1.a. **Full Title**: Genetic Analyses of Abdominal Aortic Aneurysm (AAA)

b. **Abbreviated Title (Length 26 characters)**: genetic analyses of AAA

2. **Writing Group**:

   Writing group members: Weihong Tang, Weihua Guan, Jack Pattee, Pamela Lutsey, James Pankow, Nathan Pankratz, and Eric Boerwinkle, others are welcome

   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: Weihong Tang**
   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: James Pankow
   Address: Division of Epidemiology and Community Health
   University of Minnesota
   Phone: 612-624-2883       Fax: 612-624-0315
   E-mail: panko001@umn.edu

3. **Timeline**:

   Data analysis: as soon as possible upon the approval of the manuscript proposal
   First draft of the manuscript: 3-4 months from manuscript approval date.

4. **Rationale**:

   Abdominal aortic aneurysm (AAA) affects about 1-2% of women and 4%-8% of men over 65 years old.\(^1\) Rupture of AAA is associated with high mortality.\(^2\) Identification of risk factors for AAA may help reduce the incidence and death rate of AAA. Several risk factors for AAA have been reported, including greater age\(^3,4\), male sex\(^4,5\), white race\(^5-8\), cigarette smoking, and hyperlipidemia.\(^9-14\) Genetic influence\(^15\) was also demonstrated for AAA, with heritability being as high as 0.7.\(^16\)
A few large candidate-gene and GWAS studies, based on primarily European American (EA) samples and case-control design, have associated common variants at about 10 loci with AAA risk.\textsuperscript{17-23} The most recent GWAS,\textsuperscript{23} which was the largest, identified 4 new loci: 1q32.3 (SMYD2), 13q12.11 (LINC00540), 20q13.12 (near PCIF1/MMP9/ZNF335), and 21q22.2 (ERG). These new findings need to be replicated in independent populations. Furthermore, the identified variants explained a modest proportion of the heritability of AAA,\textsuperscript{23} suggesting that additional genetic influences are yet discovered.

The ARIC AAA ancillary R01 study (PI, Tang) has ascertained 665 AAAs in the entire ARIC cohort. We therefore can test the hypotheses outlined below.

\textbf{5. Main Hypothesis/Study Questions:}

1. The new loci identified in the recent AAA GWAS can be replicated in ARIC; the time-to-event data in ARIC will provide unique risk estimate (ie, hazards ratio) on the influence of these loci on longitudinal risk of AAA.
2. GWAS and exome chip analyses with AAA will identify additional variants and genes for AAA.

The analytical plan outlined below will utilize measured exome chip data and 1000g GWAS imputation data generated in ARIC.

\textbf{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).}

We will utilize the analytical approach and pipeline that have been established for GWAS and exome chip analyses in ARIC and applied by our team in many other genetics projects. Analyses will focus on ARIC EAs. African American samples will not be considered at the time of proposal due to limited number of AAA cases (n=44 with genetics data). We are actively collaborating with the international AAA GWAS consortium.\textsuperscript{23} Where applicable, the association results from ARIC will be meta-analyzed with the other studies of the consortium.

We will conduct the following analyses for AAA:

1) We will conduct \textbf{replication analysis} of the following four new loci/genes reported by the most recent AAA GWAS\textsuperscript{23} in the EAs of ARIC: rs1795061 in SMYD2, rs9316871 in LINC00540, rs3827066 near PCIF1/MMP9/ZNF335, and rs2836411 in ERG. All the four SNPs are available in the ARIC 1000g imputation with excellent imputation quality score (\textgreater{}0.98). We will exclude participants with known AAA surgery at baseline (n=11) or uncertain AAA status during follow-up (n=30). We will use a Cox proportional hazards model to associate SNP allele dosages with times-to-event calculated from cohort-entry to the time of clinical AAA diagnosis. We will compute hazards ratios (HRs) with adjustment for age,
sex, field center, and the first five principal components for population stratification. The association between the SNPs and AAA risk will be modeled in an additive model. As part of the collaboration between Dr. Tang and Dr. Matthew Bown (one of the senior authors of the AAA GWAS paper), Dr. Bown has provided Dr. Tang genotype data for these four SNPs in a new set of 338 cases and 292 controls ascertained in Grace. Therefore, the replication results from the two studies will be meta-analyzed by using a p value-based, sample size-weighted approach. A short report is expected to come out of the analysis.

2) **GWAS of AAA** based on the 1000g imputation data in ARIC EAs. The association between allele dosage of each SNP in the 1000g data and AAA will be assessed by using logistic regression implemented in SNPTEST. Covariate adjustment and sample exclusions will be the same as in 1) above. Drs. Bown and Cristen Willer (University of Michigan) are expanding their AAA meta-analysis and have invited ARIC study to join through Dr. Tang. Therefore, the AAA GWAS in ARIC will likely join this expanded meta-analysis project.

3) Analysis of AAA with the **exome chip data** in ARIC.
   a. **Single variant association analyses**: Single markers of a suitable frequency depending on the total sample size (e.g., >1% minor allele frequency or >10 minor allele count) will be analyzed for their association with AAA in binomial function models using seqMeta, which can meta-analyze association results across cohorts. Covariate adjustment is the same as described above.
   b. **Gene-based analyses**: For variants of low frequency (e.g., <1% minor allele frequency or another suitable cutpoint), we will evaluate the rare variants in aggregate within a gene in R using the seqMeta package (i.e., genetic burden test). The T1 (or T5) test will be performed. We will select functional variants defined as missense, stop-gain, stop-loss, or splice site changes.
   c. **Meta-analysis**: Right now, the plan is to conduct separate analysis for GWAS and exome chip data in the international AAA consortium. Some cohorts in the consortium are generating exome chip data. When data from several cohorts become available, meta-analysis of single-variant and gene-based tests will be conducted to pool association results across cohorts. A Bonferroni-correction will be applied based on the number of variants or genes tested.

4) Assess the pleiotropic effect between the established SNPs for CHD and AAA. The rationale is based on the observation of some overlap between the loci for AAA and CHD (for example, LDLR, 9p21). Given that a lot more loci have been reported for CHD than AAA, the proposed analysis might lead to identification of additional loci for AAA that did not reach the genome-wide significance threshold in the AAA GWAS.

**Phenotypes:** AAA
Covariates: age, sex, field center, and principal components of ancestry if appropriate. Other risk factors including smoking and total cholesterol will be adjusted for in a secondary analysis.

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 ICTDER 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.cscce.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)?
    None

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
      ___ X  A. primarily the result of an ancillary study (list number* AS 2009.18: “Identifying Genetic and Epidemiological Risk Factors for Abdominal Aortic Aneurysm”, R01HL103695, PI: Weihong Tang)
      ___ 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/

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.

13. Per Data Use Agreement Addendum for the Use of Linked ARIC CMS Data, approved manuscripts using linked ARIC CMS data shall be submitted by the Coordinating Center to CMS for informational purposes prior to publication. Approved manuscripts should be sent to Pingping Wu at CC, at pingping_wu@unc.edu. I will be using CMS data in my manuscript _____ Yes __X__ No.

References:


