ARIC Manuscript Proposal #2221

PC Reviewed: 9/10/13 **Status: A** **Priority: 2**
SC Reviewed: _____ **Status: _____** **Priority: ____**

1.a. Full Title: Association between plasma lipids and risk of abdominal aortic aneurysm in ARIC: a Mendelian randomization study

b. Abbreviated Title (Length 26 characters): AAA and lipid levels

2. Writing Group:

Writing group members:

Lu-Chen Weng, Weihong Tang, Pamela Lutsey, Alvaro Alonso, Weihua Guan, James Pankow, Aaron Folsom

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

First author: Lu-Chen Weng, MS, MPH

Address: Division of Epidemiology and Community Health
University of Minnesota
1300 S. 2nd street, WBOB 300
Minneapolis, MN 55454

Phone: 612-626-8576 Fax: 612-624-0315
E-mail: wengx017@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: **Weihong Tang, MD, PhD**

Address: Division of Epidemiology and Community Health
University of Minnesota
1300 S. 2nd street, WBOB 300
Minneapolis, MN 55454

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

3. Timeline:

6 months for analysis, and 6 months for manuscript preparation
Manuscript draft is anticipated by end of summer 2014.

4. Rationale:

An abdominal aortic aneurysm (AAA) is a localized dilation of the abdominal aorta, which is at least 50% larger than the normal diameter. [1] AAA is one of the leading causes of death in United States with 15,000 deaths per year, and the prevalence of AAA is about 9% in adults over 65 years old. [2] AAA is usually asymptomatic and the mortality after its rupture may exceed 90%.[3] Therefore, screening for asymptomatic AAA and AAA prevention are important to decrease AAA formation and mortality.

In cross-sectional studies and case-control studies, a positive association between low-density lipoprotein (LDL-C) and AAA and a negative association between high-density lipoprotein (HDL-C) and AAA were observed. [4-9] However, inconsistency was also observed in some of the studies. In 206 AAA male cases and 252 male controls, LDL-C level was significantly higher in AAA cases than in controls, but no differences were observed in HDL-C level between groups. [7] On the contrary, greater HDL-C level was associated with reduced risk of AAA in a male cohort, but no association was found between LDL-C level and AAA risk in this population. [8] Recently, a meta-analysis, which contained 8 case-control studies, demonstrated that HDL-C level is likely lower and LDL-C level is likely higher in AAA patients than in controls. [9] Prospective studies supported the inverse relationship between HDL-C and AAA, and they additionally reported a positive association between total cholesterol and AAA. [10-12] In the Atherosclerosis Risk in Communities (ARIC) Study, higher LDL-C and total cholesterol was associated with the risk of future AAA (hazard ratio (HR)=1.94 and 1.49 in the highest versus lowest tertile, respectively), and higher HDL-C was related to 59% lower risk of AAA, after 23 years of follow-up. [13] Notably, the association between triglyceride (TG) and AAA has not been well investigated. TG level was not associated with AAA in two cross-sectional studies, [6,8] but a positive association was observed in a case-control study[14] and in ARIC prospective study as well.

Various large-scale gene-centric and genome-wide association studies (GWAS) have reported lipid-related single nucleotide polymorphisms (SNPs) in humans. [15-28] Recently, some meta-analysis combined multiple cohorts and further verified associations between genetic variants and HDL-C, LDL-C, TG, as well as total cholesterol levels. [19,20,22,23,25] To date, the largest study, including >100,000 individuals of European descent from 46 lipid GWAS, identified a total of 95 SNPs related to lipid levels and replicated most of these signals across ethnic groups. [19] A meta-analysis of 32 studies in 66,240 individuals of European ancestry based on gene-centric genotyping array further identified 11, 5, 12, and 6 novel SNPs for HDL-C, LDL-C, total cholesterol, and TG, respectively. [15]

Mendelian randomization is a specific method to estimate the causal effect of an exposure on an outcome of interest, which uses genetic variant(s) as instrumental variable(s) of the exposure. Based on several assumptions (association between genetic variant(s) and exposure; no association between genetic variants and confounders; no other pathways between genetic variants and outcomes besides through exposure), the causal effect of exposure on outcome can be estimated via the instrument variable (genetic variants), regardless of the presence of confounders. [29] Several causal relationships between lipids and disease outcomes have been investigated via the Mendelian randomization approach, [30-32] but none of them focused on AAA risk. This year, one study demonstrated the associations between lipid-related genetic variants and AAA risk using genetic score in Dutch. [33] Weighted genetics scores of LDL-C and total

cholesterol based on previous GWAS findings were positively related to AAA risk in 807 AAA cases and 1905 controls (odds ratio (OR)= 1.21 for LDL-C and OR= 1.24 for total cholesterol in the highest versus the lowest quartile of genetic score), while genetic scores of HDL-C were negatively associated with AAA risk. Although this study may partially support the role of lipids on AAA risk, it was not able to demonstrate whether the observed effect of SNPs on AAA risk was through lipids. Therefore, we propose to conduct a large Mendelian randomization study to investigate the association between lipids and AAA in the ARIC cohort, using lipid-related SNPs as instrumental variables. Data from this study will provide important information to clarify the causal link between lipid levels and risk of AAA. Because most of the known genetic variants were identified in subject of European descents, this study plans to use the known genetic variants in European Americans.

5. Main Hypothesis/Study Questions:

There is a causal relationship between plasma lipids and AAA risk, which can be demonstrated by a Mendelian randomization 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).

Study Design: A prospective cohort study design will be used to test the causal relationship between lipids and risk of AAA by a Mendelian randomization approach. GWAS data and lipid levels will be from baseline (Visit 1), and AAA status will be ascertained from Visit 1 through the 2013 events follow-up, and the abdominal aortic ultrasound exam at Visit 5.

Population: white participants with GWAS data, plasma lipid measurements, and information on hospital AAAs and ultrasound-detected asymptomatic AAAs.

Exclusions: participants who have missing values for plasma lipid measurements, genotype information, and AAA status will be excluded from this study. In addition, non-whites and extreme outliers in the lipid measurements will be excluded as well.

Independent variable: genetic variants related to plasma lipid levels, including LDL-C, HDL-C, and triglyceride, which were identified and replicated in previous meta-analyses of genome-wide association and gene-centric association studies, will be used as independent variables in this study. A total of 95 loci were identified in the GWAS, and an additional 15 SNPs were reported in the gene-centric association study (Appendix).

Dependent variables: dependent variable will be incident AAAs, including hospital/clinical AAAs and ultrasound-detected asymptomatic AAAs.

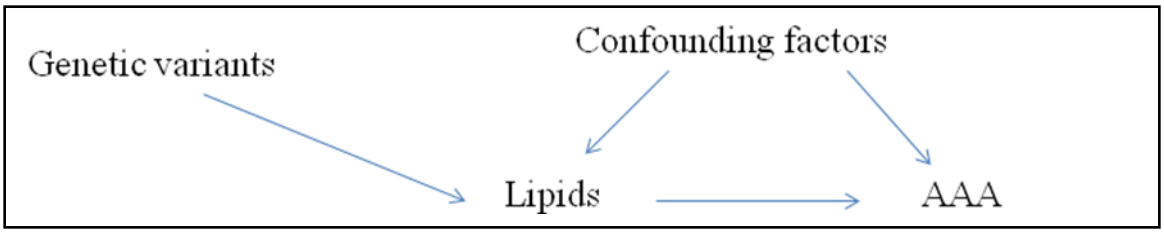
1. Hospital AAAs were defined using the definite ICD diagnostic codes 441.3, 441.4, 441.02, 38.44 and 39.71, and mortality code I71.02, I71.3, I71.4, 441.3 and 441.4. Other diagnostic codes that indicate probable diagnosis of AAA will be

investigated case-by-case to clarify or rule out AAA diagnosis. A total of 325 white subjects have been confirmed as incident hospital AAAs.

2. Asymptomatic AAAs by the abdominal aortic ultrasound exam: We will use commonly used criteria, infrarenal abdominal aortic diameter (IAD) ≥ 30 mm, to define asymptomatic AAA by Visit 5 ultrasound exam. So far in the interim data (n=5989), a total of 108 white ultrasound AAA cases have been identified. We estimate that at least 110 ultrasound AAAs will be identified from Visit 5 ultrasound exam. In total, at least 435 white AAA cases will be identified and included in this analysis.

Covariates of interest: Age, sex, field centers, principal components of population structure, and other variables that may confound the association between lipids and AAA, such as smoking, hypertension, diabetes, coronary heart disease and waist/hip ratio will be considered as potential confounders in this study.

Theoretical model:
Based on the assumption of Mendelian randomization approach, we assume 1) genetic variants are associated with plasma lipids levels, 2) genetic variants are independent of the confounders that confound the association between plasma lipids and AAA, and 3) genetic variants are independent of AAA given confounders and plasma lipids (directed acyclic graph is shown as follows). Because genetic variants are independent to the confounding factors between lipids and AAA, we do not adjust for potential confounders in our basic model, except principal components. Adjustment of potential confounders will be discussed later.



Analysis Plan:
All of the analyses that involve AAA (ie, AAA with lipids and genetic risk score) ***will be conducted and reported separately for clinical and ultrasound-detected AAAs.*** **If a test of the homogeneity of associations for the two case groups is not rejected, results will be pooled using meta-analysis. In addition, a logistic regression will be conducted to pool the data at individual level. The pooled results will also be reported.**

1. Reassurance of the association of interest:

In the first step, we will verify the association of plasma lipids (HDL-C, LDL-C, TG, and total cholesterol) with AAA risk. The Cox proportional hazard model will be used to analyze the clinical AAAs. Follow-up time will be the time period from baseline exam to the date of first hospital AAA event, death, loss to follow-up, or the end of the ARIC Visit 5 exam, whichever occurs earlier. For ultrasound-detected asymptomatic AAAs, the

outcome is the presence of AAA based on IAD ≥ 3.0 cm, ascertained during the Visit 5 ultrasound exam. People with known clinical AAA will be excluded. Since the time to incidence cannot be accurately determined, we will use logistic regression to estimate the odds ratios (OR) and 95% CI for the association of each lipids with asymptomatic AAA. In both analyses, age, sex, field centers, and principal components will be adjusted in these regression models. In addition, because the estimate of 2-stage estimation will be evaluated by probit regression model, we will also use probit regression model to test the associations between lipids and AAA for comparison.

Because participants may take lipid-lowering medications, we will account for these medications by adding constants to lipid levels for all lipid-related analysis. This approach was recommended and investigated in previous studies.[34-37] Following constants will be added for different type of medications.

Medication	LDL-C	HDL-C	TG	Total cholesterol
Statins	+19.9 mg/dL	-2.3 mg/dL	+18.4 mg/dL	+52.1 mg/dL
Fibrates	+40.1 mg/dL	-5.9 mg/dL	+57.1 mg/dL	+46.1 mg/dL
Bile acid sequestrants	+40.5 mg/dL	-1.9 mg/dL	0	
Niacin	+24.7 mg/dL	-9.9 mg/dL	+89.4 mg/dL	

2. Genetic Risk Score (GRS):

Next, we will test the association of unweighted and weighted genetic risk scores (un-wGRS and wGRS) with AAA risk. The analytic approach will be similar to what was described for clinical AAA and ultrasound-detected AAA in 1 above. The GRS will be calculated based on the finding from previous GWAS and gene-centric association study.[19,38] The GWAS included the largest sample size and identified most lipid-related genetic variants. A total of 95 genetic loci will be used, including 47 SNPs for HDL-C, 37 for LDL-C, 31 for TG, and 52 SNPs for total cholesterol. The gene-centric association study identified 39 loci for HDL-C, 34 loci for LDL-C, 32 loci for TG, and 41 loci for total cholesterol, in which 11 SNPs for HDL-C, 5 SNPs for LDL-C, 6 SNPs for TG, and 12 SNPs for total cholesterol ($p<2.4x10\text{-}6$) were novel. In our CARe genotype array, 10 SNPs for HDL-C, 5 SNPs for LDL-C, 7 SNPs for TG, and 9 SNPs for total cholesterol are available. In total, additional 15 SNPs will be tested.

Because multiple SNPs identified in the same region may represent identical signal, LD between SNPs will be evaluated. We assume that effect of each SNP is independent; therefore, we will evaluate linkage disequilibrium (LD) between selected SNPs and exclude SNPs in high LD ($r^2>0.8$) with the top/top functional SNP in that region. Only the top signal (smallest p-value) or the top functional SNP within high LD ($r^2>0.8$) region will be included in each GRS.

Un-wGRS will sum the total number of risk alleles, with each risk allele assigned as 1 (two will be the maximum value for each SNP). wGRS will be calculated by three methods ($wGRS_1$, $wGRS_2$, and $wGRS_2$). The calculation formula of the first method is as follows.

$$wGRS_1=\Sigma(\beta i * Gi)/n \qquad (1)$$

β is effect size of each SNP on lipid levels; n is the total number of SNPs for individual i; G is the number of risk alleles.

The second method is similar to the first method. Rather than divided by the total number of SNPs and the end, $wGRS_2$ will be divided by the average effect size, and $wGRS_3$ will only sum up the products of β and G. The calculation formulas of the second and third method are as follows.

$$wGRS_2=\Sigma(\beta i * Gi)/(\Sigma \beta i /n) \quad (2)$$

$$wGRS_3=\Sigma(\beta i * Gi) \quad (3)$$

These methods were also used to test the association of lipids with longitudinal trends in lipid levels, atherosclerosis, incident coronary heart disease, and AAA in previous studies. [34,39,40] The primary method of this proposed study will be the second method, which has been applied to test the longitudinal trends of lipid levels in ARIC. Comparing different F-statistic and proportion of total variance obtained from the regression model of each GRS and lipid levels, a strong instrumental variable of lipid levels will be defined for investigating the causal relationship between plasma lipids and AAA risk.

3. 2-stage estimation

A 2-stage estimation will be used to test the causal relationship between GRS and AAA risk. The first stage will generate predicted values of intermediate phenotypes (HDL-C, LDL-C, triglyceride, and total cholesterol), by regressing of each intermediate phenotype (lipids) on the instrument (GRS). The second stage will regress the outcome (AAA) on the predicted value of intermediate phenotypes. The 2-stage estimation will be performed by an instrumental variable probit regression analysis.

An additive model will be used, and the marginal effect of lipids obtained from the second stage will be explained as the probability of AAA risk changes per 1 unit of lipid level increment. Since we hypothesize GRS to be an instrument variable of lipids, potential confounding between lipids and AAA risk will be as addressed via covariates in our analysis.

To meet the assumptions of Mendelian randomization, several tests will be required. First, the F-statistics will be used to test the strength of instrument variable. A minimum F-statistic of 10 may indicate enough strength for the instrument variable method.[37] The Pearson correlation and 1-way anova will be used to test the second assumption, which assumes the instrument variable (GRS) is not correlated with confounders. If a potential confounder is related to GRS, we will consider adjusting for this confounder at the second stage. The Durbin-Wu-Hausman statistics (DWH) for comparing the results to ordinary least square will be used to test the third assumption; it assumes the instrument variable (GRS) has an effect on outcome (AAA risk) via the predicted value of the intermediate phenotype.

4. Sensitive analysis

Some identified SNPs may not only relate to lipids but also relate to potential confounders between lipids and AAA (pleiotropy**)**, which violates the third assumption of the Mendelian randomization approach. Additionally, SNPs that relate to more than one lipid trait also violate that assumption (e.g. SNPs relate to both LDL-C and total cholesterol). Scores without these SNPs will be estimated, and the causal relation will be evaluated using the new scores. In addition, to avoid the heterogeneity of genotyping/imputation between the GWAS array and the gene-centric platform, scores without SNPs identified from gene-centric association studies will also be calculated, and the corresponding causal link will be investigated as well. At last, because the constants we add to lipid level may not be accurate, rather than adding constants to the lipid levels, we will exclude all participants who took antihyperlipidemic medications at baseline. The causal inference will be evaluated in this subgroup to reassure our findings in the whole population.

5. Limitation and challenges

Rather than studying a single SNP that has limited effect, we plan to create genetic risk scores based on multiple genetic variants to provide larger power. [41] To our knowledge, one of the weighted GRSs based on GWAS result explained only 2-6% total variances in lipid levels in ARIC. [34] In this proposed study, an unweighted score as well as three weighted scores that have been shown to relate to cardiovascular outcomes will be used, and additional SNPs identified in gene-centric studies will be added for GRS calculation. We expect at least one of them to be a good instrument vector for the current approach. On the other hand, because population stratification is a potential problem of Mendelian randomization studies, we will only include white population. In addition, 10 principal components will be adjusted for in our model to account for population stratification. We assume our result will show the true association after adjustment.

7.a. Will the data be used for non-CVD analysis in this manuscript? ____ Yes ___ No_X_

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.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)?

Folsom 1505 Risk Factors for Abdominal Aortic Aneurysm
Nettleton 1101 LIPC polymorphisms, dietary fat, and plasma HDL cholesterol in adults with and without type II diabetes.
Brautbar 1465 Relationship between single nucleotide polymorphisms previously associated with lipid levels, HDL-C or triglyceride extreme levels, and atherosclerotic events.
Volcik 1391 A Genome-Wide Association Study for HDL-Cholesterol, LDL-Cholesterol and Triglycerides in ~11,000 African Americans
Barbalic 1530 Genome-wide association study of blood lipid levels: Global Lipids Consortium\

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.

Reference

1. Prisant LM, Mondy JS,3rd. Abdominal aortic aneurysm. *J Clin Hypertens (Greenwich)*. 2004;6(2):85-89.

2. Mohan PP, Rozenfeld M, Kane RA, Calandra JD, Hamblin MH. Nationwide trends in abdominal aortic aneurysm repair and use of endovascular repair in the emergency setting. *Journal of Vascular and Interventional Radiology*. 2012;23(3):338-344.

3. Ernst CB. Abdominal aortic aneurysm. *N Engl J Med*. 1993;328(16):1167-1172. doi: 10.1056/NEJM199304223281607.

4. Alcorn HG, Wolfson SK, Sutton-Tyrrell K, Kuller LH, O'Leary D. Risk factors for abdominal aortic aneurysms in older adults enrolled in the cardiovascular health study. *Arterioscler Thromb Vasc Biol*. 1996;16(8):963-970.

5. Naydeck BL, Sutton-Tyrrell K, Schiller KD, Newman AB, Kuller LH. Prevalence and risk factors for abdominal aortic aneurysms in older adults with and without isolated systolic hypertension. *Am J Cardiol*. 1999;83(5):759-764.

6. Singh K, Bønaa K, Jacobsen B, Bjørk L, Solberg S. Prevalence of and risk factors for abdominal aortic aneurysms in a population-based study the tromsø study. *Am J Epidemiol*. 2001;154(3):236-244.

7. Hobbs S, Claridge M, Quick C, Day N, Bradbury A, Wilmink A. LDL cholesterol is associated with small abdominal aortic aneurysms. *European journal of vascular and endovascular surgery*. 2003;26(6):618-622.

8. Golledge J, Van Bockxmeer F, Jamrozik K, McCann M, Norman PE. Association between serum lipoproteins and abdominal aortic aneurysm. *Am J Cardiol*. 2010;105(10):1480-1484.

9. Takagi H, Manabe H, Kawai N, Goto S, Umemoto T. Serum high-density and low-density lipoprotein cholesterol is associated with abdominal aortic aneurysm presence: A systematic review and meta-analysis. *International angiology*. 2010;29(4):371-375.

10. Forsdahl SH, Singh K, Solberg S, Jacobsen BK. Risk factors for abdominal aortic aneurysms A 7-year prospective study: The tromsø study, 1994–2001. *Circulation*. 2009;119(16):2202-2208.

11. Rodin MB, Daviglus ML, Wong GC, et al. Middle age cardiovascular risk factors and abdominal aortic aneurysm in older age. *Hypertension*. 2003;42(1):61-68.

12. Törnwall ME, Virtamo J, Haukka JK, Albanes D, Huttunen JK. Life-style factors and risk for abdominal aortic aneurysm in a cohort of finnish male smokers. *Epidemiology*. 2001;12(1):94-100.

13. Tang W, Alonso A, Lutsey P, Lederle F, Yao L, Folsom A. **Associations between middle-age risk factors and future risk of abdominal aortic aneurysm
 the atherosclerosis risk in communities (ARIC) study**. . 2013.

14. Watt HC, Law MR, Wald NJ, Craig WY, Ledue TB, Haddow JE. Serum triglyceride: A possible risk factor for ruptured abdominal aortic aneurysm. *Int J Epidemiol*. 1998;27(6):949-952.

15. Asselbergs FW, Guo Y, van Iperen EP, et al. Large-scale gene-centric meta-analysis across 32 studies identifies multiple lipid loci. *Am J Hum Genet*. 2012;91(5):823-838. doi: 10.1016/j.ajhg.2012.08.032; 10.1016/j.ajhg.2012.08.032.

16. Chasman DI, Pare G, Zee RY, et al. Genetic loci associated with plasma concentration of low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, triglycerides, apolipoprotein A1, and apolipoprotein B among 6382 white women in genome-wide analysis with replication. *Circ Cardiovasc Genet*. 2008;1(1):21-30. doi: 10.1161/CIRCGENETICS.108.773168; 10.1161/CIRCGENETICS.108.773168.

17. Tanaka T, Shen J, Abecasis GR, et al. Genome-wide association study of plasma polyunsaturated fatty acids in the InCHIANTI study. *PLoS genetics*. 2009;5(1):e1000338.

18. Aulchenko YS, Ripatti S, Lindqvist I, et al. Loci influencing lipid levels and coronary heart disease risk in 16 european population cohorts. *Nat Genet*. 2009;41(1):47-55. doi: 10.1038/ng.269; 10.1038/ng.269.

19. Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*. 2010;466(7307):707-713. doi: 10.1038/nature09270; 10.1038/nature09270.

20. Waterworth DM, Ricketts SL, Song K, et al. Genetic variants influencing circulating lipid levels and risk of coronary artery disease. *Arterioscler Thromb Vasc Biol*. 2010;30(11):2264-2276. doi: 10.1161/ATVBAHA.109.201020; 10.1161/ATVBAHA.109.201020.

21. Kathiresan S, Manning AK, Demissie S, et al. A genome-wide association study for blood lipid phenotypes in the framingham heart study. *BMC Med Genet*. 2007;8 Suppl 1:S17. doi: 10.1186/1471-2350-8-S1-S17.

22. Willer CJ, Sanna S, Jackson AU, et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet*. 2008;40(2):161-169. doi: 10.1038/ng.76; 10.1038/ng.76.

23. Musunuru K, Romaine SP, Lettre G, et al. Multi-ethnic analysis of lipid-associated loci: The NHLBI CARe project. *PLoS One*. 2012;7(5):e36473. doi: 10.1371/journal.pone.0036473; 10.1371/journal.pone.0036473.

24. Ma L, Yang J, Runesha HB, et al. Genome-wide association analysis of total cholesterol and high-density lipoprotein cholesterol levels using the framingham heart study data. *BMC Med Genet*. 2010;11:55-2350-11-55. doi: 10.1186/1471-2350-11-55; 10.1186/1471-2350-11-55.

25. Sandhu MS, Waterworth DM, Debenham SL, et al. LDL-cholesterol concentrations: A genome-wide association study. *Lancet*. 2008;371(9611):483-491. doi: 10.1016/S0140-6736(08)60208-1; 10.1016/S0140-6736(08)60208-1.

26. Chasman DI, Pare G, Mora S, et al. Forty-three loci associated with plasma lipoprotein size, concentration, and cholesterol content in genome-wide analysis. *PLoS Genet*. 2009;5(11):e1000730. doi: 10.1371/journal.pgen.1000730; 10.1371/journal.pgen.1000730.

27. Zemunik T, Boban M, Lauc G, et al. Genome-wide association study of biochemical traits in korcula island, croatia. *Croat Med J*. 2009;50(1):23-33.

28. Smith EN, Chen W, Kahonen M, et al. Longitudinal genome-wide association of cardiovascular disease risk factors in the bogalusa heart study. *PLoS Genet*. 2010;6(9):e1001094. doi: 10.1371/journal.pgen.1001094; 10.1371/journal.pgen.1001094.

29. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Stat Med*. 2008;27(8):1133-1163.

30. Haase CL, Tybjærg-Hansen A, Qayyum AA, Schou J, Nordestgaard BG, Frikke-Schmidt R. LCAT, HDL cholesterol and ischemic cardiovascular disease: A mendelian randomization study of HDL cholesterol in 54,500 individuals. *Journal of Clinical Endocrinology & Metabolism*. 2012;97(2):E248-E256.

31. Voight BF, Peloso GM, Orho-Melander M, et al. Plasma HDL cholesterol and risk of myocardial infarction: A mendelian randomisation study. *The Lancet*. 2012;380(9841):572-580.

32. Trompet S, Jukema JW, Katan MB, et al. Apolipoprotein e genotype, plasma cholesterol, and cancer: A mendelian randomization study. *Am J Epidemiol*. 2009;170(11):1415-1421.

33. van 't Hof FN, Ruigrok YM, Baas AF, et al. Impact of inherited genetic variants associated with lipid profile, hypertension, and coronary artery disease on the risk of intracranial and abdominal aortic aneurysms. *Circ Cardiovasc Genet*. 2013;6(3):264-270. doi: 10.1161/CIRCGENETICS.113.000022; 10.1161/CIRCGENETICS.113.000022.

34. Lutsey PL, Rasmussen-Torvik LJ, Pankow JS, et al. Relation of lipid gene scores to longitudinal trends in lipid levels and incidence of abnormal lipid levels among individuals of european ancestry: The atherosclerosis risk in communities (ARIC) study. *Circ Cardiovasc Genet*. 2012;5(1):73-80. doi: 10.1161/CIRCGENETICS.111.959619; 10.1161/CIRCGENETICS.111.959619.

35. Tobin MD, Sheehan NA, Scurrah KJ, Burton PR. Adjusting for treatment effects in studies of quantitative traits: Antihypertensive therapy and systolic blood pressure. *Stat Med*. 2005;24(19):2911-2935. doi: 10.1002/sim.2165.

36. Wu J, Province MA, Coon H, et al. An investigation of the effects of lipid-lowering medications: Genome-wide linkage analysis of lipids in the HyperGEN study. *BMC Genet*. 2007;8:60. doi: 10.1186/1471-2156-8-60.

37. National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). Third report of the national cholesterol education program (NCEP) expert panel on

detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel III) final report. *Circulation*. 2002;106(25):3143-3421.

38. Asselbergs FW, Guo Y, van Iperen E, et al. Large-scale gene-centric meta-analysis across 32 studies identifies multiple lipid loci. *The American Journal of Human Genetics*. 2012.

39. van't Hof FN, Ruigrok YM, Baas AF, et al. Impact of inherited genetic variants associated with lipid profile, hypertension, and coronary artery disease on the risk of intracranial and abdominal aortic aneurysms. *Circulation: Cardiovascular Genetics*. 2013;6(3):264-270.

40. Isaacs A, Willems SM, Bos D, et al. Risk scores of common genetic variants for lipid levels influence atherosclerosis and incident coronary heart disease. *Arterioscler Thromb Vasc Biol*. 2013. doi: 10.1161/ATVBAHA.113.301236.

41. Palmer TM, Lawlor DA, Harbord RM, et al. Using multiple genetic variants as instrumental variables for modifiable risk factors. *Stat Methods Med Res*. 2012;21(3):223-242. doi: 10.1177/0962280210394459; 10.1177/0962280210394459.

Appendix I-V: SNPs for genetic score calculation.
SNPs in Appendix I-IV were identified in the Teslovich GWAS, and the table were derived from the previous ARIC lipid gene score study (adapted from Lutsey *et al*.).

Appendix I. SNPs related to HDL-C

SNP	Chromosome	Position	Major Allele *	Minor Allele *	Effect Size*	ARIC MAF	Genotyping Status
rs4660293	1	39,800,767	A	**G**	-0.48	0.24	Directly genotyped
rs1689800	1	180,435,508	A	**G**	-0.47	0.35	Imputed
rs4846914	1	228,362,314	A	**G**	-0.61	0.40	Directly genotyped
rs1042034	2	21,078,786	**T**	C	0.9	0.22	Imputed
rs12328675	2	165,249,046	**T**	C	0.68	0.13	Imputed
rs1515100	2	226,837,161	**A**	C	0.46	0.36	Directly genotyped
rs13107325	4	103,407,732	C	**T**	-0.84	0.10	Imputed
rs6450176	5	53,333,782	G	**A**	-0.49	0.26	Imputed
rs2814944	6	34,660,775	G	**A**	-0.49	0.15	Directly genotyped
rs605066	6	139,871,359	T	**C**	-0.39	0.41	Imputed
rs1084651	6	161,009,807	G	**A**	-0.56	0.16	Imputed
rs17145738	7	72,620,810	**C**	T	0.57	0.12	Directly genotyped
rs4731702	7	130,083,924	**C**	T	0.59	0.48	Imputed
rs9987289	8	9,220,768	G	**A**	-1.21	0.08	Imputed
rs12678919	8	19,888,502	**A**	G	2.25	0.11	Imputed
rs2293889	8	116,668,374	G	**T**	-0.44	0.42	Imputed
rs10808546	8	126,565,000	**C**	T	0.61	0.44	Imputed
rs643531	9	15,286,034	A	**C**	-0.72	0.13	Directly genotyped
rs1883025	9	106,704,122	C	**T**	-0.94	0.26	Imputed
rs2923084	11	10,345,358	A	**G**	-0.41	0.18	Directly genotyped
rs3136441	11	46,699,823	**T**	C	0.78	0.14	Imputed
rs174601	11	61,379,716	C	**T**	-0.73	0.36	Imputed
rs964184	11	116,154,127	C	**G**	-1.5	0.14	Directly genotyped

rs7115089	11	122,035,801	**C**	G	0.31	0.38	Imputed
rs7134375	12	20,365,025	**C**	A	0.4	0.39	Imputed
rs3741414	12	56,130,316	**C**	T	0.46	0.25	Imputed
rs7134594	12	108,484,576	T	**C**	-0.44	0.47	Imputed
rs4759375	12	122,362,191	**C**	T	0.86	0.06	Imputed
rs4765127	12	123,026,120	**G**	T	0.44	0.34	Directly genotyped
rs838880	12	123,827,546	**T**	C	0.61	0.32	Directly genotyped
rs1532085	15	56,470,658	**G**	A	1.45	0.36	Imputed
rs2652834	15	61,183,920	G	**A**	-0.39	0.19	Imputed
rs3764261	16	55,550,825	**C**	A	3.39	0.34	Imputed
rs16942887	16	66,485,543	**G**	A	1.27	0.12	Imputed
rs2925979	16	80,092,291	C	**T**	-0.45	0.30	Imputed
rs881844	17	35,063,744	G	**C**	-0.51	0.34	Imputed
rs4148008	17	64,386,889	C	**G**	-0.42	0.31	Imputed
rs4082919	17	73,889,077	T	**G**	-0.4	0.47	Imputed
rs7241918	18	45,414,951	T	**G**	-1.31	0.17	Imputed
rs12967135	18	56,000,003	G	**A**	-0.42	0.22	Imputed
rs7255436	19	8,339,196	A	**C**	-0.45	0.47	Imputed
rs737337	19	11,208,493	T	**C**	-0.64	0.09	Imputed
rs4420638	19	50,114,786	A	**G**	-1.06	0.17	Directly genotyped
rs386000	19	59,484,573	**G**	C	0.83	0.20	Imputed
rs1800961	20	42,475,778	C	**T**	-1.88	0.03	Directly genotyped
rs6065906	20	43,987,422	T	**C**	-0.93	0.18	Imputed
rs181362	22	20,262,068	C	**T**	-0.46	0.20	Imputed

*Alleles in bold type refer to coded alleles; MAF: minor allele frequency; Effect sizes that refer to the change of phenotype per minor allele were identified in the Teslovich GWAS

Appendix II. SNPs related to LDL-C

SNP	Chromosome	Position	Major Allele	Minor Allele	Effect Size*	ARIC MAF	Genotyping Status

rs12027135	1	25,648,320	**T**	A	-1.1	0.45	Imputed
rs2479409	1	55,277,238	A	**G**	2.01	0.32	Directly genotyped
rs3850634	1	62,823,186	**T**	G	-1.59	0.33	Imputed
rs629301	1	109,619,829	**T**	G	-5.65	0.22	Directly genotyped
rs2807834	1	219,037,216	**G**	T	-1.09	0.31	Directly genotyped
rs514230	1	232,925,220	**T**	A	-1.13	0.48	Imputed
rs1367117	2	21,117,405	G	**A**	4.05	0.31	Imputed
rs4299376	2	43,926,080	T	**G**	2.75	0.30	Imputed
rs12916	5	74,692,295	T	**C**	2.45	0.39	Imputed
rs6882076	5	156,322,875	**C**	T	-1.67	0.36	Imputed
rs3757354	6	16,235,386	**C**	T	-1.43	0.23	Imputed
rs1800562	6	26,201,120	**G**	A	-2.22	0.06	Directly genotyped
rs3177928	6	32,520,413	G	**A**	1.83	0.15	Directly genotyped
rs11153594	6	116,461,284	**C**	T	-0.89	0.40	Directly genotyped
rs1564348	6	160,498,850	T	**C**	1.95	0.16	Directly genotyped
rs12670798	7	21,573,877	T	**C**	1.26	0.24	Imputed
rs217386	7	44,567,220	**G**	A	-1.17	0.44	Imputed
rs2126259	8	9,222,556	**C**	T	-2.22	0.09	Imputed
rs1030431	8	59,474,251	G	**A**	0.95	0.35	Imputed
rs2954022	8	126,551,803	**C**	A	-1.84	0.47	Imputed
rs11136341	8	145,115,531	A	**G**	1.4	0.40	Imputed
rs649129	9	135,144,125	C	**T**	2.05	0.23	Imputed
rs1129555	10	113,900,711	G	**A**	1.08	0.28	Directly genotyped
rs174583	11	61,366,326	**C**	T	-1.71	0.34	Directly genotyped
rs964184	11	116,154,127	C	**G**	2.85	0.14	Directly genotyped
rs11220462	11	125,749,162	G	**A**	1.95	0.14	Imputed
rs11065987	12	110,556,807	**A**	G	-0.97	0.43	Imputed
rs1169288	12	119,901,033	A	**C**	1.42	0.32	Imputed
rs2332328	14	23,952,898	C	**T**	1.17	0.48	Imputed
rs247616	16	55,547,091	**C**	T	-1.45	0.34	Imputed

rs2000999	16	70,665,594	G	**A**	2	0.20	Imputed
rs7225700	17	42,746,803	**C**	T	-0.87	0.36	Imputed
rs6511720	19	11,063,306	**G**	T	-6.99	0.10	Imputed
rs10401969	19	19,268,718	**T**	C	-3.11	0.08	Directly genotyped
rs4420638	19	50,114,786	A	**G**	7.14	0.17	Directly genotyped
rs2902941	20	38,524,928	**A**	G	-0.98	0.34	Imputed
rs909802	20	39,370,229	C	**T**	1.41	0.48	Imputed

*Alleles in bold type refer to coded alleles; MAF: minor allele frequency; Effect sizes that refer to the change of phenotype per minor allele were identified in the Teslovich GWAS

Appendix III. SNPs related to triglyceride

SNP	Chromosome	Position	Major Allele	Minor Allele	Effect Size*	ARIC MAF	Genotyping Status
rs2131925	1	62,798,530	**T**	G	-4.94	0.33	Imputed
rs1321257	1	228,371,935	A	**G**	2.76	0.39	Imputed
rs1042034	2	21,078,786	**T**	C	-5.99	0.22	Imputed
rs1260326	2	27,584,444	C	**T**	8.76	0.41	Imputed
rs10195252	2	165,221,337	**T**	C	-2.01	0.41	Imputed
rs2943645	2	226,807,424	**T**	C	-1.89	0.36	Imputed
rs645040	3	137,409,312	**T**	G	-2.22	0.22	Directly genotyped
rs442177	4	88,249,285	**T**	G	-2.25	0.41	Imputed
rs9686661	5	55,897,543	C	**T**	2.57	0.20	Imputed
rs1553318	5	156,411,901	**C**	G	-2.63	0.35	Imputed
rs2247056	6	31,373,469	**C**	T	-2.99	0.26	Imputed
rs13238203	7	71,767,603	**C**	T	-7.91	0.04	Imputed
rs7811265	7	72,572,446	**A**	G	-7.91	0.20	Imputed
rs11776767	8	10,721,339	G	**C**	2.01	0.37	Imputed
rs1495743	8	18,317,580	C	**G**	2.97	0.23	Directly genotyped
rs12678919	8	19,888,502	**A**	G	-13.64	0.11	Imputed
rs2954029	8	126,560,154	**A**	T	-5.64	0.47	Imputed

rs10761731	10	64,697,616	**A**	T	-2.38	0.42	Imputed
rs2068888	10	94,829,632	**G**	A	-2.28	0.48	Imputed
rs174546	11	61,326,406	C	**T**	3.82	0.33	Imputed
rs964184	11	116,154,127	C	**G**	16.95	0.14	Directly genotyped
rs11613352	12	56,078,847	**C**	T	-2.7	0.23	Imputed
rs12310367	12	123,052,631	**A**	G	-2.42	0.35	Imputed
rs2412710	15	40,471,079	G	**A**	7	0.03	Imputed
rs2929282	15	42,033,223	A	**T**	5.13	0.05	Imputed
rs261342	15	56,518,445	C	**G**	2.99	0.21	Imputed
rs11649653	16	30,825,988	**C**	G	-2.13	0.39	Directly genotyped
rs7205804	16	55,562,390	**G**	A	-2.88	0.46	Imputed
rs10401969	19	19,268,718	**T**	C	-7.83	0.08	Directly genotyped
rs439401	19	50,106,291	**C**	T	-5.5	0.39	Imputed
rs4810479	20	43,978,455	T	**C**	3.32	0.24	Imputed
rs5756931	22	36,875,979	**T**	C	-1.54	0.40	Imputed

*Alleles in bold type refer to coded alleles; MAF: minor allele frequency; Effect sizes that refer to the change of phenotype per minor allele were identified in the Teslovich GWAS

Appendix IV. SNPs related to total cholesterol

SNP	Chromosome	Position	Major Allele	Minor Allele	Effect Size*	ARIC MAF	Genotyping Status
rs12027135	1	25,648,320	**T**	A	-1.22	0.45	Imputed
rs2479409	1	55,277,238	A	**G**	1.96	0.32	Directly genotyped
rs3850634	1	62,823,186	**T**	G	-2.6	0.33	Imputed
rs7515577	1	92,782,026	**A**	C	-1.18	0.21	Directly genotyped
rs629301	1	109,619,829	**T**	G	-5.41	0.22	Directly genotyped
rs2807834	1	219,037,216	**G**	T	-1.38	0.31	Directly genotyped
rs514230	1	232,925,220	**T**	A	-1.36	0.48	Imputed
rs1367117	2	21,117,405	G	**A**	4.16	0.31	Imputed
rs1260326	2	27,584,444	C	**T**	1.91	0.41	Imputed

rs4299376	2	43,926,080	T	**G**	3.01	0.30	Imputed
rs6759321	2	136,039,146	G	**T**	1.18	0.29	Imputed
rs2290159	3	12,603,920	**G**	C	-1.42	0.22	Imputed
rs12916	5	74,692,295	T	**C**	2.84	0.39	Imputed
rs6882076	5	156,322,875	**C**	T	-1.98	0.36	Imputed
rs3757354	6	16,235,386	**C**	T	-1.46	0.23	Imputed
rs1800562	6	26,201,120	**G**	A	-2.16	0.06	Directly genotyped
rs3177928	6	32,520,413	G	**A**	2.31	0.15	Directly genotyped
rs2814982	6	34,654,538	**C**	T	-1.86	0.10	Imputed
rs9488822	6	116,419,586	**A**	T	-1.18	0.35	Imputed
rs1564348	6	160,498,850	T	**C**	2.18	0.16	Directly genotyped
rs2285942	7	21,549,442	C	**T**	1.7	0.16	Imputed
rs2072183	7	44,545,705	G	**C**	2.01	0.27	Imputed
rs2126259	8	9,222,556	**C**	T	-3.14	0.09	Imputed
rs1961456	8	18,299,989	A	**G**	1.07	0.33	Imputed
rs1030431	8	59,474,251	G	**A**	1.26	0.35	Imputed
rs2737229	8	116,717,740	**A**	C	-1.11	0.30	Imputed
rs2954022	8	126,551,803	**C**	A	-2.3	0.47	Imputed
rs11136341	8	145,115,531	A	**G**	1.34	0.40	Imputed
rs581080	9	15,295,378	**C**	G	-1.57	0.19	Imputed
rs1883025	9	106,704,122	**C**	T	-2.24	0.26	Imputed
rs651007	9	135,143,696	C	**T**	2.3	0.23	Directly genotyped
rs2255141	10	113,923,876	G	**A**	1.14	0.28	Imputed
rs10832963	11	18,620,817	**G**	T	-1.06	0.27	Imputed
rs174550	11	61,328,054	**T**	C	-1.78	0.33	Imputed
rs964184	11	116,154,127	C	**G**	4.68	0.14	Directly genotyped
rs7941030	11	122,027,585	T	**C**	0.97	0.39	Imputed
rs11220463	11	125,753,421	A	**T**	2.01	0.10	Imputed
rs11065987	12	110,556,807	**A**	G	-0.96	0.43	Imputed
rs1169288	12	119,901,033	A	**C**	1.45	0.32	Imputed

rs1532085	15	56,470,658	G	**A**	1.54	0.36	Imputed
rs3764261	16	55,550,825	C	**A**	1.67	0.34	Imputed
rs2000999	16	70,665,594	G	**A**	2.34	0.20	Imputed
rs7206971	17	42,780,114	G	**A**	1.01	0.48	Imputed
rs7239867	18	45,418,715	**G**	A	-1.94	0.17	Imputed
rs6511720	19	11,063,306	**G**	T	-7.09	0.10	Imputed
rs10401969	19	19,268,718	**T**	C	-4.74	0.08	Directly genotyped
rs4420638	19	50,114,786	A	**G**	6.83	0.17	Directly genotyped
rs492602	19	53,898,229	A	**G**	1.27	0.47	Imputed
rs2277862	20	33,616,196	**C**	T	-1.19	0.15	Imputed
rs2902940	20	38,524,901	**A**	G	-1.38	0.29	Imputed
rs4297946	20	39,244,689	G	**C**	1.52	0.48	Directly genotyped
rs1800961	20	42,475,778	**C**	T	-4.73	0.03	Directly genotyped

*Alleles in bold type refer to coded alleles; MAF: minor allele frequency; Effect sizes that refer to the change of phenotype per minor allele were identified in the Teslovich GWAS

Appendix V. Additional SNPs related to lipids from large-scale gene-centric association study*

Trait	SNP	Chromosome	Position	Major Allele	Minor Allele	Effect Size*	ARIC MAF
HDL-C	rs4970834	1	109,616,403	**C**	T	0.0147926	0.19
	rs28645722	8	19,847,174	G	**A**	-0.053755	0.02
	rs28575919	8	19,847,249	G	**C**	-0.059755	0.02
	rs765547	8	19,910,554	**G**	A	0.0396713	
	rs7388248	8	144,376,728	**G**	C	0.0100259	0.27
	rs2066718	9	106,629,076	**C**	T	0.0369465	0.03
	rs4759361	12	121,744,233	**T**	A	0.0121857	0.19
	rs3922628	12	121,775,248	**A**	T	0.0121919	0.22

	rs583662	15	56,509,568	A	**G**	-0.033982	0.02
	rs12720873	16	55,563,573	**G**	A	0.0404438	0.03
	rs12966382	18	45,339,438	**C**	T	0.0183067	0.15
LDL-C	rs11591147	1	55,278,235	**G**	T	-0.357963404	0.02
	rs3798220	6	160,881,127	T	**C**	0.089303332	0.02
	rs4725984	7	150,299,447	C	**T**	0.029922353	0.35
	rs11024739	11	18,602,419	**A**	C	-0.031721519	0.27
	rs1801689	17	61,641,042	A	**C**	0.10322347	0.03
TG	rs389883	6	32,055,439	**T**	G	-0.015634275	0.29
	rs17211510	6	32,710,408	C	**A**	0.015192776	0.26
	rs28645722	8	19,847,174	G	**A**	0.062747593	0.02
	rs28575919	8	19,847,249	G	**C**	0.064827007	0.02
	rs3135507	11	116,166,698	C	**T**	0.041063735	0.03
	rs5142	11	116,207,060	C	**T**	0.066189148	0.09
	rs133029	22	36,906,261	**C**	T	-0.023849579	0.10
TC	rs11591147	1	55,278,235	**G**	T	-0.360567	0.02
	rs2516448	6	31,498,389	**C**	T	-0.024023	0.46
	rs17211510	6	32,710,408	C	**A**	0.0354316	0.26
	rs28635570	8	126,575,814	C	**T**	0.0429184	0.23
	rs2280845	9	130,622,991	**C**	T	-0.033044	0.25
	rs7396835	11	116,189,238	C	**T**	0.0598212	
	rs7396851	11	116,189,374	C	**T**	0.0595118	
	rs5142	11	116,207,060	C	**T**	0.0645833	0.09
	rs6602910	13	113,564,928	A	**G**	0.0338499	0.38
	rs289716	16	55,566,877	A	**T**	0.0308744	0.32
	rs289718	16	55,567,433	T	**C**	0.0260831	0.32
	rs2228603	19	19,190,924	G	A	-0.122697	

*Alleles in bold type refer to coded alleles; MAF: minor allele frequency; Effect sizes that refer to the change of phenotype per minor allele were identified in the Asselbergs gene-centric meta-analysis