Candidate Gene Association Resource (CARe) Project
Database Application

NHLBI CARe Project Number: TBD

Submission Date: July 18, 2012

All proposals must be submitted by email to the CARe DAC COORDINATOR at care-proposals@broad.mit.edu

All sections of this application must be completed. Incomplete applications will be returned.

Sections I-IV should be no more than 3 pages in length, not including references.

I. INVESTIGATOR INFORMATION:

Title of Proposed Project: Does variation in CYP2A6 mediate the relationship between cigarette smoking, incident type 2 diabetes, and glycemic traits?: the Candidate Gene Association Resource (CARe) Project

Name of proposed *Lead Author: Jason L. Vassy, MD, MPH, SM
Email Address: jvassy@partners.org
Telephone Number: (314) 239-3741
Fax Number: (617) 732-7072
Mailing Address: Division of General Medicine, Brigham and Women’s Hospital, 1620 Tremont Street, BC-3, Boston, MA, 02120

Name of proposed *Senior Author: TBD by the consortium writing group.
Initially Senior and ultimately responsible for the project: James B. Meigs MD, MPH
Email Address: jmeigs@partners.org
Telephone Number: (617) 724-3203
Fax Number: (617) 724-3544
Mailing Address: General Medicine Division, Massachusetts General Hospital, 9th floor, 50 Staniford Street, Boston, MA 02114

Institution/Company of Principal Investigator: Brigham and Women’s Hospital

Name of Lead Analysts / Statisticians: Ching-Ti Liu, PhD, and Josée Dupuis, PhD
Josée Dupuis, PhD:
Email Address: dupuis@bu.edu
Telephone Number: (617) 638-5880
Fax Number: (617) 638-5299
Mailing Address:
Josée Dupuis, PhD
Dept of Biostatistics, BUSPH
715 Albany Street, Talbot Building
Boston, MA 02118
II. SCIENTIFIC RATIONALE (~250 WORDS)

A. Please provide an abstract describing the rationale and design of the proposed research project. The abstract must include major hypotheses, an outline of the research methods and analytical approach, and phenotypes to be studied. It should state clearly the objectives of the proposed project and provide the background rationale that would justify them. It should also address why the CARe database is appropriate for answering the research question. (Abstracts for approved projects will be posted, with the name of the Principal Investigator, on the NHLBI/NCBI CARe website.)

Cigarette smoking is an independent risk factor for incident type 2 diabetes (T2D). A meta-analysis of 25 prospective studies by Willi, et al., found that smokers had an adjusted risk ratio for incident T2D of 1.44 (95% CI 1.31, 1.58) over a follow-up of 5 to 30 years in adulthood. This relationship appears to be dose-dependent\(^1\) and may be mediated by nicotine. Nicotine might increase abdominal adiposity and insulin resistance while decreasing insulin secretion from pancreatic beta-cells\(^2\). The relationship between smoking and T2D has considerable heterogeneity. In the meta-analysis by Willi, the risk ratio for T2D comparing active smokers to non-smokers ranged from 0.82 to 3.74 across the 25 studies, and statistical tests of heterogeneity were significant (Q statistic, 98.08, P<0.01; \(I^2\) 75.5%)\(^1\). This variation suggests the presence of modifiers of the effect of smoking on T2D risk.

Genetic variation might be such an effect modifier. The cytochrome p450 (CYP) 2A6 enzyme metabolizes the majority of inhaled nicotine to cotinine in the liver\(^3\). Numerous \(CYP2A6\) polymorphisms have been identified, many of which result in slower nicotine metabolism\(^4\). A recent cross-sectional study in Chinese smokers found that heavy smokers with \(CYP2A6\) polymorphisms known to be associated with intermediate, slow, and poor nicotine metabolism had adjusted odds ratios for T2D of 2.44 (95% CI 0.59,10.09), 5.12 (95% CI 1.08, 24.23), and 8.54 (95% CI 1.28, 57.02), respectively, compared to light smokers with normal \(CYP2A6\) metabolism\(^2\). This cross-sectional analysis, however, offers only weak evidence for a causal pathway linking smoking to T2D. Moreover, genetic loci other than \(CYP2A6\) may contribute to the heterogeneity in the relationship between smoking and T2D.

We propose using data from the CARe consortium of prospective cohort studies to test 2 primary hypotheses:

1. \(CYP2A6\) variants modify the effect of smoking on T2D risk and fasting insulin and fasting glucose (glycemic traits):
   a. \(CYP2A6\) variants do not affect T2D risk among non-smokers. Among smokers, \(CYP2A6\) variants resulting in poorer nicotine metabolism are associated with greater T2D risk from smoking, compared to \(CYP2A6\) variants causing normal nicotine metabolism.
b. Among non-diabetics, CYP2A6 variants do not affect glycemic traits in non-smokers. Among smokers, CYP2A6 variants resulting in poorer nicotine metabolism are associated with a greater increase in glycemic traits from smoking, compared to CYP2A6 variants causing normal nicotine metabolism.

2. Beyond CYP2A6, there are other genetic loci that mediate the relationship between smoking and T2D risk and glycemic traits (genome-wide gene-by-smoking interaction analyses).

The CARe database is an ideal setting in which to test these hypotheses. It consists of several well-characterized cohorts (ARIC, CARDIA, CHS, FHS, JHS and MESA) with T2D prevalent and incident phenotypes, smoking status, and GWAS array data. Through the CARe consortium, limited data sets from these cohorts have been combined to enable genetic analyses for which the individual studies would not have enough power to undertake. The authors already have limited CARe data sets through prior approved CARe proposals. For the present proposal, we will pool the data from the participating CARe studies to test the hypotheses above.

To address Specific Aim 1a, we will model time-to-incident T2D in these pooled data as a function of baseline smoking status, CYP2A6 genotype (as previously categorized\(^2\)), and the interaction between them, adjusted for covariates including age, sex, and study. These models will be performed in non-smokers alone and then in the overall cohort, and each model will be performed in whites and blacks separately. For Specific Aim 1b, cross-sectional adjusted linear regression models will similarly model fasting insulin and fasting glucose as a function of smoking status, CYP2A6 genotype, and their interaction. The analytic plan for Specific Aim 2 will be patterned on recent CARe analyses by Nora Franceschini, in which genome-wide association analyses were performed on study- and race-specific residuals from regression models for quantitative renal traits as a function of age, sex, and study site, stratified by smoking status. To detect SNP-by-smoking interaction, heterogeneity between smoking strata was assessed by a heterogeneity test and \(I^2\) metric.

III. PRIOR EXPERIENCE OF THE PI AND ASSOCIATES (~250 WORDS)

Please describe the experience and expertise of your team to complete the research project.

The proposed first author (Dr. Vassy) has recently completed a research fellowship in general medicine at Massachusetts General Hospital as a mentee of the proposed senior author (Dr. Meigs). Pertinent to this proposal, Dr. Vassy currently has two publications in press examining genetic prediction models for T2D, one using data from the CARDIA study\(^6,7\). Dr. Meigs has co-authored more than 250 peer-reviewed papers in the fields of diabetes, cardiovascular, and genetic epidemiology. Drs. Liu and Dupuis has considerable expertise in analyzing genetic data in the Framingham Heart Study, and Dr. Liu has recently assisted in the analysis of CARe data to look at genotype-by-smoking interactions in renal quantitative traits (CARe proposal by Nora Franceschini referenced above).

IV. DATA REQUESTED FOR THE PROPOSED ANALYSES (Provide rationale for any requested data whose relevance to these analyses is not obvious):

**Genotype data (check all that apply):**

[ ] CARe IBC Candidate Gene array

[ ] Affymetrix 6.0

[X] Both IBC array and Affymetrix 6.0

[ ] Other (please specify; e.g. SNPs within candidate genes, specific regions, etc.)

**Phenotype data:**
From each CARe cohort we request the following traits (if available, as already defined by the Diabetes Phenotype Group):

1. Sex
2. Age
3. Smoking status
4. BMI (derived from weight and height)
5. Waist circumference
6. Diabetes diagnosis (yes/no)
7. Hypertension diagnosis (yes/no)
8. Medication: diabetes oral medication (yes/no)
9. Medication: diabetes insulin therapy (yes/no)
10. Medication: hypertension medication (yes/no)
11. Medication: cholesterol medication (yes/no)
12. Family history of diabetes (yes/no)
13. Fasting glucose (mmol/l)
14. Fasting insulin (pmol/l)
15. HOMA-IR (derived from fasting glucose and insulin)
16. HOMA-B (derived from fasting glucose and insulin)
17. HbA1c (%)
18. Fasting triglycerides
19. Total cholesterol
20. HDL cholesterol
21. LDL cholesterol
22. Physical activity
23. Dietary variables including food frequency questionnaire data or 24-hour recall data, as available

V. ORGANIZATION APPLYING FOR DATA ACCESS:

Name of Applicant Organization: Brigham and Women’s Hospital

Address of Applicant Organization: 1620 Tremont Street, BC-3, Boston MA, 02120

Name and Title of Applicant Organization’s Authorized Institutional Business Official:

Karen Lodigiani, Director of Contracting, Partners of Clinical Research Office

Name of Each Institution/Company whose investigators will receive access to the requested data:

Brigham and Women’s Hospital, Massachusetts General Hospital, and Boston University School of Public Health

I, Jason L. Vasy, PI for the proposed project, request the following data:

a. [x] Yes [ ] No: Genotype Data
b. [x] Yes [ ] No: Phenotype Data

Please answer the following questions:

a. [ ] Yes [x] No: This research using the CARe database may be used for development of a commercial product or for commercial purposes.
b. [ ] Yes [x] No: Data accessed through this application will be used by or shared with individuals from a for-profit company.

NOTE: Some participants in the CARe Study have not provided consent for use of their CARe data for development of a commercial product or to be accessed by a for-profit entity. For each approved project, a dataset will be generated by a computer algorithm that incorporates the consent options of each of the CARe Cohorts, such that data access will only be provided as specified by the informed consent document of each study participant.

VI. ANNUAL REPORTING OF RESULTS

Continued access to the CARe database after one year will only be permitted if an annual report is submitted to the DAC Coordinator (care-dac@nhlbi.nih.gov) that describes the product of your research using the CARe database, and includes a listing of presentations and publications resulting from that research. Please complete the following:

[x] Yes [ ] No: I understand that use of the CARe dataset includes the timely completion of an annual report.

Complete results of CARe analyses are being web-posted by the National Center for Biotechnology Information (NCBI). Please complete the following:

[x] Yes [ ] No: I am willing to provide my completed analysis results for such web-posting.

VII. ADDITIONAL DOCUMENTS:

Return by email to the DAC Coordinator: care-proposals@broad.mit.edu

(Signed or other non-electronic documents can be faxed.)

Please include:

(1) Application Form

(2) Supporting documentation, including current human studies training certification for all key personnel*

(3) IRB Approval (Note: Full board or expedited review is required; “exempt” status is not acceptable.)*

(4) Data Distribution Agreement*

*Can be submitted after DAC Committee approval

Applications can be submitted at any time. Applicants will be notified of the DAC Committee decision within four weeks.

Upon approval, a Data Recipient at each institution that will receive data will be required to submit a signed Data Distribution Agreement to the CARe DAC Coordinator. The CARe Data Distribution Agreement can be obtained from Website to be added. The Data Recipient(s) will be responsible for obtaining signatures on behalf of the Recipient Entities and returning the signed Data Distribution Agreement to the CARe DAC Coordinator before access can be provided to the CARe database.

VII. ADDITIONAL INFORMATION (Required only for “CARe Investigators”)
Name and CARe Cohort Affiliation(s) of proposed Lead Author(s): N/A

Names and CARe Cohort Affiliation(s) of proposed Senior Author(s): James B. Meigs, FHS

Table of CARe Cohort Representation: Please see the Table in: “Name(s) of all other Professional Participants.”

References