual follow-up

All incident cases, prevalent hybrid cases (defined below in Classification of Diabetes Type), and those with known genetic beta cell defects are followed with annual in-person visits. These include the physical examination and laboratory tests and questionnaire information about all factors that could change over time, except DAA and stimulated C-peptide testing in type 1A cases. All prevalent cases and all cases initially classified as other specific diabetes types are asked to complete an annual survey by mail. This survey gathers information on health care utilization and updates contact information to facilitate ancillary studies.

1.8. Classification of diabetes type

As recommended by the American Diabetes Association (ADA) Expert Committee on the Diagnosis and Classification of Diabetes Mellitus [22], SEARCH has developed a systematic approach to triage for classification of diabetes based on pathogenesis. The guiding principles for this triage system, described below and in Fig. 1, follow a hierarchical procedure. Data collected using SEARCH laboratory values will supercede those from other laboratories that will in turn supercede clinical data alone.

Initially, data measured in the SEARCH laboratory (DAA {antibodies to glutamic acid decarboxylase, IA-2, and insulin autoantibodies} and fasting plasma C-peptide concentration) define diabetes type, and an assignment to type 1A, type 1, type 2, or hybrid diabetes is made for incident and prevalent cases. All other cases in which this determination cannot be made are considered untypeable and investigated further as outlined below.

Stimulated C-peptide tests are done on prevalent cases >8 years that are initially untypeable and on cases with mixed features of autoimmunity and insulin resistance. The initial study plan was to use the same criteria for stimulated C-peptide testing in incident cases. During the second year of data collection, it was decided to invite all incident cases >8 years to undergo stimulated C-peptide testing, because there was an unexpectedly large number of incident cases that could not be categorized based solely on autoantibody data and fasting C-peptide.

Cases age >8 years old that are initially untypeable will undergo a stimulated C-peptide test using cut points that will assure making misdiagnosis highly unlikely [23]. Diabetes type will then be assigned by comparing the biochemical and clinical phenotypes of the untypeables to the biochemical and clinical phenotypes of two reference groups: (1) all laboratory-defined incident type 2 cases from SEARCH >8 years old, and (2) a sample of cases that are drawn from all incident SEARCH type 1A cases, which will be matched to untypeable cases by age and date of ascertainment of specimens. Subjects who decline measurement of DAA and C-peptide as part of the study but who have had DAA and/or C-peptide measured in the past in a non-SEARCH laboratory are classified as shown in Fig. 1.

Participants for whom only clinical data are available are classified based on the clinical phenotype defined in the reference group described previously. In those who still cannot be typed by this method, the following clinical definitions will be used: type 1 diabetes—diagnosis of diabetes made when the

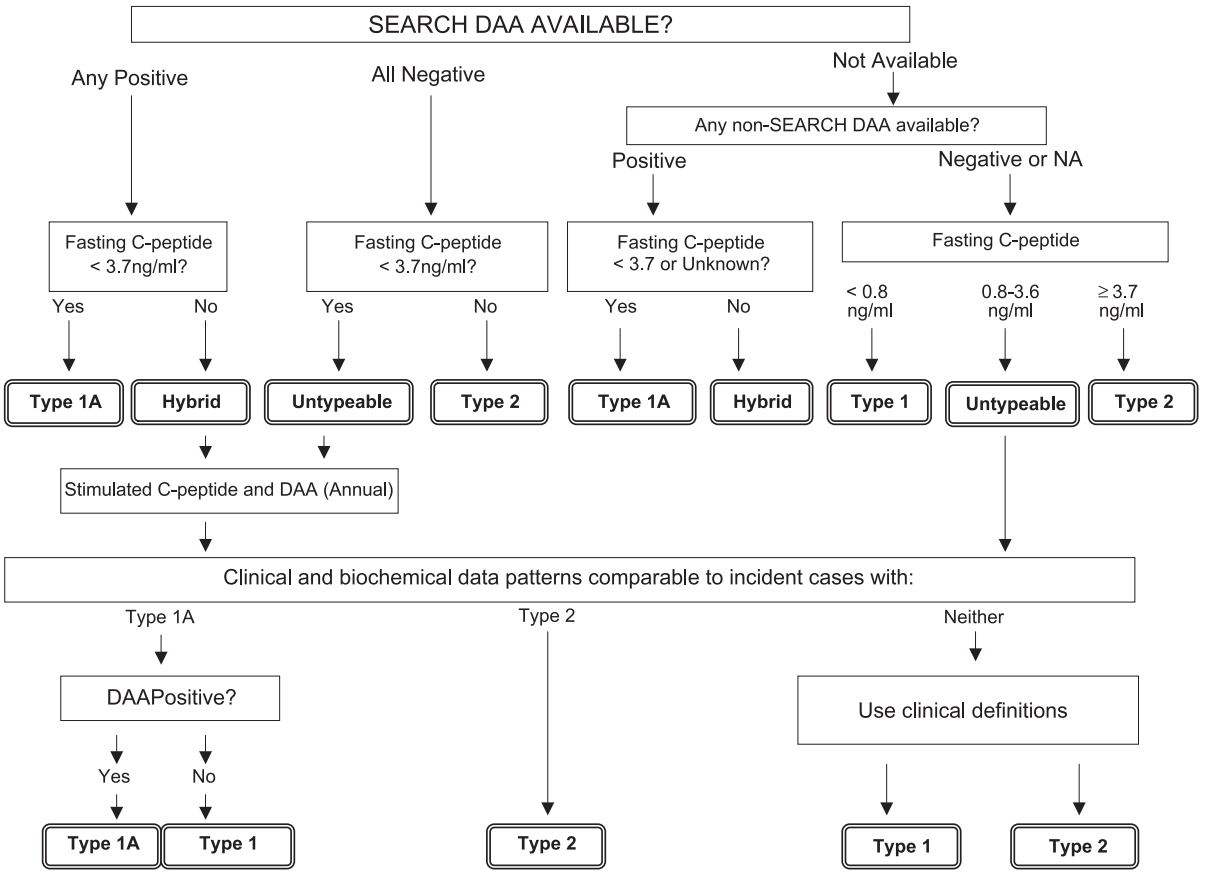

Fig. 1. SEARCH initial diabetes classification approach.

subject was <10 years old with a weight less than the 25th percentile for chronologic age or body mass index less than the 50th percentile for chronologic age at diagnosis; type 2 diabetes—duration of diabetes greater than 1 year and no insulin therapy for 1 month without an episode of diabetic ketoacidosis, or duration of diabetes greater than 6 months and never treated with insulin.

Subjects previously identified as having a genetic defect in beta cell functioning will undergo yearly DAA measurement and a stimulated C-peptide test. For subjects classified with other specific types of diabetes [22], except those with a previously diagnosed genetic defect in beta cell functioning, the type of diabetes is recorded and no further testing for typology performed.

As new information, such as new DAA or new markers of beta cell destruction, becomes available over the course of the study, additional testing will be performed, and the approach to classification of diabetes type will be modified to reflect the most accurate and current methods of classifying the types of diabetes.

1.9. Statistical considerations

Based on presumed negligible changes in population size and characteristics between 2000 and 2001, prevalence estimates are calculated per 1000 persons less than 20 years of age in 2001 by dividing the

total number of validated, eligible prevalent cases in 2001 by the total number of persons aged less than 20 years who were residents or members of the denominator population in 2000. Incidence estimates are calculated per 100,000 persons less than 20 years of age per year by dividing the total number of validated, eligible cases of diabetes with onset in the incidence year by the estimated total of persons resident or member of a given populations in the incidence year. Prevalence in 2001 and annual incidence are also calculated by type, age, race/ethnicity (non-Hispanic White, Hispanic, African-American, Asian, Pacific Islander, Native American) and gender. National estimates of prevalence and incidence are derived by applying the age, gender, and race/ethnicity specific estimates of prevalence and incidence derived from SEARCH to counts of the U.S. national population in 2000 (for prevalence) and projections of the U.S. national population for subsequent years.

The number of expected prevalent and incident cases cannot be estimated with precision using published data, especially considering the ethnic diversity of the base population. It is anticipated that there will be 6200 to 6800 prevalent cases and at least 800 incident cases annually.

Furthermore, SEARCH aims to develop efficient and practical approaches to classification of diabetes type for prevalent and incident cases (aim 2). An assessment of the accuracy of different approaches is based on misclassification ratios. The number of misclassifications is used to estimate the proportion of false positives, false negatives, sensitivity, and specificity. In evaluation of potential measures that are continuous or ordinal, receiver operator curves (ROC) is used to evaluate and test the usefulness of the diagnostic measure.

The third major aim of the study is to describe and compare clinical presentation and course of type 1, type 2, and other types or hybrids of diabetes. Patient characteristics and clinical presentation will be compared between types of diabetes. The statistical significance of these comparisons will be tested using the chi-square tests for categorical measures, Wilcoxon rank-sum tests for ordinal measures, and analysis of variance for continuous measures. Analyses are performed separately for prevalent and incident cases. Analysis of covariance procedures will be used to compare complications and risk factors for complications (e.g., hypertension, microalbuminuria, hyperglycemia) between types of diabetes adjusting for possible confounders (e.g., age, sex, race/ethnicity).

The clinical course of incident diabetes cases (acute complications, quality of life, insulin production, glycemia, lipidemia, development of microalbuminuria, presence of autoantibodies) is described using longitudinal data collected once per year. The statistical significance of differences in clinical course by diabetes type will be tested using repeated measures analysis of covariance and mixed models [24]. Maximum likelihood will be used to fit these models, because this increases precision and minimizes bias associated with varying lengths of follow-up among participants.

2. Discussion

SEARCH is uniquely poised to generate vital information required to develop clinical interventions and public health policies designed to reduce the incidence and improve the outcomes of diabetes in youth. The six study centers provide a study population that is larger and more diverse in terms of race/ethnicity, geography, and age than any previous study of diabetes in children. Importantly, this study systematically employs a uniform methodology of diabetes classification based on pathogenic criteria, reflecting the recommendations of the American Diabetes Association (ADA) Expert Committee on the Diagnosis and Classification of Diabetes Mellitus [22]. The use of this systematic method of

classification with this large and diverse study population will enable this study to achieve its primary objective, to provide accurate estimations of the incidence and prevalence of diabetes by type, age, gender and race/ethnicity, and enhances the generalizability of this approach in other settings. The study is powered to identify differences in the presentation and course of each type of diabetes by age, sex, and race/ethnicity if they are present. In addition, the study is designed to facilitate the longitudinal study of this pediatric population, which, if further funding is obtained, has the potential to illuminate factors associated with the onset and progression of diabetes-related complications. Finally, DNA and biological samples obtained by this study will provide an invaluable resource for genetic and biochemical studies.

SEARCH does however have limitations. The study will determine the prevalence and incidence of diagnosed diabetes only. No attempt will be made to determine how much undiagnosed diabetes exists in youth or whether undiagnosed cases vary by age or ethnicity. While it is recognized that type 2 diabetes may be present for years prior to diagnosis, SEARCH will not screen for undiagnosed cases. Growing awareness of the presence of type 2 diabetes among youth may influence screening or diagnostic approaches by health care providers over time. In the absence of an extension beyond the year 2005, the study will have a limited ability to address temporal trends in the incidence of diabetes given that it will only encompass 3 years.

SEARCH is also faced with a number of practical challenges. The completeness of case ascertainment depends on the cooperation of multiple health care providers and organizations. To assure complete cases ascertainment, all potential sources of care for children with diabetes have been enumerated in the geographic sites, and the study monitors the ability of sites to obtain data from these sites. Linkage of computer-stored records from a large number of sources (pharmacy, hospitalizations, outpatient records, laboratory) is used in an attempt to assure complete case ascertainment in the membership-based sites. Capture–recapture methods are being explored as a way to assess completeness of case ascertainment in settings where all potential cases that derive from each data source can be determined to be validated, unique cases of diabetes or ruled out as validated unique cases and where reasonable independence of information from the various sources can be assured. Concern about the completeness of case ascertainment is unlikely to be completely laid to rest.

The recruitment of this study population also presents some unique obstacles. The study population is by definition a vulnerable population due to the age composition. It is anticipated that there will be missing data resulting from patients who refuse enrollment. This is particularly problematic in calculating prevalence and incidence, and distribution of diabetes type, if refusal is more common among specific groups of youth (e.g., ethnic minorities versus whites, older versus younger, girls versus boys). Although the study is likely to yield substantial benefits in terms of the knowledge it will generate, the benefit to individual participants is somewhat limited, and participation entails a significant amount of time and minimally invasive procedures.

Issues of data confidentiality are addressed in the study design by reducing the data required to document a case to a minimal set of variables without individual identifiers. The study design also incorporates the use of multiple data sources as needed and as possible at each of the sites, including information from parents and patients, health care providers, medical records, and administrative health care data, in an effort to reduce the amount of missing data.

The recent marked increases in type 2 diabetes in youth are occurring in parallel with increases in adults and appear to be directly related to nationwide increases in obesity in all ages and ethnic groups. Reasons for the regional increases in type 1 diabetes have not been clearly identified. The SEARCH

study will increase our understanding of the public health burden of diabetes for youth, who are at high risk for long-term complications from this disease.

Acknowledgements

SEARCH for Diabetes in Youth is funded by the Centers for Disease Control and Prevention (PA number 00097) and supported by the National Institute of Diabetes and Digestive and Kidney Diseases.

References

[1] Rewers M, LaPorte RE, King H, Tuomilehto J. Trends in the prevalence and incidence of diabetes: insulin-dependent diabetes mellitus in childhood. World Health Stat Q 1988;41:179–89.
[2] Fagot-Campagna A, Pettitt DJ, Engelgau MM, Burrows NR, Geiss LS, Valdez R, et al. Type 2 diabetes among North American children and adolescents: an epidemiologic review and a public health perspective. J Pediatr 2000;136:664–72.
[3] Rosenbloom AL, Joe JR, Young RS, Winter WE. Emerging epidemic of type 2 diabetes in youth. Diabetes Care 1999;22:345–54.
[4] Dabelea D, Pettitt DJ, Jones KL, Arslanian S. Type 2 diabetes mellitus in minority children and adolescents: an emerging problem. In: Arslanian S, editor. Endocrinology and metabolism clinics of North America, vol. 28. 1999. p. 709–29.
[5] EURODIAB ACE Study Group. Variation and trends in incidence of childhood diabetes in Europe. Lancet 2000;355: 873–6.
[6] Onkamo P, Vaananen S, Karvonen M, Tuomilehto J. Worldwide increase in incidence of Type I diabetes—the analysis of the data on published incidence trends. Diabetologia 1999;42:1395–403.
[7] Lipton R, Keenan H, Onyemere KU, Freels S. Incidence and onset features of diabetes in African-American and Latino children in Chicago, 1985–1994. Diabetes Metab Res Rev 2002;18:135–42.
[8] Ferrara A, Imperatore G, Quesenberry CP, Karter AJ, Fagot-Campagna A, Ergas IJ, et al. Characteristics and quality of health care of type 1 and type 2 diabetic youths: the Northern California Kaiser Permanente Diabetes Registry. Diabetes 2002;51(Suppl. 2):A26.
[9] Barrett TG, Ehtisham S, Smith A, Hattersley AT. UK Diabetes Survey shows type 2 diabetes present in 0.4% of newly diagnosed children, associated with overweight, female, and ethnic minority status. Diabetes 2002;51(Suppl. 2):A25.
[10] American Diabetes Association: report of the expert committee on the diagnosis and classification of diabetes mellitus. Diabetes Care 1997;20:1183–97.
[11] Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. Diabet Med 1998;15:539–53.
[12] US Census Bureau. United States Census 2000, available at: http://www.census.gov/).
[13] Fienberg SE, Johnson MS, Junker BW. Classical multilevel and Bayesian approaches to population size estimation using multiple lists. J R Stat Soc, Ser A 1999;162:383–406.
[14] Verlato G, Muggeo M. Capture–recapture method in the epidemiology of type 2 diabetes. Diabetes Care 2000;23:759–64.
[15] Capture–recapture and multiple-record systems estimation. I: history and theoretical development. International Working Group for Disease Monitoring and Forecasting. Am J Epidemiol 1995;142:1047–58.
[16] Hook EB, Regal RR. The value of capture–recapture methods even for apparent exhaustive surveys. The need for adjustment for source of ascertainment intersection in attempted complete prevalence studies. Am J Epidemiol 1992;135:1060–7.
[17] Falorni A, Ortqvist E, Persson B, Lernmark A. Radioimmunoassays for glutamic acid decarboxylase (GAD65) and GAD65 autoantibodies using ^{35}S or ^{3}H recombinant human ligands. J Immunol Methods 1995;186:89–99.
[18] Marcovina SM, Landin-Olsson M, Essen-Moller M, Palmer JP, Lernmark A. Evaluation of a novel radioimmunoassay using ^{125}I labeled human recombinant gad65 for the determination of glutamic acid decarboxylase (GAD65) autoantibodies. Int J Clin Lab Res 2000;30:21–6.

[19] Gruben CF, Daniels T, Toivola B. A novel radioligand binding assay to determine diagnostic accuracy of isoform specific glutamic acid decarboxylase antibodies in childhood IDDM. Diabetologia 1994;37:344–50.
[20] Yu L, Robles DT, Abiru N, Kaur P, Rewers M, Kelemen K, et al. Early expression of antiinsulin autoantibodies of humans and the NOD mouse: evidence for early determination of subsequent diabetes. Proc Natl Acad Sci 2000;97:1701–6.
[21] Heding LG. Radioimmunological determination of human c-peptide in serum. Diabetologia 1975;11:541–8.
[22] American Diabetes Association. Report of the Expert Committee on the diagnosis and classification of diabetes Mellitus. Diabetes Care 2000;23:S4–19.
[23] Allen AF, Klingensmith GJ, Jensen P, Simoes E, Hayward A, Chase HP. Effect of Bacillus Calmette–Guerin vaccination on new-onset type 1 diabetes. A randomized clinical study. Diabetes Care 1999;22(10):1703–7.
[24] Laird NM, Ware JH. Random-effects models for longitudinal data. Biometrics 1982;38:963–74.