1.a. Full Title: Relationship between single nucleotide polymorphisms previously associated with lipid levels, HDL-C or triglyceride extreme levels, and atherosclerotic events.

2. Writing Group:
   Writing group members:
   Ariel Brautbar, MD
   Christie Ballantyne, MD
   Alanna Morrison PhD
   Ron Hoogeveen PhD
   Salim Virani MD
   Weihong Tang, MD, PhD
   Lloyd Chambless PhD
   Eric Boerwinkle, PhD

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

First author: Ariel Brautbar
Address: Ariel Brautbar, Baylor College of Medicine, One Baylor Plaza, M.S. BCM225, Houston, TX 77030

   Phone: 713-907-6222   Fax:
   E-mail: brautbar@bcm.tmc.edu

Corresponding/senior author (if different from first author correspondence will be sent to both the first author & the corresponding author):
   Christie Ballantyne
   Address: Baylor College of Medicine, 6565 Fannin, M.S. A656, Houston, TX 77030

   Phone: 713-798-5034   Fax: 713-798-3057
3. **Timeline**: Analysis to start as soon as approval is obtained. Manuscript is to be prepared as soon as analysis is available. We hope that the analysis and manuscript preparation will take place within 1 year from approval of the proposal.

4. **Rationale**: A number of epidemiological studies including the Atherosclerosis Risk in Communities (ARIC) study have shown association between high-density lipoprotein cholesterol (HDL-C), triglycerides (TG) and atherosclerotic disease.

1. **HDL-C**: A strong inverse association exists between levels of HDL-C and the development of incident coronary heart disease (CHD) events. Clinical trials have shown that high levels of HDL-C on statin therapy are associated with lower frequency of cardiovascular events and less atherosclerosis. In addition to the epidemiological evidence and data from clinical trials that suggest that high levels of HDL-C are protective against atherosclerosis, there is a wealth of experimental data that suggest that therapies which modulate HDL-C levels or functions may prove beneficial in the treatment of atherosclerosis and may enhance regression of disease and reduce CHD events.

   Although the epidemiological data shows that the inverse relationship between levels of HDL-C and development of CHD is at least as strong as the positive relationship between levels of low-density lipoprotein cholesterol (LDL-C) and development of CHD, the clinical phenotype of patients with mutations in genes that involve HDL metabolism is far more heterogeneous than the clinical phenotypes of patients who have mutations in genes involving LDL metabolism. For example, mutations in the gene that encodes the LDL receptor and mutations in the gene that encodes for apolipoprotein (apo) B are both associated with increased plasma levels of LDL-C and increased risk for CHD. Thus far, all genetic mutations that lead to very high levels of plasma LDL-C are associated with increased risk for CHD. In contrast, some mutations that lead to very low levels of HDL-C, such as a deletion of the gene encoding for apo A-I, are associated with an increased risk for CHD, whereas others, such as the apo A-I Milano mutation, appear to confer no risk and may be protective against atherosclerosis. Furthermore, mutations which are associated with very high levels of HDL-C such as in the mutations in the gene that encodes for CETP have not been consistently associated with protection against atherosclerosis.

   We hypothesize that investigation of groups divided according to upper and lower levels of blood HDL-C and then by the presence or absence of atherosclerosis, for frequency of SNP’s in genes that have been previously associated with HDL-C levels, will demonstrate significant differences in allele frequency between the groups, for certain tested SNPs. Establishing association between SNP’s in genes which are known to be associated with HDL-C levels and the presence or absence of atherosclerosis in the tails of HDL-c levels in ARIC may provide insight into molecular pathways which regulate both HDL metabolism and atherosclerosis.

2. **TG**: 
Recently multiple association studies have identified new SNP’s associated with TG blood levels. However, clinical significance of these SNP’s, in context of high or low levels of TG has not been examined. This would seem an important question to address as certain clinical syndromes that are characterized by elevated triglycerides have no association with atherosclerotic disease (type 1 hyperlipidemia) while others do (type 5 hyperlipidemia). We hypothesize that investigation of groups divided according to upper and lower levels of blood TG and then to the presence or absence of atherosclerosis, for frequency of SNPs in genes that have been previously associated with TG levels will demonstrate significant differences in allele frequency between the groups, for certain tested SNPs. Establishing association between SNP’s, in genes previously associated with TG levels and the presence or absence of atherosclerosis in the tails of TG levels in ARIC may provide insight into molecular pathways which regulate both metabolism of TG and triglyceride enriched lipoproteins and atherosclerosis.

5. Main Hypothesis/Study Questions:
For the purpose of this manuscript

1. Atherosclerosis or atherosclerotic event will be defined as:
   1. Prevalent/Incident CHD
   2. ABI <0,9
   3. Visible plaque on carotid ultrasound
   4. Intima–media thickness (IMT) >75% for age, race and gender.

2. SNPs in genes that have been previously (or are thought to be) associated with HDL-C/TG levels will be used for the analysis described in this manuscript.

Hypothesis:

A. HDL-C

A1. Out of the whole study population define the groups with the upper and lower 15 percentile by age and gender:
Hypothesis A1: Certain SNP’s will have significant difference in frequency between the group in the upper 15 percent of blood HDL-C level and the group in the lower 15 percent of blood HDL-C levels

A2. Out of the whole study population define the groups with the upper 15 percent of blood HDL-C level with:
   a. Absence of atherosclerotic disease or clinical event.
   b. Presence of atherosclerotic disease or clinical event.
Hypothesis A2: Certain SNP’s will have significant difference in frequency between the group in the upper 15 percent of blood HDL-C level with atherosclerotic disease, and the group in the upper 15 percent of blood HDL-C level without atherosclerotic disease.
A3. Out of the whole study population define the groups with the lower 15 percent of blood HDL-C level with:
   a. Absence of atherosclerotic disease or clinical event.
   b. Presence of atherosclerotic disease or clinical event.

**Hypothesis A3:** Certain SNP’s will have significant difference in frequency between the group in the lower 15 percent of blood HDL-C level with atherosclerotic disease, and the group in the lower 15 percent of blood HDL-C level without atherosclerotic disease.

B. Triglycerides (TG)

B1. Out of the whole study population define the groups with the upper and lower 15 percentile of blood TG level as defined by age and gender:

**Hypothesis B1:** Certain SNP’s will have significant difference in frequency between the group in the upper 15 percent of blood TG level and the group in the lower 15 percent of blood TG levels.

B2. Out of the whole study population define the groups with the upper 15 percent of blood TG level with:
   a. Absence of atherosclerotic disease or clinical event.
   b. Presence of atherosclerotic disease or clinical event.

**Hypothesis B2:** Certain SNP’s will have significant difference in frequency between the group in the upper 15 percent of blood TG level with atherosclerotic disease, and the group in the upper 15 percent of blood TG level without atherosclerotic disease.

B3. Out of the whole study population define the groups with the lower 15 percent of blood TG level with:
   a. Absence of atherosclerotic disease or clinical event.
   b. Presence of atherosclerotic disease or clinical event.

**Hypothesis B3:** Certain SNP’s will have significant difference in frequency between the group in the lower 15 percent of blood TG level with atherosclerotic disease, and the group in the lower 15 percent of blood TG level without atherosclerotic disease.

**Questions to be addressed in a stepwise manner:**

1. What are the demographics and traditional risk factors (TRF) of subjects included in the following groups:
   a. The upper 15 percent of HDL levels
   b. The lower 15 percent of HDL levels
   (Comparison between demographics/TRF of a,b will be done using the chi square statistics)
c. The upper 15 percent of TG levels
d. The lower 15 percent of TG levels  
(Comparison between demographics/TRF of c,d will be done using the chi square statistics)

e. Upper 15 percent of HDL with atherosclerosis.
f. Upper 15 percent of HDL without atherosclerosis
g. Lower 15 percent of HDL with atherosclerosis.
h. Lower 15 percent of HDL without atherosclerosis  
(Comparison between demographics/TRF of e,f,g,h will be done using the chi square statistics)
i. Upper 15 percent of TG with atherosclerosis.
j. Upper 15 percent of TG without atherosclerosis
k. Lower 15 percent of TG with atherosclerosis.
l. Lower 15 percent of TG without atherosclerosis  
(Comparison between demographics/TRF of i,j,k,l will be done using the chi square statistics)

2. What are the different allele frequencies of predefined selected SNP’s previously associated with HDL-C, in subjects in the following groups: High HDL-C levels, low HDL-C levels, High HDL-C levels without atherosclerosis, high HDL-C levels with atherosclerosis, low HDL-C levels with atherosclerosis, low HDL-C levels without atherosclerosis

3. What are the different allele frequencies of predefined selected SNP’s previously associated with TG, in subjects in the following groups: High TG levels, low TG levels, High TG levels without atherosclerosis, high TG levels with atherosclerosis, low TG levels with atherosclerosis, low TG levels without atherosclerosis

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 population will exclude race other than white or black and those for whom DNA use is restricted.

Subjects will be divided for analysis according to
A. HDL-C levels(as defined by age and gender).
B. HDL-C levels and the presence or absence of atherosclerosis.
C. TG levels (as defined by age and gender).
D. TG levels and the presence or absence of atherosclerosis

Demographics and traditional risk factors will be analyzed according to the tables below and will be done separately for HDL-C and TG:

**Characteristics of individuals in the upper (lower) percentage of high (low) HDL-C/TG with (without) evidence for atherosclerosis**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking, years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI, m/kg²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waist/hip ratio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDL-C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total cholesterol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Triglyceride</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMT, mm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**HDL-C/TG within race and sex groups for N=**

<table>
<thead>
<tr>
<th>Race and sex group</th>
<th>Obs</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>White women</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White men</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SNP genotyping data will be derived from previously reported genes associated with HDL-C/TG candidate SNP’s using the Affymetrix 6.0. The SNP frequency data will be compared according to the following tables:

SNP frequency will be compared for each to adjacent columns for each SNP using chi square statistics. Calculations will be separately based on sex and race depending on total number of participants and ability to obtain statistical significance.

Adjustment will be in the context of logistic regression.

<table>
<thead>
<tr>
<th>Allele, Frequency</th>
<th>Lower 15%</th>
<th>Upper 15%</th>
<th>P</th>
<th>Lower 15%</th>
<th>Upper 15%</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SNP 1
SNP 2
<table>
<thead>
<tr>
<th>Allele Frequency</th>
<th>SNP 1</th>
<th>SNP 2</th>
<th>SNP 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>15% Lower</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ath</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower 15%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-ath</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P 15%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ath</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper 15%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-ath</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P 15%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ath</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The following Genes will be included in the analysis (others may be added):
- ABCA1, APOA1/2/3/4/5, APOC1/2/3/4, APOE, CETP, LCAT, PLTP,
- LXR, PPARAlpha, PPARGamma,
- SMPD1, GBA, SCARB2, LIPE, SOAT2, ABCG2, CUBN,
- HDLBP, MTP, NPC, SCARB1/SRB1, LPL, LIPC, LIPG, GALNT2, MMAB, MVK,
- TCF1, MPO, PPARD, TRIB1, PCRC1, SORT1, NCAN, CLIP2, PBX4, VNN1, SIRT1
BMP-1, CTalpha, PON1, THRB, NR1H4, CNR1, GPR109A, MCP-1, ANGPTL4
GCKR, BCL7B, TBL2, MLXIPL, ANGPTL3, DOCK7, ATG4C, MC4R, ABHD5, PNPLA2 (ATGL),
PNPLA3, PNPLA4, CES1, LYPLAL1, PLIN.

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

b. If Yes, is the author aware that the file ICTDER02 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 ICTDER02 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?  ____ 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 ICTDER02 must be used to exclude those with value RES_DNA = “No use/storage DNA”?  ____ 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

____ 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)?

11. a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data?  ____ Yes  ____ 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.