SEARCH FOR DIABETES IN YOUTH STUDY

OMB SUPPORTING STATEMENT: PART B

Submitted by:

Sharon Saydah, PhD MHS

Project Officer

Epidemiology and Statistics Branch

Division of Diabetes Translation

Centers for Disease Control and Prevention

Atlanta, GA 30341

Telephone: (301) 458-4183

Fax: (301) 458-4025

e-mail: ssaydah@cdc.gov

May 4, 2011

Table of Contents

Attachments

B.1. Respondent Universe and Sampling Methods

Registry Study- *Case Ascertainment Processes utilized by the SEARCH 3 Clinical Sites*

Ongoing Case Ascertainment: SEARCH 3 will to use the reporting network of clinics and health care providers that has been established as the primary approach to case-finding for incident cases of diabetes for the period 2010-2014. Additionally, the case ascertainment approach involves existing validated pediatric diabetes databases, hospital and health plan databases, and other health care organizations.

Case Validation: Cases of diabetes will be validated based on physician reports, medical record reviews, or self-report of a physician diagnosis of (non-gestational) diabetes. A physician-diagnosed case of diabetes is established if any of the following criteria are met: (1) medical record review indicating a physician diagnosis of diabetes, (2) the diagnosis of diabetes is directly verified by a physician, (3) the physician referred a youth with diabetes to the study, or (4) the case was included in a clinical database that had a requirement for verification of diagnosis of diabetes by a physician.

Eligibility Criteria: Eligibility criteria will remain the same. As in SEARCH 1 and 2, the study will be confined to children/youth who, in addition to having an onset of physician-diagnosed of diabetes the index year, are also are < 20 years of age on December 31 of the index year and 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 membership-based centers at diagnosis, and 3) are not active duty military personnel or institutionalized. Protected Health Information (PHI) will be obtained in order to validate and confirm eligibility and uniqueness of cases in keeping with HIPAA and the procedures and approvals required by the local IRB.

De-duplication: Duplicates will be identified using both electronic files and manually, both within and between case sources, using the name or initials, gender, date of birth, ethnicity, zip code, or other available information, in keeping with HIPAA requirements to use the least amount of PHI in conducting research. The number of duplicates identified will be used to estimate completeness of ascertainment with the capture-recapture method among the geographic centers.

Systematic re-ascertainment: A systematic case re-ascertainment process will be conducted in 2012 for cases incident in 2006-2010 and in 2014 for 2008-2013 cohorts. The purpose of re-ascertainment is to assure as complete ascertainment of diabetes cases as possible, especially those with type 2 diabetes, who, due to different care patterns, are more likely to be missed by the SEARCH network prior to the close of the registration window. Additional cases will be registered with the center and the Coordinating Center with a flag to indicate that they were registered outside of the standard window.

Case Registration: Cases that are valid, eligible and unique will be registered by the center with information being uploaded to the Coordinating Center. Names, addresses, date of diagnosis and date of birth are not provided to the Coordinating Center. In cases where duplicates and cases that are not valid or eligible are identified at a later date, they

will be unregistered by both the local center and the Coordinating Center.

Cohort Study: Eligibility and Anticipated Sample Size

Individuals registered in the SEARCH cohort with a diagnosis of non-gestational diabetes in incident years 2002-2005, 2006 and 2008 and who completed a SEARCH baseline in-person visit in five of the SEARCH clinical centers will be eligible for the new SEARCH Cohort Study once they have had diabetes for at least 5 years. We will determine an algorithm that optimizes duration of disease over the data collection years of the Cohort Study, considering both diagnosis year and the timing of the most recent SEARCH visit during the SEARCH 2 protocol. We anticipate that the average duration of diabetes will be 8-9 years (range 6-14 yrs diabetes duration), and the average age of participants at the time of the SEARCH Cohort Visit will be about 20 years (range 8-32 yrs old). Our goal is to have 80% of eligible members of the cohort attend the proposed SEARCH Cohort Study visit, for a total of 3145 participants with a Cohort Study visit.

Of the 20% not attending the research clinic visit, an additional 10% will be participate by completion of surveys by telephone or internet, for a total of 90% inclusion or 3550 individuals. Based on proportions of youth in sub-groups of diabetes type and race/ethnicity, the anticipated numbers and characteristics of anticipated participants in the SEARCH Cohort Study data collection are displayed in Table 1.

Table 1. Anticipated Sample Size and Characteristics: SEARCH Cohort Study (2002 - 2006, 2008)

	SC		OH		CO		CA		WA		All	
Type Race/Ethnicity	80%	90%	80%	90%	80%	90%	80%	90%	80%	90%	80%	90%
Type 1												
NHW	405	456	386	434	711	800	98	110	410	461	2010	2261
Other	175	197	58	65	188	211	185	209	94	106	700	788
Type 2												
NHW	38	42	30	34	16	17	7	8	16	18	107	119
Other	106	120	42	47	79	95	83	97	18	23	328	382
ALL	724	815	517	581	999	1124	377	424	541	608	3145	3550

B.2. Procedures for the Collection of Information

Registry study

Collection of Core Variables: A minimum amount of demographic and clinical information is needed for all registered cases in order for the study to be able to provide population-based rates of diabetes mellitus by age, gender, diabetes type and

race/ethnicity for the entire population of cases. This information is also critical in assessing possible response bias to the in person research visit.

Medical Record Abstraction (MRA) serves the following purposes: a) validation of diabetes diagnosis; b) main source of core demographic and diagnostic information, and c) secondary data source for race/ethnicity information. In SEARCH 2, an additional set of items pertinent to clinical presentation was added to the medical record abstraction effort: weight/height at diagnosis, DKA at diagnosis and insulin use history. We will continue to collect these data through MRA in SEARCH 3 and will seek 100% completion.

Initial Patient Survey (IPS) contains key data, including the core information described above, and serves to: a) verify of case eligibility (e.g., residence in the year of diagnosis); and b) is the main source for self-reported race/ethnicity information. Additional information includes: symptoms at presentation, potential secondary causes of the diabetes, use of insulin and other medications, treatment history, family structure, usual language spoken, and contact information (for local use only).

In-Person Research Visit (IPV) is designed to collect data on relevant dimensions of diabetes type (presence of autoimmunity, genetic susceptibility to autoimmunity, insulin sensitivity, insulin secretion) and data informing the clinical presentation of diabetes. The following will stored for future analyses (by separate consent) blood, serum, plasma and urine for future genetic and non-genetic analyses. Only cohorts incident in 2011 will be eligible to participate in the IPV. An additional sampling approach will be implemented in SEARCH 3, in order to reduce participant burden and maximize study resources, without compromising the statistical power to detect trends in clinical characteristics over time. To maximize the number of minority participants and youth with T2D, eligible cases for SEARCH 3 IPV are 100% of minority (non-Caucasian) youth, regardless of age, and 100% of youth age ≥10 years at diagnosis, regardless of race/ethnicity, but only 50% of Non-Hispanic white youth with onset age < 10 years. We will seek a 70% completion of the IPV among eligible youth.

Cohort study

Risk Factors for Diabetes Complications (details on measures Appendix B.1)

Laboratory measures:

- Laboratory assays will measure markers of autoimmunity (diabetes autoantibodies, GAD65, IA2; plus the addition of the recently identified ZnT8 antibody once assays are standardized), fasting c-peptide, fasting glucose, hemoglobin A1c, lipid profile (total, LDL- and HDL-cholesterol, and triglycerides), apolipoprotein B, LDL particle size and density, adipocytokines (CRP, IL-6, fibrinogen, adiponectin, leptin).

Physical examination measures:

- Standardized anthropometry methods include height, weight, waist circumference
- Systolic and diastolic blood pressure

- Evaluation for acanthosis nigricans
- Assessment of health behaviors(including dietary intake by food frequency, physical activity, alcohol, and active and passive smoking)
- Retinopathy
- Nephropathy
- Vascular dysfunction (marker of macrovascular disease)
- Markers of neuropathy
- Acute complications (hypoglycemia and diabetic ketoacidosis)

.

Diabetes Treatment & Psychosocial Factors:

- Diabetes treatment regimen and related technologies
- Psycho-social factors
- Socio-cultural factors (household income, per capita income, family structure, preferred language, parental and participant-attained education.
- Processes to care and barriers to care

B.3. Methods to Maximize Response Rates and Deal with Nonresponse

Registry study

Methods for case re-ascertainment and capture-recapture methods are utilized to maximize case registration.

Cohort study retention strategies

Retention strategies will include traditional, proven, cohort retention strategies such as: birthday cards, study newsletters, updating contact information annually, utilizing internet-based search systems to locate individuals lost to follow-up, using cell-phone text messaging and e-mail, offering flexible study date appointments including home visits, offering assistance with transportation, mailing pre-visit instructions, a reminder call prior to the visit, acknowledgement of participation, and participant remunerations that are appropriate for the length and the respondent burden of the proposed study visit. Investigators and study personnel will 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.

B.4. Tests of Procedures or Methods to be Undertaken

Incidence Rate Estimation

Because all SEARCH Cohort Study participants will have had at least one previous SEARCH in-person visit, we will be able to define a group of participants who were free from the event of interest (i.e., normotensive) at “baseline”. Multiple logistic regression methods will be employed to examine the incidence rates of binary measures (e.g.,

hypertension) of interest. Predictors can include categorical or continuous variables. A continuous variable that measures the time between visits for each participant (to account for the fact that individuals will have different lengths of follow-up) and the predictor- by- time interaction will be included. Next, we will expand the logistic regression model to include other participant level characteristics (e.g., SEARCH clinical center, age, and gender (a "demographically adjusted model")). We will then expand the model to adjust for other covariates. In addition, we will examine potential interactions; if significant interaction is present, analyses will be performed stratified by that characteristic.

Prevalence Estimation

Some of the outcomes of interest will not have been measured during SEARCH 1 or 2, such as outcomes including retinopathy and neuropathy. Therefore, prevalence of these outcomes will be estimated. Models to evaluate cross-sectional associations of risk factors will use logistic regression and will proceed as described above to account for potential confounding or effect modification.

Longitudinal Models

All participants in the SEARCH Cohort Study will have already had at least one in-person visit during SEARCH 1 and 2, and ~75% of the 2002-2005 incidence cases have at least 2 in-person visits per the SEARCH 2 protocol. Since SEARCH 2 also included longitudinal data (there are over 2000 SEARCH participants already with at least one follow-up visit), our team developed a plan for modeling longitudinal data. Specifically, we will use longitudinal mixed effects analysis of covariance models that always include duration of diabetes as a time-varying covariate. This approach correctly models the varying durations of disease prior to the initial SEARCH in-person visit, and the varying durations of time allowed via the SEARCH data collection windows between the initial and subsequent visits.

The initial model will examine outcomes (measured previously between 1 (baseline) and 4 times (baseline, 12, 24, 60 mo visits) and once during the SEARCH Cohort Study visit), the predictor of interest (e.g., DM type), the duration of diabetes at each measurement time and the predictor-by-diabetes duration interaction. These models will then be expanded to include demographic information (e.g., sex) that would be considered as fixed/non-time varying effects. In addition, based on our experience with performing these longitudinal analyses on the SEARCH 2 cohort, we also propose to consider treating the exposure (predictor) of interest as a time-varying covariate in these models as well. This will allow the time-varying correlation of the predictor to the outcome of interest to be modeled correctly. We will also consider adding other time-varying covariates (e.g., BMI z-score) into these models as needed to examine the specific relationships being studied. These mixed effects models also are flexible to allow for potentially non-linear relationships to be modeled over time, and permit random rates of progression, consistent with a perspective that different participants progress through time at different rates. Use of random intercepts and/or slopes provides a source of autocorrelation between repeated measures. More flexible structures for the

correlation between repeated measures will be investigated using combination mixed models that allow the specification of separate parameters representing variation between experimental units, and serial correlation within units. Our choice of methods for accounting for serial correlation depends on the plausibility of the model, and the number of outcomes relative to the number of participants. For example, with many participants and few repeated measurements, an unstructured covariance matrix can often provide for the most efficient estimation of model parameters.

For analysis of longitudinal discrete outcomes (e.g. transfer of care from a pediatric to adult provider), we will use the generalized estimating equation (GEE) approach to fit logistic or log-linear models that account for the dependency between repeated measures. GEE techniques allow estimation of model parameters and their standard errors from longitudinal data having continuous and categorical responses and potentially missing observations. An advantage of this technique is that the assumptions required are weaker than those of maximum likelihood techniques: one need not specify the distribution of the dependent variable, just the relationships between the marginal mean and variance, and between the marginal mean and covariates.

B.5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

CDC will consult with the SEARCH Coordinating Center, SEARCH clinical sites and CDC partners. The SEARCH Clinical Sites are responsible for the data collection from the participants. The Coordinating Center is responsible collecting the data from the Clinical Sites. Data management and analysis will be performed by the SEARCH Coordinating Center at Wake Forest University. Specific data analysis plans are developed in collaboration with the SEARCH Clinical Sites, the CDC and the Coordinating Center

Ronny Bell, PhD. and Ralph D'Agostino, PhD.
Principal Investigators for the Coordinating Center, Responsible for overseeing the data collection from the SEARCH Clinical sites and the data analysis.
Wake Forest University School of Medicine– Coordinating Center
Wake Forest University Health Sciences Medical Center Blvd.
Winston-Salem, NC 27157
rbell@wfubmc.edu
336-716-9736

Ralph D'Agostino, PhD,
Wake Forest University School of Medicine– Coordinating Center
Wake Forest University Health Sciences Medical Center Blvd.
Winston-Salem, NC 27157
336-716-9410
rdagosti@wfubmc.edu

Sharon Saydah, PhD
Senior Scientist
LCDR USPHS
Centers for Disease Control and Prevention
Division of Diabetes Translation
Koger Center, Williams Blg
Atlanta, GA 30341
301-458-4183
ssaydah@cdc.gov

Giuseppina Imperatore, MD PhD
Epidemiology Team Lead
Centers for Disease Control and Prevention
Division of Diabetes Translation
Koger Center, Williams Bldg.
Atlanta, GA 30341
770-4888-5821
Gai5@cdc.gov