ARIC Manuscript Proposal #2379

PC Reviewed: 6/10/14	**Status: A**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: **Genome-wide association study (GWAS) meta-analysis for gallstone disease in CHARGE**

b. Abbreviated Title (Length 26 characters): Gallstone GWAS in CHARGE

2. Writing Group: CHARGE Gallstone working group

Writing group members: Weihong Tang, Pamela L. Lutsey, Lu-Chen Weng, and Aaron R. Folsom, others welcome. Other authors from additional CHARGE cohorts.

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. __WT__ **[please confirm with your initials electronically or in writing]**

First author:
Address: Division of Epidemiology and Community Health
University of Minnesota

Phone: 612-626-9140 Fax: 612-624-0315
E-mail: tang0097@umn.edu

ARIC author to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (this must be an ARIC investigator).
Name: Aaron Folsom
Address: same as above

Phone: 612-626-8862 Fax: 612-624-0315
E-mail: folso001@umn.edu

3. Timeline: Finish by Fall 2014

4. Rationale:

Gallstone disease represents one of the most common digestive disorders in the US and worldwide.[1] Gallstone disease is influenced by genetic factors, with up to 25% of risk attributable to genetic influence.[2] In 2007, a genome-wide association study (GWAS) consisting of 280 cases identified the hepatic cholesterol transporter *ABCG8* locus as a

susceptibility factor for human gallstone disease.[3] Additional genes likely contribute to the risk. No other population-based GWAS has been reported for this phenotype. CHARGE is doing a meta-analysis of GWAS findings related to gallstone disease (please see listed cohorts below). ARIC data analysis will take place in Minnesota. A meta-analysis will be conducted by Amit D. Joshi, Charlotte Andersson, and Andrew Johnson at Harvard and NIH.

5. Main Hypothesis/Study Questions:

Gene variants can be identified that associate with the risk of gallstone disease by GWAS approach.

6. Design and analysis (study design, inclusion/exclusion, outcome and other variables of interest with specific reference to the time of their collection, summary of data analysis, and any anticipated methodologic limitations or challenges if present).

Participating groups with GWAS data and gallstone disease phenotype:

Discovery studies:
Women's Genome Health Study (WGHS)
Nurses' Health Study (NHS-1/2) and Health Professionals Follow-Up Study (HPFS)
Study of Health in Pomerania (SHIP)
Atherosclerosis Risk In Communities (ARIC) prevalence study
Rotterdam Study
ARIC incidence study
Framingham Heart Study (FHS)
BioVU - (Vanderbilt University)
SPC (PopGen cohort)
SHIP-TREND (Germany)

Replication studies:
Copenhagen City Heart Study
Copenhagen general population study
NHS1/HPFS-replication set
ARIC African American sample

Phenotype: prevalent and incident gallstone disease

Imputation:

- Imputation to HapMap 2.5 M for 22 autosomal chromosomes

Analysis Approach

1) Gallstones trait definitions for GWAS scans:

a. Questionnaire (yes/no): cases will be defined as having answered yes at least once currently or historically to questions such as:
   "Have you ever been told you have gallstones?"
   "Do you or have you ever had gallstones?"
   "Have you ever had gallbladder surgery or gallstones removed?"
   "Does patient/participant have signs of cholesysectomy scar?"

   Controls will be defined as individuals who have always answered no to these questions.

b. Hospital admissions record or other medical record: Genotyped individuals from medical information warehouses may be included if based on ICD-9 codes for cholecystectomy and matching controls can be defined. A separate case definition will be considered based on confirmed gallstones via ultrasound imaging.

c. Exclusions: If other indication for cholecystectomy is known (e.g., other surgery hepatobilliary carcinoma, gallbladder polyp) these individuals should be excluded. These conditions are expected to be rare in population-based or EMR samples.

2) GWAS scans: dichotomous GEE analyses will be conducted with gallstones as the outcome variable and imputed SNP dosage as the independent variable adjusting for age, sex, and cohort-specific variable (eg study site) in a simple model. A second adjusted model will include additional adjustment for BMI.
   - Additive genetic model will be used.

3) Meta-analysis: After we finalize individual cohort results, at least 2 designated analysts will download individual cohort results and conduct weighted meta-analyses to confirm that similar results are obtained by both analysts. Depending on available case/control cohort samples for all definitions (a,b,c) above we may conduct GWAS meta-analysis combining all categories and/or separately meta-analyze specific categories. Meta-analysis will be conducted centrally and top signals prioritize for replication and possible further functional study. Replication and generalization (non-European samples) cohorts have already been contacted and confirmed.

4) Secondary conditional analyses: For all genome-wide significant loci ($P < 5x10^{-8}$) we will conduct conditional analyses adjusting for genotype dose of the top variant at each locus (e.g., top ABCG5/8 SNP) whether there is evidence for multiple, independent signals at the locus. For this analysis we may rely on Peter Visscher et als' method which circumvents the need for individual cohorts to repeat analyses. This will be done centrally based on meta-analysis results using the option for dichotomous traits in GCTA, thus no further work should be necessary at the cohort level (http://www.complextraitgenomics.com/software/gcta/massoc.html).

7.a. Will the data be used for non-CVD analysis in this manuscript? ____ Yes _X___ No

b. If Yes, is the author aware that the file ICTDER03 must be used to exclude persons with a value RES_OTH = "CVD Research" for non-DNA analysis, and for DNA analysis RES_DNA = "CVD Research" would be used? ____ Yes ____ No
(This file ICTDER03 has been distributed to ARIC PIs, and contains the responses to consent updates related to stored sample use for research.)

8.a. Will the DNA data be used in this manuscript? __X__ Yes ____ No

8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the file ICTDER03 must be used to exclude those with value RES_DNA = "No use/storage DNA"?
__X__ Yes ____ No

8.c. If yes, is the author aware that the participants with RES_DNA = 'not for profit' restriction must be excluded if the data are used by a for profit group?
__X__ Yes ____No

9.The lead author of this manuscript proposal has reviewed the list of existing ARIC Study manuscript proposals and has found no overlap between this proposal and previously approved manuscript proposals either published or still in active status. ARIC Investigators have access to the publications lists under the Study Members Area of the web site at: http://www.cscc.unc.edu/ARIC/search.php

__X__ Yes _____ No

10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new proposal or collaboration)?

Boland LL, Folsom AR, Rosamond WD; Atherosclerosis Risk in Communities (ARIC) Study Investigators. Hyperinsulinemia, dyslipidemia, and obesity as risk factors for hospitalized gallbladder disease. A prospective study. Ann Epidemiol. 2002 Feb;12(2):131-40.

Boland LL, Folsom AR, Boerwinkle E; Atherosclerosis Risk in Communities (ARIC) Study Investigators. Apolipoprotein E genotype and gallbladder disease risk in a large population-based cohort. Ann Epidemiol. 2006 Oct;16(10):763-9.

11. a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? ____ Yes __X__ No

11.b. If yes, is the proposal

___ A. primarily the result of an ancillary study (list number* _____)
___ B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* _____)

*ancillary studies are listed by number at http://www.cscc.unc.edu/aric/forms/

12. **Manuscript preparation is expected to be completed in one to three years. If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.**

References:

1. Everhart JE, Khare M, Hill M, et al. Prevalence and ethnic differences in gallbladder disease in the United States. *Gastroenterology.* Sep 1999;117(3):632-639.
2. Katsika D, Grjibovski A, Einarsson C, et al. Genetic and environmental influences on symptomatic gallstone disease: a Swedish study of 43,141 twin pairs. *Hepatology.* May 2005;41(5):1138-1143.
3. Buch S, Schafmayer C, Volzke H, et al. A genome-wide association scan identifies the hepatic cholesterol transporter ABCG8 as a susceptibility factor for human gallstone disease. *Nat Genet.* Aug 2007;39(8):995-999.