ARIC Manuscript Proposal #2258

PC Reviewed: 11/12/13  Status: A  Priority: 2
SC Reviewed: _________  Status: _____  Priority: ____

1.a. Full Title: The associations between PCSK9 loss-of-function variants with stroke, coronary heart disease, and low-density lipoprotein cholesterol

b. Abbreviated Title (Length 26 characters): PCSK9, CVD & LDL-C

2. Writing Group: Kabisa ST, Muntner P, Boerwinkle E, Whitsel EA, on behalf of the of the CHARGE Drug-Gene GWAS consortium

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. _SK_

Sylvie T. Kabisa, PhD PhD
Biostatistics Research Manager
UNC - Chapel Hill
Depts of Epidemiology
Cardiovascular Disease Program
CVS Center, Suite 301-C
137 East Franklin Street
Chapel Hill, NC 27514
(T) 919-966-6652
(F) 919-966-9800
tchumtch@email.unc.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).

Eric A. Whitsel
Research Associate Professor
UNC - Chapel Hill
Depts of Epidemiology and Medicine
Cardiovascular Disease Program
CVS Center, Suite 301-B
137 East Franklin Street
Chapel Hill, NC 27514
(T) 919-966-3168 or 1967
(F) 919-966-9800
eric_whitsel@unc.edu

3. Timeline:
4. Rationale:
Previous studies have established that proprotein convertase subtilisin/kexin type 9 (PCSK9) loss-of-function single-nucleotide polymorphism (SNP) variants are associated with lower low density lipoprotein cholesterol (LDL-C) concentrations.\textsuperscript{1-3} The Y142X (rs67608943) and C679X (rs28362286) nonsense variants more common in blacks have been associated with 30-40 mg/dL lower LDL-C concentrations and the R46L (rs11591147) missense variant more common in whites has been associated with 15-20 mg/dL lower LDL-C concentrations.\textsuperscript{1,3} Also, PCSK9 loss-of-function variants have been associated with a lower risk of coronary heart disease (CHD).\textsuperscript{1,2} The few results available on the association between PCSK9 variants and stroke have produced inconsistent results.\textsuperscript{1,4} The studies of PCSK9 variant and outcomes conducted to date have included few individuals taking statins or excluded these individuals. Statin use has increased substantially over the past 15 years. Therefore, investigating the association of PCSK9 variants and outcomes among statin users is relevant to the current clinical climate.

We have recently completed an analysis of the association between PCSK9 variants and incident CHD and stroke, separately, in the Reasons for Geographic and Racial Differences in Stroke (REGARDS) study (N=11,699), a nationwide cohort study of US adults $\geq 45$ years old, 42% African American/58% white and 56% in the stroke belt region in the southeast US. Approximately, 1/3rd of REGARDS participants were taking statins at baseline and all analyses were conducted separately for individuals taking and not taking a hydroxy-methylglutaryl-coenzyme A (HMG CoA) reductase inhibitor, aka “statin”.

5. Main Hypothesis/Study Questions:
Our current goal is to examine replication of the findings from REGARDS in the CHARGE Drug-Gene Consortium cohorts. Therefore, we propose the following questions: Are the PCSK9 variants listed above (each assayed by the Illumina exome chip) associated with LDL-C concentrations in cross-sectional analyses and with incident CHD and stroke (separately) in time-to-event analyses? Do the overall associations differ among users and non-users of statins? More specifically, we will:

1. Determine if PCSK9 loss-of-function variants are associated with LDL-C concentrations for people taking and not taking statins, separately.
2. Determine if PCSK9 loss-of-function variants are associated with the incidence of CHD for people taking and not taking statins, separately.
3. Determine if PCSK9 loss-of-function variants are associated with the incidence of stroke for people taking and not taking statins, separately.

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).

**Genotypes of interest**
PCSK9 loss-of-function variants for this analysis are nonsense mutations rs67608943 (426C→G; Y142X) and rs28362286 (2037C→A; C679X) for blacks and missense mutation rs11591147 (137G→T; R46L [replacement of the arginine at position 46 with leucine]) for whites.

Analyses will compare those carrying a PCSK9 loss-of-function SNP variant compared to not carrying these variants. The first two variants will be analyzed in blacks only, and the third in whites only. Variants in blacks will be analyzed both separately and together as a single “PCSK9 variant” category.

**Drug exposure definition**
Exposure to statin at baseline (Yes/No) will be defined by the use of atorvastatin, simvastatin, lovastatin, rosuvastatin, pravastatin, fluvastatin, or pitavastatin.

**Dependent variables/outcomes**
Cross-sectional: Fasting LDL-C
Cohort: Time to first stroke event and separately time to first CHD event

**Exclusions**
For all analyses:
- Non-consenters
- Missing information on statin use or PCSK9 genotype at baseline
- Missing covariate information (only for adjusted analyses: see below)

For LDL-C analyses only:
- Did not fast for 8 or more hours prior to blood draw

For incident CHD outcome analyses only:
- Has a history of CHD or missing history of CHD
- Lack of information post-baseline

For incident stroke outcome analyses only:
- Has a history of stroke or missing history of stroke
- Lack of information post-baseline

**Covariates**
All covariates represent participant status at baseline

Continuous variables: Age, systolic blood pressure

Categorical variables: gender, smoking (current versus former/never), physical activity (not regularly physically active [in REGARDS this is coded as not engaging in activity that results in sweat at least once a week]), antihypertensive use (yes/no), kidney function (estimated glomerular filtration rate < 60 ml/min/1.73m² using the CKD-EPI formula), diabetes (fasting glucose \( \geq 126 \) or non-fasting glucose \( \geq 200 \) or taking glucose lowering
pills or insulin), history of CHD (for LDL-C and stroke outcome analyses), and history of stroke (for LDL-C and CHD outcome analyses).

Study specific: principal components (PCs) for population stratification, study site

ANALYSES

Since the pre-1995 prevalence of statin use is low, the first available visit in each participating cohort after January 1st, 1995 will be chosen as the baseline for this study.

All analyses will be performed both stratified and un-stratified by statin use. Each analysis will be stratified by race, with rs67608943 and rs28362286 analyzed only in blacks, and rs11591147 analyzed only in whites. All descriptive analyses and regression models for blacks will be analyzed both separately for each SNP, and by a single “any PCSK9 variant” category. All analyses will be complete case, so that adjusted models may have fewer observations than unadjusted analyses.

Descriptive analyses
Baseline characteristics will be calculated by the four groups defined by PCSK9 loss-of-function variant status and statin use (1) not carrying any variant and not taking a statin, (2) carrying any variant and not taking statin, (3) not carrying any variant and taking a statin, and (4) carrying any variant and taking a statin). Characteristics to be calculated include age, gender, smoking, physical activity, systolic blood pressure, antihypertensive use, kidney function, fasting LDL-C, diabetes, history of CHD, and history of stroke. Continuous characteristics will be calculated as mean (standard deviation) and categorical characteristics will be calculated as sample size (percentage).

Association of PCSK9 variant status and the incidence of stroke.
This analysis will be conducted overall (adjusted for statin use) and stratified by statin use. Participants with a history of stroke at baseline will be excluded. Within each race-statin strata, calculate the number of participants that are stroke follow-up cases with variants, stroke follow-up cases without variants, non-cases with variants, and non-cases without variants. Stroke incidence rates (per 1,000 person-years) with standard errors will be calculated for participants with and without PCSK9 variants. Cox regression models will be used to calculate the hazard ratios for stroke and associated with PCSK9 variants. Hazard ratios will be calculated from unadjusted models, adjusted for age and gender, and then further adjusted for smoking, physical activity, antihypertensive use, systolic blood pressure, kidney function, diabetes, and history of CHD. All three models should be adjusted for PCs and/or study site, as appropriate.

Model (race-stratified, both unstratified and stratified by statin use): log \( h_i(t) = \beta_0 + \beta_{P\text{SNP}} + \beta_{C_i} \)

Where \( h_i(t) \) is hazard of stroke for the \( i^{th} \) participant, \( \beta_0 \) is the intercept, SNP is the presence of the PCSK9 loss-of-function genetic variant, and \( C_i \) is the vector of adjustment variables (not used in unadjusted models).
**Estimation:** Cox regression with robust and model-based estimates of standard errors in cohorts of unrelated individuals; mixed-effects models in cohorts with related individuals.

**Association of PCSK9 variant status and the incidence of CHD by statin use.**
This analysis will be conducted overall (adjusted for statin use) and stratified by statin use. Participants with a history of CHD will be excluded. Within each race-statin stratum, calculate the number of participants that are CHD follow-up cases with variants, CHD follow-up cases without variants, non-cases with variants, and non-cases without variants. CHD incidence rates (per 1,000 person-years) with standard errors will be calculated for participants with and without PCSK9 mutations. Cox regression models will be used to calculate the hazard ratios for CHD associated with PCSK9 variants. Hazard ratios will be calculated from unadjusted models, adjusted for age and gender, and then further adjusted for smoking, physical activity, antihypertensive use, systolic blood pressure, kidney function, diabetes, and history of stroke. All three models should be adjusted for PCs and/or study site, as appropriate.

**Model** (race-stratified, both unstratified and stratified by statin use): \[ \log h_i(t) = \beta_0 + \beta_{pSNP} + \beta_C C_i \]

Where \( h_i(t) \) is hazard of either CHD for the \( i^{th} \) participant, \( \beta_0 \) is the intercept, SNP is the presence of the PCSK9 loss-of-function genetic variant, and \( C_i \) is the vector of adjustment variables (not used in unadjusted models).

**Estimation:** Cox regression with robust and model-based estimates of standard errors in cohorts of unrelated individuals; mixed-effects models in cohorts with related individuals.

**Association of PCSK9 variant status and LDL-C concentrations by statin use.**
This analysis will be conducted overall (adjusted for statin use) and stratified by statin use. Only participants who fasted will be included. Within each race-statin strata, calculate the number of participants that have or do not have variants. General linear models will be used to calculate the cross-sectional difference in LDL-C concentrations comparing participants with versus without (reference) PCSK9 variants. Parameter estimates with standard errors will be calculated from unadjusted models, models adjusted for age and gender, and then further adjusted for smoking, physical activity, antihypertensive use, systolic blood pressure, kidney function, diabetes, and history of CHD and history of stroke. All three models should be adjusted for PCs and/or study site, as appropriate.

**Model** (race-stratified, both unstratified and stratified by statin use): \[ Y_i = \beta_0 + \beta_{pSNP} + \beta_C C_i \]

Where \( Y_i \) is continuous LDL-C for the \( i^{th} \) participant, \( \beta_0 \) is the intercept, SNP is the presence of a PCSK9 loss-of-function genetic variant, and \( C_i \) is the vector of adjustment variables (not used in unadjusted models).

**Estimation:** Linear regression with robust and model-based estimates of standard errors in
cohorts of unrelated individuals; mixed-effects models in cohorts with related individuals

Meta-analysis

27 meta-analyses will be conducted for each variant: one for each of the three outcomes (stroke, CHD, LDL-C), each with three different adjustment models and three different populations (overall, in statin users, and in those not using statins).

Given the small number of variants of interest, accuracy at very low p-values is not necessary in these analyses. Therefore fixed effects inverse variance weighted meta-analysis will be used. Because exposure is binary, and analyses are stratified by statin use, primary analyses will use model-based standard errors, in accordance with recommendations for CHARGE analyses of rare variants. However, robust standard errors will also be reported so that sensitivity analyses can be run on significant results.

7.a. Will the data be used for non-CVD analysis in this manuscript? X Yes _ 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? X 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

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.cscirc.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)?

MS #1093 Nonsense mutations in PCSK9 confer protection against coronary heart disease
MS #1275 PCSK9 variants and PAD
Global fine-mapping of genome-wide association study (GWAS)-identified variants for HDL-C, LDL-C, and triglycerides using the Metabochip in the Population Architecture using Genomics and Epidemiology (PAGE) study

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

11.b. If yes, is the proposal  

___   A. primarily the result of an ancillary study (list number* __________)  

__X__   B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* 2009.10 __________)  

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

12a. 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.

12b. The NIH instituted a Public Access Policy in April, 2008 which ensures that the public has access to the published results of NIH funded research. It is your responsibility to upload manuscripts to PUBMED Central whenever the journal does not and be in compliance with this policy. Four files about the public access policy from http://publicaccess.nih.gov/ are posted in http://www.cscc.unc.edu/aric/index.php, under Publications, Policies & Forms. http://publicaccess.nih.gov/submit_process_journals.htm shows you which journals automatically upload articles to Pubmed central.

References


