1.a. **Full Title**: Genomewide association study of blood pressure trajectories from midlife to older age: the Atherosclerosis Risk in Communities (ARIC) Study

b. **Abbreviated Title (Length 26 characters)**: GWAS of BP trajectories

2. **Writing Group**:
   Poojitha Balakrishnan, Kunihiro Matsushita, Elizabeth Colantuoni, Anna Kucharska-Newton, Nisa Maruthur, J Hunter Young, Aravinda Chakravarti, Terri Beaty, Others welcome

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

First author: Poojitha Balakrishnan
Address: Johns Hopkins Bloomberg School of Public Health
615 N. Wolfe Street, Suite W6517
Baltimore, MD 21205
Phone: 510-417-0783
E-mail: pbalakr2@jhu.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: Kunihiro Matsushita
Address: Johns Hopkins Bloomberg School of Public Health
615 N. Wolfe Street, Room W6017
Baltimore, MD 21205
Phone: 443-287-8766
Fax: 443-683-8358
E-mail: kmatsush@jhsph.edu

3. **Timeline**:
We aim to submit the manuscript for ARIC publications committee within 12 months from the date of approval of this proposal.

4. **Rationale**:
Hypertension is defined as systolic blood pressure $\geq 140$ mm Hg or diastolic blood pressure $\geq 90$ mm Hg. $^1$ Hypertension is associated with a variety of adverse clinical outcomes including myocardial infarction, congestive heart failure, stroke and kidney disease. $^2$ It affects approximately 80 million adults in the U.S., making it one of the most prevalent modifiable risk factors of cardiovascular disease. $^{1,3}$ In addition, the National Health and Nutrition Examination
Survey (NHANES) 2007-2010 reported 6% of adults over 20 years have undiagnosed hypertension.1

Family history of hypertension is a well-established risk factor for elevated blood pressure.4 The heritability of quantitative measures of blood pressure has been estimated to be 30-80%.5–7 By definition, this heritability could reflect shared genes or shared familial environmental factors. Early physiologic and genetic studies highlighted genes responsible for sodium reabsorption.8–13 In fact, genes regulating sodium reabsorption impact 4 of the 5 major categories of monogenetic hypertension.7,11 However among the 100+ genes identified in genome-wide association studies (GWAS), most genes do not involve sodium reabsorption.7,14 Moreover, these studies have been successful in explaining only about 10% of variation in blood pressure.7,15,16

Part of the unexplained variation in blood pressure may be attributed to the use of single visit blood pressure in most epidemiological studies.17–22 The investigation of blood pressure changes over time is limited, especially in genetic association studies. Early linkage studies implicated overlapping and novel regions of the genome for longitudinal compared to single blood pressure measurement.23–25 Some preliminary work on blood pressure trends and genetic markers has been reported.26,27 For example, the absence of change in blood pressure over time was associated with some of the recognized candidate genes for blood pressure.26 Some investigators have also used long term averaging for BP measurements over multiple time points and report an increase in statistical power to detect associations with common genetic markers.27 To our best knowledge, the genetics of patterns of blood pressure changes over time have not been assessed in a longitudinal cohort of middle-aged to older participants.

Our objective is to investigate this novel phenotype, blood pressure trajectory classes - patterns of blood pressure changes over time. Identifying common variants associated with blood pressure trajectories can help account for more of the variation in blood pressure and possible pathogenesis of hypertension.

5. **Main Hypothesis/Study Questions:**
Common genetic variants are associated with blood pressure trajectory classes.

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:** Genome-wide association study

**Inclusion criteria:** All ARIC participants with genetic data and blood pressure measured in at least two of visits 1-5 with data on covariates and whether participants were taking blood pressure lowering medications (N = 11,533; 8,767 European Americans and 2,766 African Americans)

**Outcome:** Blood pressure trajectory classes derived from blood pressure measurements from visits 1-5 using latent class analysis, as detailed subsequently, and will be investigator-labeled
(e.g., elevated at baseline – increasing over time, elevated-stable, moderate-increasing, moderate-stable, and low-stable).

**Predictor:** Imputed genotype dosage using 1000 Genomes reference panels

**Covariates:** Gender, center, body mass index, principal components of genetic ancestry associated with outcome at p<0.05

**Statistical analysis:** All of the statistical analyses described below will be performed separately for systolic and diastolic blood pressure, but given its stronger association with clinical outcomes, we are primarily interested in systolic blood pressure. 28-30

We will use latent class analysis to create the blood pressure trajectory class phenotype - individuals grouped according to patterns of changes in blood pressure over time. For those taking antihypertensive medication, blood pressure will be calibrated using the most reliable method (e.g., constant calibration factor and truncated normal regression31-34) currently under investigation in our previous ARIC manuscript proposal #2394 entitled “Determinants of blood pressure trajectories from midlife to older age: the Atherosclerosis Risk in Communities (ARIC) Study”. The latent class models provide posterior probabilities based on the measured blood pressure for individuals without treatment and the calibrated blood pressure for individuals with treatment. 35 These posterior probabilities will be used to cluster individuals and each individual will be assigned to the trajectory class with the largest posterior probability. 35 The best fitting latent class model will be determined with a modified Bayesian information criterion (BIC) as outlined by Andurf, et al. 36

We will then use the blood pressure trajectory classes as the phenotype in multinomial logit models to test for association with common variants, stratified by race group. Significance threshold will be set at the conventional genome-wide significance of 5x10-8. As secondary analysis we will also perform a meta-analysis of the results from the European Americans and African Americans. We will also aim to explore gene set enrichment analysis using these GWAS results. The p-values from the GWAS will be used as input. We will prioritize candidate pathways including genes with the top GWAS hits in the analysis. The results will demonstrate whether these candidate pathways are enriched for statistical significance more than expected by chance.

The latent class analysis will be performed using the ‘traj’ plugin in STATA version 12.37,38 The multinomial logit regression and GWAS will be done with R. 39 MAGENTA will be used for the gene set enrichment analysis. 40

**Strengths and Limitations:** This study proposes to investigate the genetics of blood pressure trajectory classes, which has not been investigated before. Furthermore the use of latent class analysis in the setting of GWAS provides a novel method to investigate longitudinal phenotypes. A limitation is the availability of limited number of studies for replication.

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

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)?

No previous ARIC proposal focuses on predicting blood pressure trajectory classes using genome-wide marker information. ARIC Manuscript Proposal #2146 (Systolic blood pressure trajectories and incident cardiovascular disease) explores the association of blood pressure trajectories between visits 1 and 4 with subsequent cardiovascular outcomes. Our ARIC Manuscript Proposal #2394 (Determinants of blood pressure trajectories from midlife to older age: the Atherosclerosis Risk in Communities (ARIC) Study) investigates the association between blood pressure trajectories and non-genetic risk factors. We believe this proposal will complement both MP2146 and MP2394 since the focus is to evaluate genetic determinants of blood pressure trajectory classes. Also key investigators from those two proposals participate in this proposal.

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

__X__ A. primarily the result of an ancillary study
(list number* 2006.04)

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:


38. StataCorp. 2011. *Stata statistical software: Release 12*. College Station, TX: StataCorp LP.
