Blood Pressure Polygenic Scores, Target Organ Damage, and Antihypertensive Drug Effects: Insights from longitudinal Real-World Clinical Data

by

Yusuph M. Mavura

DISSERTATION

Submitted in partial satisfaction of the requirements for degree of

DOCTOR OF PHILOSOPHY

in

Epidemiology and Translational Science

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Neil Risch
E3D0E6171DC74F3...

Neil Risch

Chair

DocuSigned by:
Maria Glymour
[illegible]424...

Maria Glymour

DocuSigned by:
Akinyemi Oni-Orisan
[illegible]7A0341B...

Akinyemi Oni-Orisan

DocuSigned by:
[illegible]492...

Catherine Tcheandjieu

DocuSigned by:
99419627041F4D9...

Franklin Huang

Committee Members

Dedication

This dissertation is dedicated to my family, the late Prof. Mavura, Dr. Mavura and Beresha Mavura

Acknowledgements

I am deeply indebted to my dissertation committee—Dr. Neil Risch, Dr. Maria Glymour, Dr. Akinyemi Oni-Orisan, Dr. Catherine Tcheandjieu, and Dr. Franklin Huang—for your guidance, constructive feedback, and encouragement, all of which were essential to this dissertation. Indeed, I stand on the shoulders of giants. I offer special thanks to Neil and Maria for believing in me, welcoming me into the program, and training me in Human Genetics, Epidemiology and Public Health. I am profoundly grateful for the steadfast support and belief of my extended family and friends in the USA, Tanzania, and Kenya. To my parents, thank you for your endless love, motivation, and sacrifices, which have made this journey possible. My late dad, Prof. Mavura, was a major reason I pursued a PhD, and his influence has shaped me into the person I am today. As he often told me, "show them what you are made of," and I believe this dissertation is a testament to that. To my sister, Beresha, your support and motivation have been a constant source of strength. To my partner, Sally Ireri, thank you for your patience, understanding, and unceasing encouragement. I extend my sincere appreciation to the University of California San Francisco Department of Epidemiology and Biostatistics for their invaluable support, and to my colleagues in the department—especially my cohort-mates—whose camaraderie made this endeavor possible. Finally, I acknowledge the contributions of everyone whose prior work on the Kaiser Permanente (KP) Research Program on Genes, Environment, and Health (RPGEH) Genetic Epidemiology Research on Adult Health and Aging (GERA) has informed this dissertation. To all who have contributed in any way to my life and academic journey, thank you.

Blood Pressure Polygenic Scores, Target Organ Damage, and Antihypertensive Drug Effects: Insights from longitudinal Real-World Clinical Data

Yusuph Mnaya Mavura

Abstract

High blood pressure is a leading cause of disability-adjusted life years (DALYs) worldwide and a major risk factor for adverse health outcomes, including cardiovascular disease, kidney disease, heart failure, and all-cause mortality. Blood pressure (BP) is highly heritable, with narrow-sense heritability estimates of at least 0.2 for systolic and diastolic blood pressure and over 1,000 genome-wide association study (GWAS) loci identified. Polygenic risk scores (PRS) summarize an individual's genetic predisposition to a trait or disease based on the combined effects of genetic variants. While PRS have potential applications in disease risk prediction and population screening, their performance varies based on the context of evaluation. Factors such as age, sex, medication use, ancestry, and their interactions influence the predictive power of BP PRS, which often perform worse in underrepresented populations. This dissertation utilizes longitudinal, real-world clinical data to examine the prediction accuracy of BP PRS across sex, age, and ancestry groups, assess their associations with target organ damage outcomes such as chronic kidney disease, and evaluate BP response to first line monotherapy antihypertensive medications by ancestry and drug class.

In the first chapter of my dissertation, I investigate the associations between multi-ancestry polygenic risk scores (PRSs) for blood pressure (BP) traits in a real-world EHR-based clinical cohort by sex, antihypertension medication use, and ancestry. I then looked at the effect of age on PRS association with BP traits and how variance explained by PRS changes with age across

ancestry groups. Summary statistics from the meta-analysis of UKB, ICBP, and BioVU, COGENT and Biobank Japan were used to train multi-ancestry PRSs for diastolic blood pressure (DBP) and systolic blood pressure (SBP), using PRS-CSx. I then tested these PRSs in the Kaiser Permanente (KP) Research Program on Genes, Environment, and Health (RPGEH) Genetic Epidemiology Research on Adult Health and Aging (GERA) cohort, in which individuals had multiple office visit BP measures across their life-course. The primary outcomes were DBP and SBP measurements. I tested the exposures (DBP and SBP PRS) on the outcomes using linear mixed models, stratified by sex (female/ male), ancestry group (African American, East Asian, European, Latino(a)) controlling for relevant covariates. I further stratified analyses by evenly sized age-bins. In a sensitivity analysis using individuals with both treated and untreated measures, I further stratified analyses by antihypertension (AHT) medication use (whether the BP measure was treated or not). I found significant differences in BP PRS prediction accuracy, by ancestry, sex, age across ancestries and antihypertension medication use. The estimated associations and prediction accuracy measured by percent variance explained (PVE) by the PRSs for BP traits varied across strata of ancestry, sex. Generally, females had higher PVE then males. The highest PRS PVE was in Latino(a) ancestry (female) with PVE of 2.27% in SBP and 2.14% DBP, followed by European, 1.92 % SBP and 1.7% DBP, then East Asian at 1.13% SBP and 1.7% DBP, and African American ancestries at 0.42 % SBP, and 0.76% DBP. PVE peaked at midlife and tailed off, except in Afr SBP-PRS where the accuracy dropped off drastically from early age. The PVE was higher in BP outcome measures that were not treated compared to treated across all the ancestry groups.

In a real-world clinical multi ancestry cohort of ~100 000 US adults (GERA) within the same healthcare system, I show significant differences in multi-ancestry BP PRS prediction accuracy (measured by PVE) by ancestry, sex, age and antihypertension treatment status within Afr, Eas, Eur, Lat ancestry groups. This highlights the challenges of using multi-ancestry BP-PRS in real-world clinical settings, even though the PRS is constructed using multiple ancestry GWAS with large samples sizes.

In the second chapter of my dissertation, I investigated the association between multi-ancestry PRSs BP and the incidence of target organ damage (TOD) clinical phenotypes at any age: chronic kidney disease (CKD), and cardiovascular disease (CVD) outcomes: Ischemic stroke (IS), heart failure (HF), atrial fibrillation (Afib), cardiomyopathy (CMP), coronary artery disease (CAD) and myocardial infarction (MI) in a real-world clinical cohort across African (Afr), East Asian (Eas), European (Eur), and Latino(a) (Lat) ancestry groups. I constructed multi-ancestry PRSs for SBP and DBP using summary statistics from genome-wide association studies (GWAS) of the UKB, ICBP, BioVU meta-analysis, the COGENT collaboration, and Biobank Japan. I then used these PRSs to test their association with TOD incidence in the Kaiser Permanente (KP) Research Program on Genes, Environment, and Health (RPGEH) Genetic Epidemiology Research on Adult Health and Aging (GERA) cohort. The primary outcome was incidence of each TOD, with age as the time scale in the analysis. Individuals entered the risk set at their age of entry into the cohort and were followed until the age at which they experienced the TOD or were censored, excluding those with prevalent disease at their age of entry. I used Cox proportional hazard models, stratified by ancestry group (Afr, Eas, Eur, Lat) and adjusted for sex, genetic principal components, diabetes, and respective PRS of the target organ damage

phenotype obtained from the polygenic score catalogue (PGS catalogue, which is an open database of published polygenic scores that provides consistently annotated metadata for each score). An increase of 1 SD of the SBP PRS resulted in a 22%, 21%, 13%, 11% higher risk in incidence of CKD at any age in Lat, Afr, Eur, Eas ancestries respectively. This was statistically significant associated. However, the BP PRS association with the incidence of the CVD related TOD phenotypes was weaker, and not consistently statistically significant across ancestries. The effect of SBP PRS was found to be largely independent of PRS of the target organ damage phenotypes and diabetes status at start of follow up for CKD as the association only slightly deteriorated when added to the models. However, for ischemic stroke, the HR_{adj} became statistically insignificant after addition of stroke TOD PRS. The correlation coefficient between SBP PRS and stroke TOD PRS was observed to be high across all ancestries. BP-PRS is associated with increased risk in incident CKD TOD at any age across ancestries, independent of TOD PRS and diabetes status at start of follow up. This association was less apparent in CVD related TOD outcomes.

Finally, in the final chapter of my dissertation, I investigated ancestry differences in individual level response to monotherapy blood pressure medication treatment in a real-world clinical cohort with longitudinal EHR data by drug class type and dose in individuals of African, East Asian, European and Latino(a) ancestry with hypertension. I used prescription data from the Kaiser Permanente Research Program on Genes, Environment, and Health (RPGEH) Genetic Epidemiology Research on Adult Health and Aging (GERA) cohort, linking it with BP measurements where individuals had multiple internal medicine office visit BP measures across their life-course. We investigated drug classes—Angiotensin-Converting enzyme (ACE)

inhibitors, hydrochlorothiazide (HCTZ), and dihydropyridine calcium channel blockers (CCBs)—and calculated the average change in systolic (SBP) and diastolic blood pressure (DBP) due to treatment by monotherapy AHT. I used ANOVA and Tukey's Post hoc tests to test the differences in average change by ancestry group. I found that estimated average BP reductions varied depending on ancestry, drug class. Generally, estimated SBP reductions were more modest compared to those seen in clinical trials. The largest average reduction in SBP was seen with HCTZ in individuals of African ancestry, while the smallest was with ACE inhibitors in the same group. African ancestry individuals showed larger SBP reductions with HCTZ, but statistically significant smaller reductions with ACE inhibitors compared to other ancestry groups. The findings suggest that estimated BP reduction from antihypertensive medications vary by ancestry, drug class, in real-world clinical cohort settings. This challenges the common assumption in GWAS studies that BP medication effects are uniform (e.g., a 15 mm Hg reduction in SBP) across all ancestry populations and drug types, highlighting the need for more tailored approaches in clinical and genetic research e.g. in genome-wide association studies (GWAS) studies correcting for BP medication use.

Table of Contents

1. Chapter 1: Association of polygenic risk scores (PRS) with blood pressure in a real-world electronic health record (EHR) based clinical cohort. 1

Background *2*

Methods *4*

GWAS Summary Statistics Datasets 4

GERA Dataset 4

Multi-Ancestry PRS construction 5

Statistical analysis: Testing PRS across strata of Ancestry, Sex, AHT treatment status in GERA 7

Results *10*

Characteristics of GERA participants 10

PRS association with BP by ancestry group, sex, and antihypertension medication use 11

Discussion *15*

Conclusion *18*

2. Chapter 2: Association of multi-ancestry polygenic risk scores for blood pressure with the incidence of target organ damage in a real-world clinical cohort. 35

Introduction *36*

Methods *38*

GERA population 38
Exposure and outcomes 38
Multi-Ancestry PRS construction 38
Inclusion / Exclusion criteria 40
Statistical Analyses 40
Results *42*
PRS Association with incidence of target organ damage across ancestry groups 42
Discussion *45*
Conclusion *47*
3. Chapter 3: Estimated changes in blood pressure from monotherapy antihypertension medication treatment in a real-world cohort. 57
Background *58*
Methods *60*
Dataset 60
Main outcomes and measures 60
Statistical Analysis 61
Results *63*
Characteristics of GERA cohort (those diagnosed with essential hypertension) 63
Average SBP and DBP changes due to monotherapy treatment by ancestry 64
Discussion *67*
Conclusion *70*

References 85

Appendix 92

Table A.1. List of ICD codes used to define outcome incidence: *92*

Table A.2. Drug conversions for 1 daily dose (DD) by different drug classes. *97*

List of Figures

Figure 1.1: Blood pressure measurements inclusion and exclusion criteria in GERA cohort..... 21

Figure 1.2: Number of imputed SNPs from each sequencing array in GERA, from summary statistics, and intersection, used in training data for the PRS. 22

Figure 1.3: Schematic showing steps taken for PRS construction and testing 23

Figure 1.4: Percent variance explained (PVE) by SBP and DBP PRS using difference in marginal R^2 in nested models, stratified by sex. 24

Figure 1.5: SBP PRS estimates by ancestry, sex 25

Figure 1.6: DBP PRS estimates by ancestry, sex and antihypertension treatment status 26

Figure 1.7: SBP PRS percent variance explained within each of the age bins, by ancestry..... 27

Figure 1.8: DBP PRS percent variance explained within each of the age bins, by ancestry. 28

Figure 1.9: SBP across age bins in GERA cohort 29

Figure 1.10: DBP across age bins in GERA cohort..... 30

Figure 1.11: SBP PRS estimates by ancestry, across 10 age bins 31

Figure 1.12: DBP PRS estimates by ancestry, across 10 age bins..... 32

Figure 1.13: SBP PRS percent variance explained within each of the age bins, by ancestry and treatment status, among those who had both treated and untreated measures. 33

Figure 1.14: DBP PRS percent variance explained within each of the age bins, by ancestry and treatment status, among those who had both treated and untreated measures. 34

Figure 2.1: SBP and DBP PRS standardized within ancestry groups in GERA cohort 52

Figure 2.2: Standardized scored PRS of select target organ damage and BP phenotypes in GERA. The PRSs were downloaded from the PGS catalogue, an online database with published PRS 53

Figure 2.3: HR_{adj} for incident TOD outcomes with 1 SD increase in SBP PRS stratified by ancestry 54

Figure 2.4: HR_{adj} for incident TOD outcomes with 1 SD increase in DBP PRS stratified by ancestry 55

Figure 2.5: Extra incident cases per 10,000 person years with 1 SD increase of SBP PRS by ancestry groups in GERA cohort 56

Figure 3.1: Schematic of blood pressure measurements and AHT treatments included in this study from the GERA cohort. 78

Figure 3.2: Histogram of average within-individual systolic blood pressure (SBP) changes (treated vs. untreated) for ACE-inhibitors across ancestries, and P-value for Anova. 79

Figure 3.3: Histogram of average within-individual diastolic blood pressure (DBP) changes (treated vs. untreated) for ACE-inhibitors across ancestries, and P-value for Anova. 80

Figure 3.4: Histogram of average within-individual systolic blood pressure (SBP) changes (treated vs. untreated) for HCTZs across ancestries, and P-value for Anova. 81

Figure 3.5: Histogram of average within-individual diastolic blood pressure (DBP) changes (treated vs. untreated) for HCTZs across ancestries, and P-value for Anova. 82

Figure 3.6: Histogram of average within-individual systolic blood pressure (SBP) changes (treated vs. untreated) for CCBs across ancestries, and P-value for Anova. 83

Figure 3.7: Histogram of average within-individual diastolic blood pressure (DBP) changes (treated vs. untreated) for CCBs across ancestries, and P-value for Anova. 84

List of Tables

Table 1.1: Summary statistics for multi-ancestry PRS construction 19

Table 1.2: Characteristics of GERA cohort 20

Table 2.1: Characteristics of GERA cohort at time of start of follow, and primary outcomes ... 48

Table 2.2: Correlation between SBP PRS and TOD PRS in GERA cohort, estimated using Pearson correlation coefficients. 51

Table 3.1: Characteristics of dataset by analysis group 71

Table 3.2: Extended characteristics of dataset by analysis group 72

Table 3.3: Tukey's post hoc test results comparing average systolic blood pressure responses between ancestry groups for ACE inhibitors, CCBs, and HCTZ monotherapies. 74

Table 3.4: Tukey's post hoc test results comparing average diastolic blood pressure responses between ancestry groups for ACE inhibitors, CCBs, and HCTZ monotherapies. 75

Table 3.5: Estimated effect of doubling dose of monotherapy on SBP in mm HG by different drug classes, controlled for sex, genetic PCs, age, diabetes, chronic kidney disease and heart failure status. 76

Table 3.6: Estimated effect on DBP in mm HG of doubling dose of monotherapy by different drug classes, controlled for sex, genetic PCs, age, diabetes, chronic kidney disease and heart failure status. 77

1. Chapter 1: Association of polygenic risk scores (PRS) with blood pressure in a real-world electronic health record (EHR) based clinical cohort.

Background

Recently, it has been shown that multi-ancestry blood pressure (BP) polygenic risk scores (PRS) predict BP traits better than BP-PRS derived from single ancestry and are highly associated with important CVD outcomes such as coronary artery disease in research settings[1]. However, BP-PRS utility in predicting blood pressure traits, and BP sequelae across ancestral populations is not understood. There are concerted efforts to implement equitable use of BP-PRS across all ancestries in clinical contexts, not only in research[2].

PRS prediction accuracy refers to the ability of the PRS to correctly estimate an individual's genetic predisposition to traits (in this case BP) or diseases, and the higher the accuracy, the more reliable it is for risk stratification[3]. PRS are developed through an initial training dataset for example GWAS and GWAS summary statistics and subsequently validated in an independent data set[4]. Prediction accuracy of BP-PRS, and PRS in general may depend on factors linked to the training/base set or factors linked to target/ validation dataset. Several factors related to training set include but not limited to sample size of GWAS/ base data, number of SNPs included in the models, how closely the linkage disequilibrium (LD) in the GWAS summary statistics matches the target population, distributions of true effect sizes and modeling of their prior distributions. Improvements in PRS accuracy in general have been made by developing algorithms that optimize some of the mentioned factors.

Important factors linked to BP-PRS prediction accuracy in the target dataset may include sex, medication use. In addition, PRS predictive accuracy may vary by age in individuals' life-course[1,5]. BP traits measured in an individual may vary with age, due to factors such as antihypertension medication use, lifestyle factors like smoking, diet, co-morbidities, that may have a cumulative and larger effect on BP traits over time. Understanding how PRS association with BP traits and accuracy/ variance explained is influenced by factors such as individuals' measured age, sex, ancestry, and antihypertension use in a real-world clinical cohort will inform its utility for clinical use.

The objective of this study is to investigate the association of multi-ancestry PRSs with BP traits in a clinical real-world cohort (GERA) and examine differences in association and prediction accuracy by age, sex, ancestry, and antihypertension medication use.

Methods

GWAS Summary Statistics Datasets

We used Biobank Japan (BBJ)[6], Continental Origins and Genetic Epidemiology Network (COGENT) consortium[7], and BP GWAS summary statistics from a meta-analysis of UK Biobank (UKB), the International Consortium for Blood Pressure (ICBP), and Vanderbilt University's biorepository of DNA linked to de-identified medical records (Bio VU)[8] as the base data to develop the PRS for blood pressure (SBP, DBP). Details on Biobank Japan[6], COGENT[7] and UKB, ICBP, BioVU[8] are provided elsewhere. Below is a description of the GERA dataset.

GERA Dataset

We used GERA[9] as the test cohort. This dataset includes 97,769 individuals with longitudinal integer blood pressure readings, from 5 different self-reported race/ethnicity groups. In this study, systolic blood pressure (SBP) and diastolic blood pressure (DBP) traits were used as the outcome. The BP readings are EHR-linked from a clinical setting (Kaiser Permanente of North California). The blood pressure readings were restricted to integer measures (excluded ranges) and those measured in the internal medicine department and excluded other departments such as emergency room. BP readings that were considered to be duplicates were removed (if they had the same time stamp and SBP, DBP value). The total number of BP readings that passed QC was 2,329,346 for the cohort (Figure 1.1). Antihypertension treatment (AHT) information was obtained from linked EHR prescription filling/dispensing records. BP readings were considered treated if they occurred within any time-period of AHT drug prescription, that is if the age of BP

measurement falls between the age at AHT drug prescription start, and the age at which the prescription ended.

We used imputed genotype data for all the GERA individuals. Details on genotyping and chips used, QC, imputation can be found at [10–13]. Genotyping was performed using custom-designed Affymetrix Axiom arrays specific to major ancestry groups: African (Afr), East Asian (Eas), European (Eur), Latino(a) (Lat) groups[10,11]. Downstream analyses were stratified based on these groups, as well as an additional group, South Asian (Sas). Individuals in this group were either genotyped in the Eur or Eas group[12,13]. We dropped the Sas ancestry group from analyses since there weren't adequate numbers of individuals for analysis in this sub-stratum.

Multi-Ancestry PRS construction

We used GWAS summary statistics from three sources: 1) a meta-analysis of UKB, ICBP, BioVU, 2) BBJ, 3) COGENT for both SBP and DBP. The base data for the construction of the multi-ancestry PRSs consisted of medication adjusted GWAS summary statistics (Beta coefficients and P-values) from 1) 1,028,980 European ancestry individuals in the meta-analysis study that included UKB, ICBP, BioVU, 2) 31,969 African ancestry individuals from COGENT, 3) 145,515 East Asian ancestry individuals from BBJ (Table 1.1).

SNPs used in construction were intersected between base and test sets, as well as SNPs available in the UKBB LD reference panel used to control for SNP-SNP correlations. We used European, East Asian and African specific UKBB LD reference panel to match the ancestry specific BP GWAS summary statistics in the base data described above. The final number of SNPs used was 980,722 across all ancestry groups in the base and test data (Figure 1.2).

The PRSs were developed using the PRS-CSx[14] tool, a Bayesian regression method for PRS generation which considers minor allele frequency (MAF) and linkage disequilibrium (LD) information in its priors. PRS-CSx is designed for cross population prediction and performs best compared to other contemporary PRS generating tools using other algorithms such as clumping and thresholding which are limited in settings where genetic architecture of populations in base datasets notably differ (such as African and European populations)[14].

The workflow for PRS generation using PRS-CSx is shown in figure 1.3 (Figure 1.3). PRS-CSx is a Bayesian approach to computing PRS that derives each genetic variant's posterior mean effect size (weight) from the prior GWAS base data while accounting for the LD using the 1000 Genomes or UKBB as the reference population. We ran the PRS-CSx META model, with hyperparameter phi value set to auto. The hyperparameter value 'phi' controls the sparsity of genetic architecture in the Bayesian models. All analyses were performed using the META PRS (DBP and SBP) with hyperparameter phi value set to "Auto".

We generated combined SNP effect sizes across populations using an inverse-variance-weighted meta-analysis of the population-specific posterior effect size estimates using the flag '-META'. After computing posterior effect sizes, we used Plink 1.9[15] for allelic scoring, to compute the PRS for each of the test set's ancestry analysis groups Afr, Eas, Eur, Lat.

To facilitate comparison of PRS effect sizes across all ancestry/ analysis groups, we first standardized the BP PRS within each ancestry group, recentering each PRS distribution to mean

0 and SD 1. This also prevents any spurious association that may be contributed by allele frequency differences across the ancestral groups.

Statistical analysis: Testing PRS across strata of Ancestry, Sex, AHT treatment status in GERA

We estimated the effect sizes of the association between PRS and DBP/SBP using linear mixed models for both DBP and SBP with individuals' longitudinal BP measurements. The linear mixed models assume uncorrelated covariance structure between the random intercept and slope. The models were stratified by ancestry analysis groups (Eur, Afr, Lat, Eas), sex (restricted to male or female). Age at BP measurement was normalized across all ancestry analyses groups (for everyone combined), recentering each age distribution to mean 0 and SD 1. The model covariates included normalized age, normalized age-squared body mass index (BMI), smoking status (ever, never, or current smoker), the first 10 and 6 genetic PCs for Eur and Afr, Lat, Eas, respectively. An increase in 1 standard deviation therefore represents 11.97-year change from the mean of 70.29 years of age at BP measurement.

Variance explained by the predictors and interactions were evaluated using marginal pseudo R^2 [16]. The difference between marginal R^2 in nested models, with and without PRS as a predictor, corresponds to the variance explained by the PRS in the models where the dependent variable is either SBP or DBP.

In an analysis to investigate the effect of age on PRS estimates and variance explained, we fit linear mixed models as described above, stratified by ancestry analysis group, and age bins. In

these analyses sex was added as a covariate in the model, as opposed to stratifying by sex. A total of 10 evenly spaced age bins were constructed between the ages of 18 and 90 across all the ancestry analysis groups. Estimates, and PVE was then compared between the different age bins. In a sensitivity analysis, we restricted to individuals with both treated and untreated BP measures, then further stratified the analysis by treatment status.

Each model for calculating marginal R^2 was bootstrapped 25 times in order to estimate uncertainty around the mean marginal R^2 difference and to calculate differences in estimated PVE by either sex or age bins. Bootstrapping is a resampling technique used to estimate the variability of a statistic, such as marginal R^2, by repeatedly sampling with replacement from the original dataset. In this study, models were bootstrapped 25 times to generate a distribution of marginal R^2 values, capturing uncertainty around the mean and quantifying variability in the percent variance explained (PVE). This process involved creating new datasets by resampling observations, refitting the models, and calculating marginal R^2 for each iteration. Bootstrapping also allowed for the estimation of differences in PVE across subgroups, such as sex or age bins, providing robust uncertainty measures for comparisons.

Differences in PVE by sex (female vs male) were formally tested using Mann Whitney U tests, and differences in PVE across the age bins were formally tested using Anova tests. We used R statistical software in all the tests above.

In the statistical analyses above, the threshold for statistical significance was determined by an original alpha set at $P< 0.05$, followed by Bonferroni correction due to multiple testing across

strata. For the Anova test, the significance threshold was set to $P< 0.0013$, and for the Mann Whitney U tests, $P< 0.003$.

Results

Characteristics of GERA participants

Of the 97,769 individuals in GERA that passed QC, 79215 (81%) belonged to Eur analysis group, 8023 (8.2%) to Lat group, 7067 (7.2%) to Eas group, 3009 (3.1%) to Afr group and 455 (0.5%) to Sas group. Overall, 42.1% were male and 57.9% female at birth. The mean (SD) age at baseline survey was 61.5 (13.0). The mean (SD) number of office visit BP measurements per individual was 20.7 (14.2), and the mean (SD) number of years of follow up was 11.1 (4.4). Of the 2,329,346 BP measurements (Figure 1.1), 52.98% were taken during treatment ('treated') with an antihypertensive medication and 47.02% were 'untreated'. The proportion of individuals with both treated and untreated BP measures differed by ancestry group. In total, 48,078 (49.2%) had both treated and untreated values. Among individuals in Afr group, 54.9% had both treated and untreated measurements, while Eas subjects had 45.7%, Eur subjects had 49.4%, Lat subjects had 48.3%, and Sas subjects had 44.2%. The characteristics of the cohort stratified by analysis group are presented in Table 1.1.

PRS association with BP by ancestry group, sex, and antihypertension medication use

The highest coefficient estimates of association (largest regression betas) were in Lat, female: 3.6 mm HG per 1 SD increase in SBP PRS and highest percent variance explained (PVE) was in Lat female: 2.3% in SBP PRS (Figure 1.4, figure 1.5, figure 1.6). The lowest estimates of association were in Afr male: 0.8 mm HG increase with 1 SD increase in DBP PRS and lowest PVE was in Afr male: less than 0.5% in SBP PRS (Figure 1.4, figure 1.5, figure 1.6). We show the PVE in figure 1.4, and estimated effect sizes, 95% C.I. for SBP and DBP PRS in Figure 1.5 and figure 1.6 respectively.

The percent variance explained by BP PRS was statistically significantly higher for females compared to males for SBP and DBP (Mann Whitney U test $P < 0.003$). The PVE by DBP PRS was also statistically significantly higher for females compared to males for DBP measures (Mann Whitney U test $P < 0.003$) (Figure 1.4).

The Lat ancestry group had the highest coefficient estimates and PVE by BP -PRS, followed by Eur, then Eas and finally a large drop in the Afr group in measures of SBP & DBP (Figures 1.4-1.6).

The DBP-PRS regression coefficient estimates were generally lower than those seen in SBP-PRS (Figure 1.5), although the PVE was generally higher across comparable strata, with a larger sex difference for DBP compared to SBP.

We investigated the effect of age in the association between the PRSs and BP outcomes and the PVE by the PRS in the models, stratified by ancestry analysis group. We stratified the mixed models into 10 evenly spaced age-bins, from age 18 to 90.

Overall, there was a statistically significant difference in the SBP PRS PVE across age bins within ancestry groups: in Afr, Eas, Eur, Lat (Anova P value <0.0013), and in DBP PRS PVE across age bins within ancestry groups: Afr, Eas, Eur, Lat (Anova P value < 0.0013) (Figures 1.7-1.8).

The coefficient estimates for the association between SBP and SBP PRS seem to be lower in young individuals, peak in mid-life and gradually weaken with age, in individuals from Eur and Lat analysis groups. However, we noted a different pattern in those in the Afr group, whereby the strongest associations, and corresponding PVE, was highest in young individuals, and dramatically weakened with age. In Eas, there appears to be bimodal peaks, one in early middle age and another in early late age.

These patterns of SBP-PRS association with age contrast with SBP measures in GERA which seem to rise linearly and gradually with age across all groups, and in both treated and untreated measures (Figure 1.9, 1.10).

The coefficient estimates for the association between DBP and DBP PRS and PVE by age seem to follow the relationship seen in SBP described above, although there appears to be an increase in association at later age, especially notable in the Lat analysis group. (Figures 1.11, 1.12).

In the sensitivity analysis whereby only those with treated and untreated measures were retained, the associations in treated measures strata were attenuated compared to untreated measures strata, and there is generally reduced variation of the estimates with age (Figures 1.13, 1.14). Overall, treatment with antihypertensive drugs resulted in reduced PVE across all ancestry groups strata as seen from the lower DBP and SBP PRS PVE (Figures 1.13, 1.14). However, there was a high SBP-PVE in the 39-46 age bin in Eas and in the 39-46 age bin for Lat DBP-PRS.

Within analysis strata of ancestry and treatment status, increase in 1 unit of BMI was consistently associated with a statistically significant increase in SBP. This was true across strata of sex, except Afr male both treated and untreated where the association did not reach statistical significance. The largest effect was seen in Eas untreated females (0.84 mmHG increase in SBP per 1 unit BMI, 95% CI 0.74-0.95, P=3.5E-54), while the smallest effect was seen in Eur treated females (0.24 mmHG increase in SBP per 1 unit BMI, 95% CI 0.22-0.27, P=6.1E-99). BMI was less associated with DBP across strata of ancestry and treatment status. Only in the Eur, Eas and Lat untreated strata was there a statistically significant positive relationship between BMI and DBP. Interestingly, there was a statistically significant negative association between BMI and DBP among Eur female treated measurements (0.05 mmHG decrease in DBP per 1 unit BMI, 95% CI 0.07-0.04, P=1.7E-12).

Normalized age at measurement was largely positively associated with SBP across ancestry strata, except Eur untreated males, while normalized age squared at measurement was largely negatively associated with DBP across ancestry strata.

Smoking status (non-smoker as reference vs former smoker, or current smoker) was largely not significantly associated with SBP or DBP, except for the Eur strata which had the most subjects. Here, smoking was associated with lower SBP and DBP. The most significant association was for Eur male treated measurements (0.9 mmHG decrease in SBP in current smoker vs non-smoker, 95% CI 1.33-0.47, P=3.79E-5).

Ancestry PCs were largely not statistically significantly associated with SBP or DBP. Eas PC2, Eur PC1 and 5 were associated with estimated SBP, while Eas PC2, Eur PC2, Eur PC5, Lat PC1 were associated with estimated DBP.

Discussion

We investigated the association of multi-ancestry BP PRS with BP outcome measures in a real-world clinical cohort with comprehensive linked antihypertension prescription data. The multi-ancestry PRS of SBP and DBP were constructed using PRS-CSx, a Bayesian tool using ~1 million SNPs from summary statistics of 3 different single ancestry BP GWAS, African, East Asian, European.

We found the associations between multi-ancestry BP-PRS and BP traits to be highest in those of Latino(a) ancestry, followed by European, then East Asian and African ancestries. The associations also differed by sex across ancestries, with female individuals having higher PVE and larger regression coefficient estimates. There was a significant difference of PRS PVE and association with BP traits with age across all ancestry strata although the patterns differed by ancestry. There appears to be a dramatic reduction of PVE by PRS with age in the Afr strata, while for Eur, Lat there appears to be an increase in PVE from younger age, which peaks in mid-life and then gradually declines with age. In Eas, there appears to be two peaks in SBP-PRS, either side of mid-life. This might indicate instability of the impact of BP PRS with age, which differs with ancestry and treatment status. The associations were significantly higher in BP outcome measures that were not treated compared to treated across all the ancestry groups, especially in Afr.

It appears that the heritability of BP traits in females is greater than in males[17]. This suggests that environmental determinants of BP traits contribute more variability in BP in males compared to females in this real-world and clinical cohort. PVE was reduced in treated measures compared to

untreated measures. This may be explained by the fact that treatment adds a random effect to the BP traits. The degree to which treatment effects are uncorrelated with the baseline BP determines how much the PVE will be reduced. In our case, the reduction appeared significant. We lumped all treatments and doses into a single stratifying variable. It is possible that if treatment effects are predictable based on drug and dose, and such predictions are taken into account, the reduction in PVE for treated measures might be less attenuated.

PVE is highest in those of Latino(a) ancestry, followed by European then East Asian and African ancestries. This might be explained by the fact that the input GWAS summary statistics were derived from majority European population, followed by East Asian and finally African. On the other hand, Latino(a) ancestry individuals are known to have varying amounts of European admixture and having higher LD among SNP markers than Eur population, due to Native American admixture[18,19]. Higher LD may lead to higher posterior betas as high effect SNPs tag many other high effect SNPs as compared to in lower LD populations such as African ancestry populations, and could also contribute to the lower PVE observed in the Afr strata[20].

It is likely that the decrease in PRS PVE after middle age in Eur, Eas, Lat and from young age in Afr could be due to higher heritability of BP traits in middle aged individuals compared to older individuals, as was shown in an age-stratified heritability study done in the Framingham Heart Study families cohort[21]. This could be due to increased cumulative effects of environmental/ non-genetic exposures on BP measures such as diet, exercise, stress. We observed peak PVE in Eur, Eas, Lat around mid-age. This could be due to gene-environment interactions that appear more prominently in middle age.

Our findings corroborate prior findings that BP PRS is more predictive of untreated BP compared to treated BP and in females compared to males[1], and we show that it is consistent across ancestry groups. Mostafavi et al[17] showed DBP PRS difference in PVE by sex in Eur ancestry individuals, with females having statistically higher R^2 compared to males. While differences in DBP and SBP PRS effect estimates by sex in Afr, Lat, Eas, Eur were observed by Kurniansyah et al[1], it was not clear whether there were statistically significant differences in variance explained across the different ancestry groups as well. Similar to our study, this study also found the difference in PVE by sex to be greater in DBP compared to SBP. Similarly**,** Kurniansyah et al[1] showed differences of PRS estimates by age groups (young, mid age and old) although it wasn't apparent if there were statistically significant differences of PVE by age. In contrast to our findings, the study found higher estimates in SBP-PRS in mid-age in Black race-ethnicity participants. It was also not apparent if there were any PVE differences by age across ancestries in that study. Our study findings may differ from other studies investigating multi-ancestry BP PRS as we tested the PRSs on a clinical, real-world cohort, and used multiple measures per individual across their life-course.

We note some limitations in this study. First, the sample size differences between ancestry strata were significant in the GERA test dataset, which increased the 95% C.I. around the effect estimates and PVE across the non-Eur strata compared to the Eur strata. Another limitation is the sample size differences between ancestry groups in the training set or input GWAS summary statistics. The number of individuals in the European summary statistics was significantly greater than for East Asian and African summary statistics. This may be one of the reasons why PVE and estimates were lower in Afr strata compared to others, although Lat had the highest PVE and

estimates even though there was no Latino(a) GWAS summary statistics input in the PRS training set.

Another is the general limited heritability of BP traits and the modest SNP-based heritability. The PRS performance largely depends on the heritability of phenotypes[2].

Nevertheless, these findings show that multi-ancestry BP-PRS associations and PVE of BP traits vary by sex, age, antihypertension use and ancestry. Therefore, caution must be taken in considering the use and implementation of BP-PRS in clinical contexts.

Future research efforts should include construction with GWAS summary statistics with significantly more samples from non-European populations including African, East Asian, South Asian and Latino(a) ancestries, and evaluation of BP-PRS in larger, diverse real-world clinical datasets.

Conclusion

In a real-world clinical multi ancestry cohort of ~100 000 US adults (GERA) within the same healthcare system, we show clear significant differences in multi-ancestry BP PRS prediction accuracy (measured by PVE) by sex, treatment status and importantly age across Afr, Eas, Eur, Lat ancestries. This highlights the challenges of using multi-ancestry BP-PRS in real-world clinical settings.

Tables

Table 1.1: Summary statistics for multi-ancestry PRS construction

Population of individuals in GWAS where estimates were calculated	Cohort	N of individuals in GWAS	LD reference panel used (N)
African (AFR)	COGENT (meta-analysis)	31,969	UKBB AFR (7,507)
East Asian (EAS)	BioBank Japan	145,515	UKBB EAS (2,181)
European (EUR)	UKBB+ ICBP+ BioVU (meta-analysis)	1,028,980	UKBB EUR (375,120)

Table 1.2: Characteristics of GERA cohort

	Analysis group				
	afr	**eas**	**eur**	**lat**	**sas**
N (% of total)	3009 (3.08)	7067 (7.23)	79215 (81.02)	8023 (8.21)	455 (0.47)
Age at survey	58.88 (12.66)	56.53 (13.77)	62.57 (12.62)	56.42 (13.99)	52.94 (13.47)
BMI at survey	29.32 (6.01)	24.58 (4.16)	26.89 (5.33)	28.08 (5.69)	25.05 (3.82)
N Female (%)	1781 (59.19)	4059 (57.46)	45782 (57.84)	4795 (59.79)	179 (39.34)
Mean number of BP measurements per individual (SD)	22.51 (15.62)	18.51 (13.02)	20.77 (14.08)	21.71 (15.41)	17.73 (12.77)
Mean range in years between first and last BP measurements per individual (SD)	11.15 (4.44)	11.39 (4.24)	11.01 (4.44)	11.44 (4.32)	11.07 (4.59)
Mean SBP (SD) Treated	131.4 (16.03)	128.14 (15.55)	128.38 (15.92)	128.86 (15.61)	125.83 (15.84)
Mean SBP (SD) Untreated	127.47 (15.21)	122.9 (15.32)	125.02 (14.9)	123.77 (14.9)	120.45 (14.45)
Mean DBP (SD) Treated	72.47 (10.77)	70.73 (10.52)	70.09 (10.5)	70.28 (10.66)	69.98 (11.08)
Mean DBP (SD) Untreated	73.95 (10.05)	71.48 (10.07)	71.66 (9.87)	71.88 (10.0)	71.19 (9.72)
% Diagnosed with Hypertension	80.09 %	60.91 %	66.25%	61.33 %	54.73 %

Figures

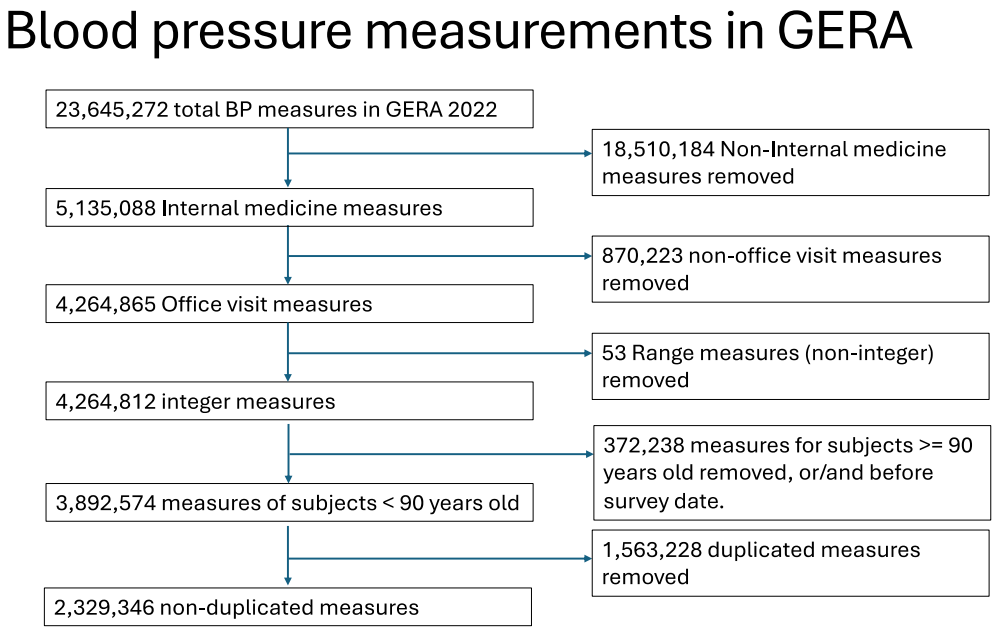

Figure 1.1: Blood pressure measurements inclusion and exclusion criteria in GERA cohort

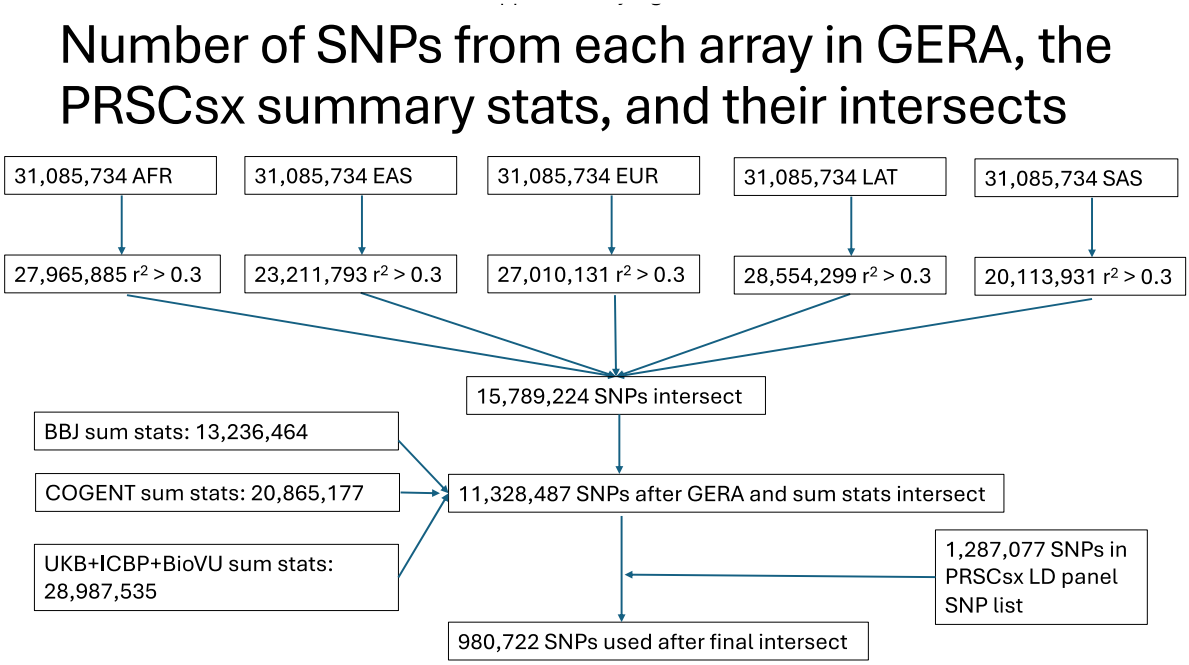

Figure 1.2: Number of imputed SNPs from each sequencing array in GERA, from summary statistics, and intersection, used in training data for the PRS.

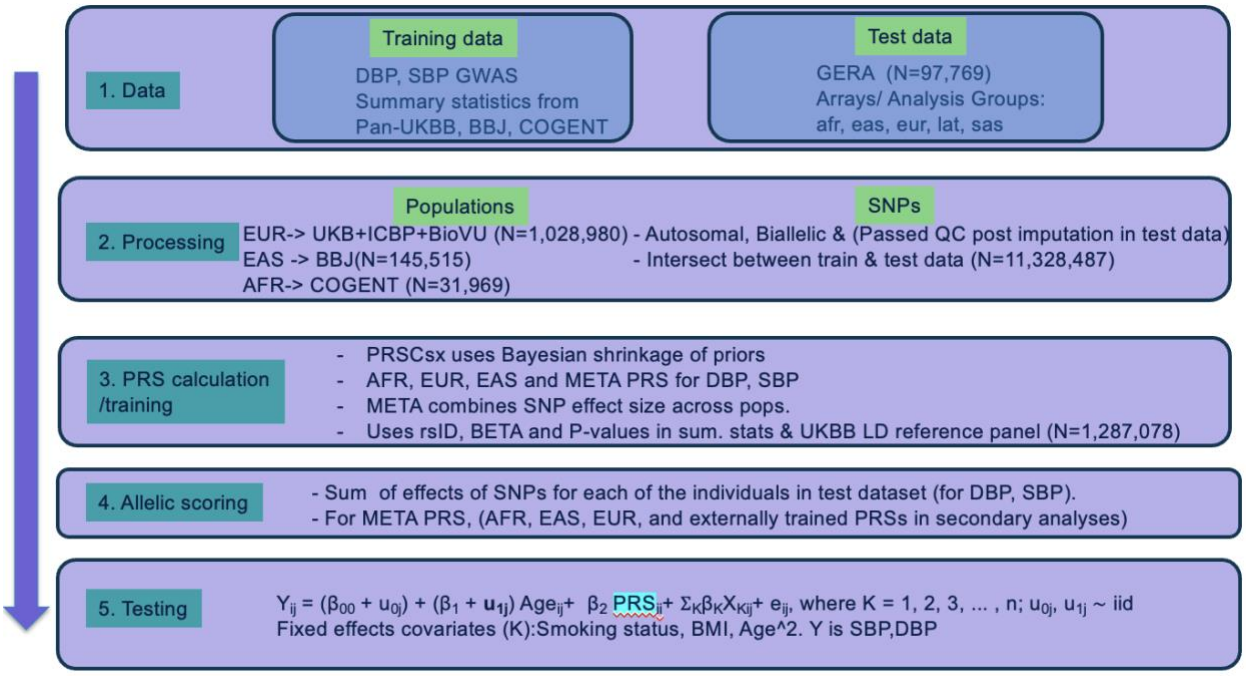

Figure 1.3: Schematic showing steps taken for PRS construction and testing

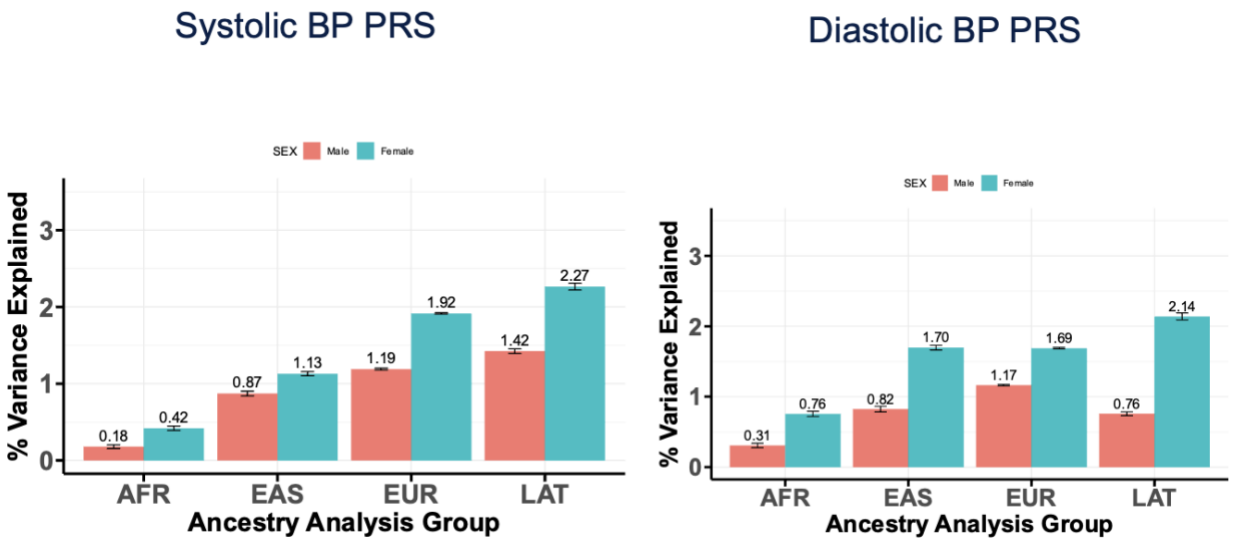

Figure 1.4: Percent variance explained (PVE) by SBP and DBP PRS using difference in marginal R^2 in nested models, stratified by sex.

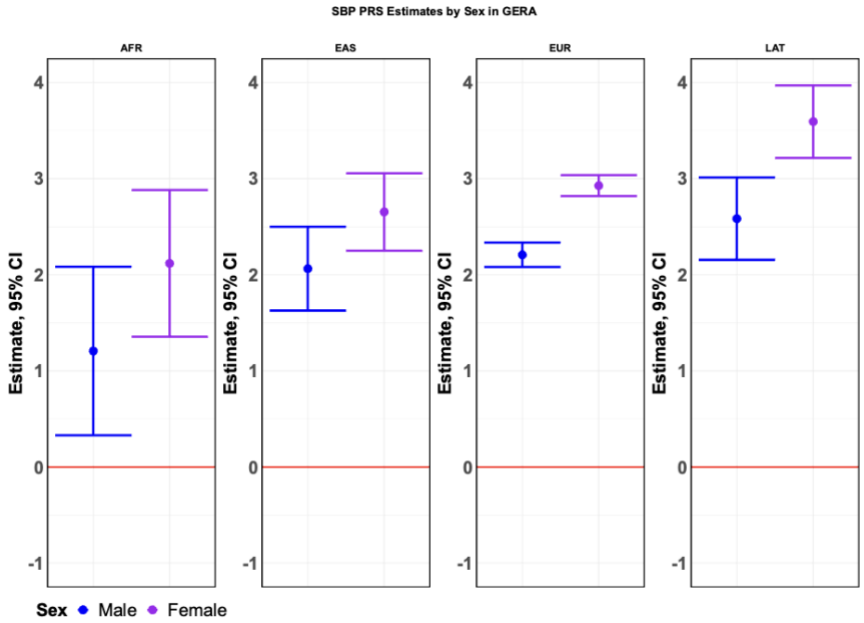

Figure 1.5: SBP PRS estimates by ancestry, sex

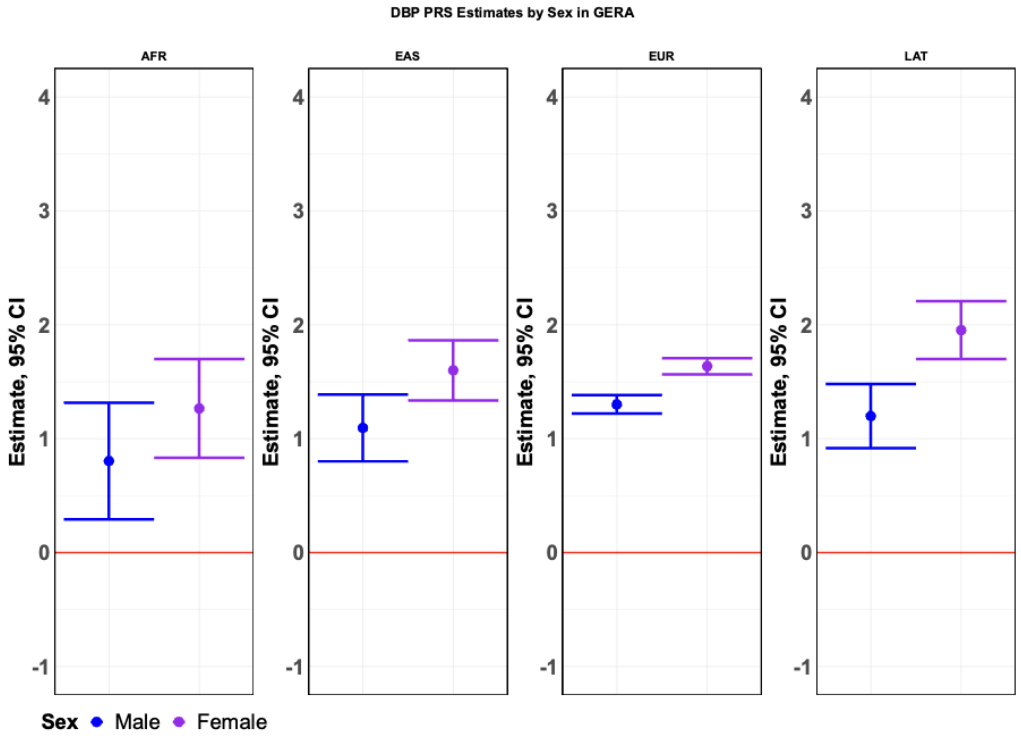

Figure 1.6: DBP PRS estimates by ancestry, sex and antihypertension treatment status

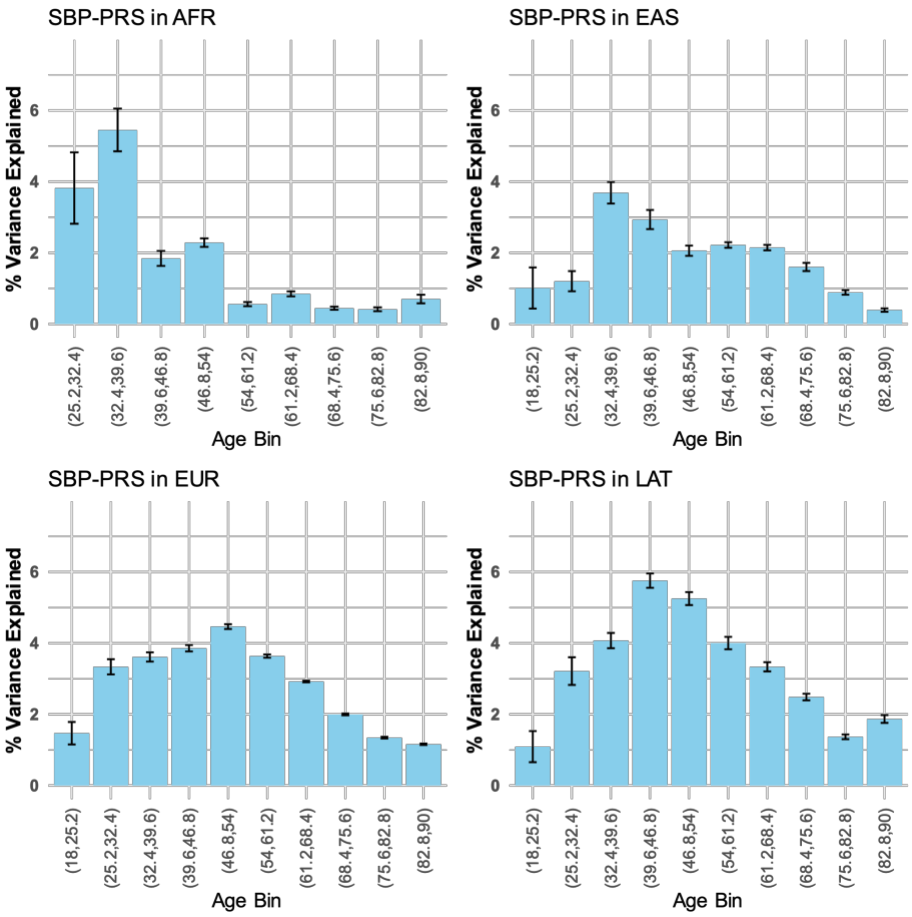

Figure 1.7: SBP PRS percent variance explained within each of the age bins, by ancestry.

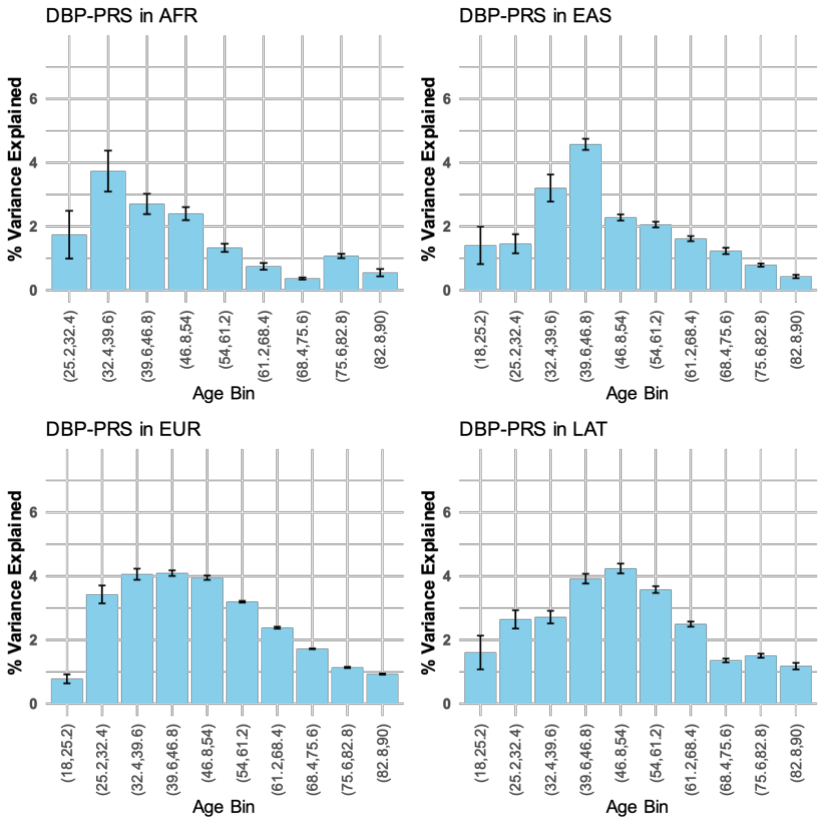

Figure 1.8: DBP PRS percent variance explained within each of the age bins, by ancestry.

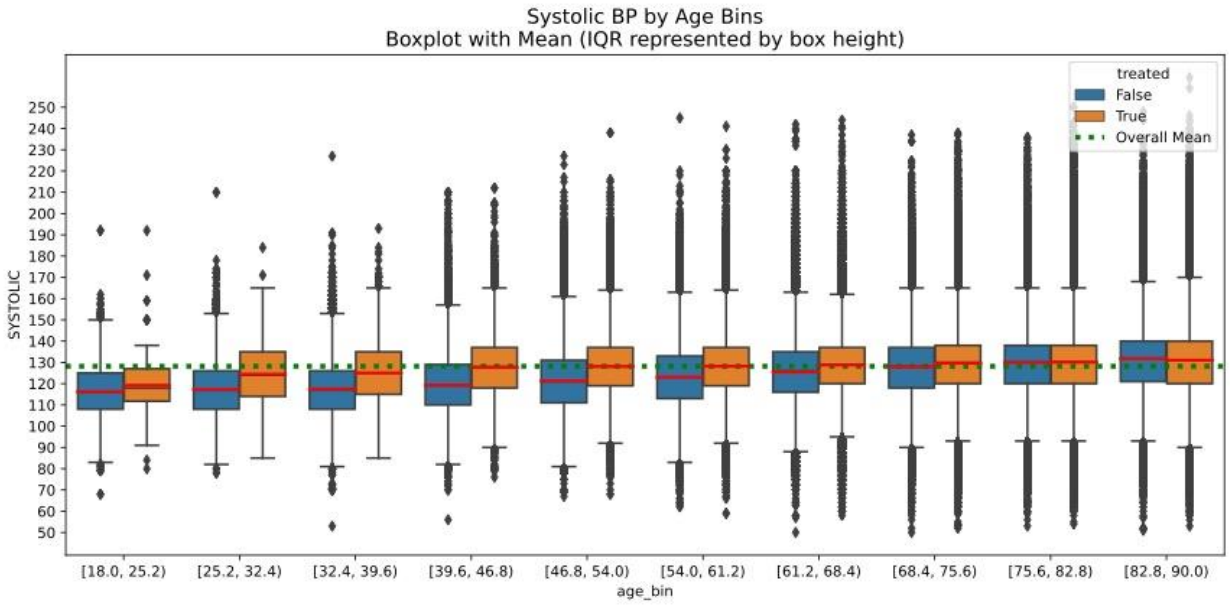

Figure 1.9: SBP across age bins in GERA cohort

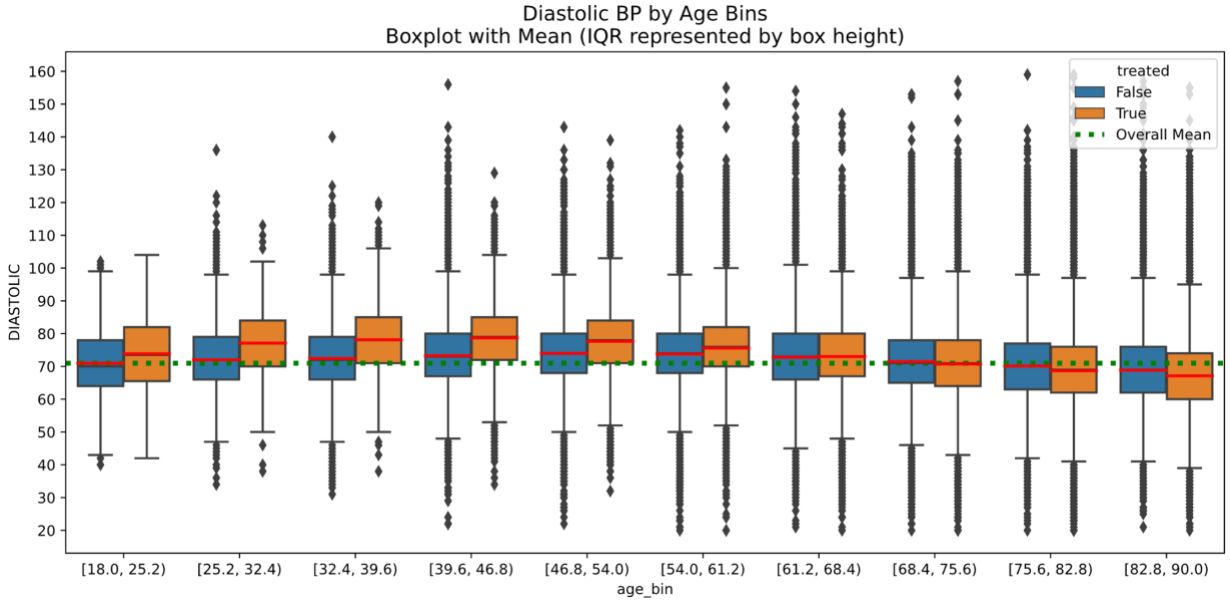

Figure 1.10: DBP across age bins in GERA cohort

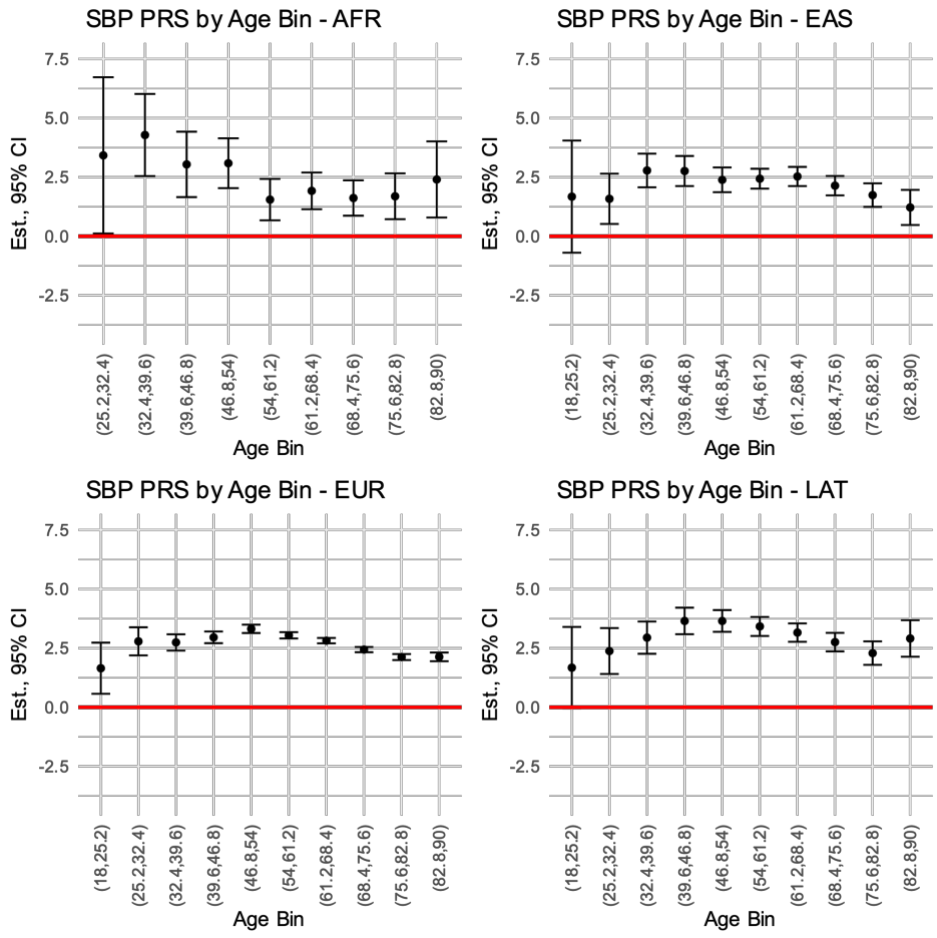

Figure 1.11: SBP PRS estimates by ancestry, across 10 age bins

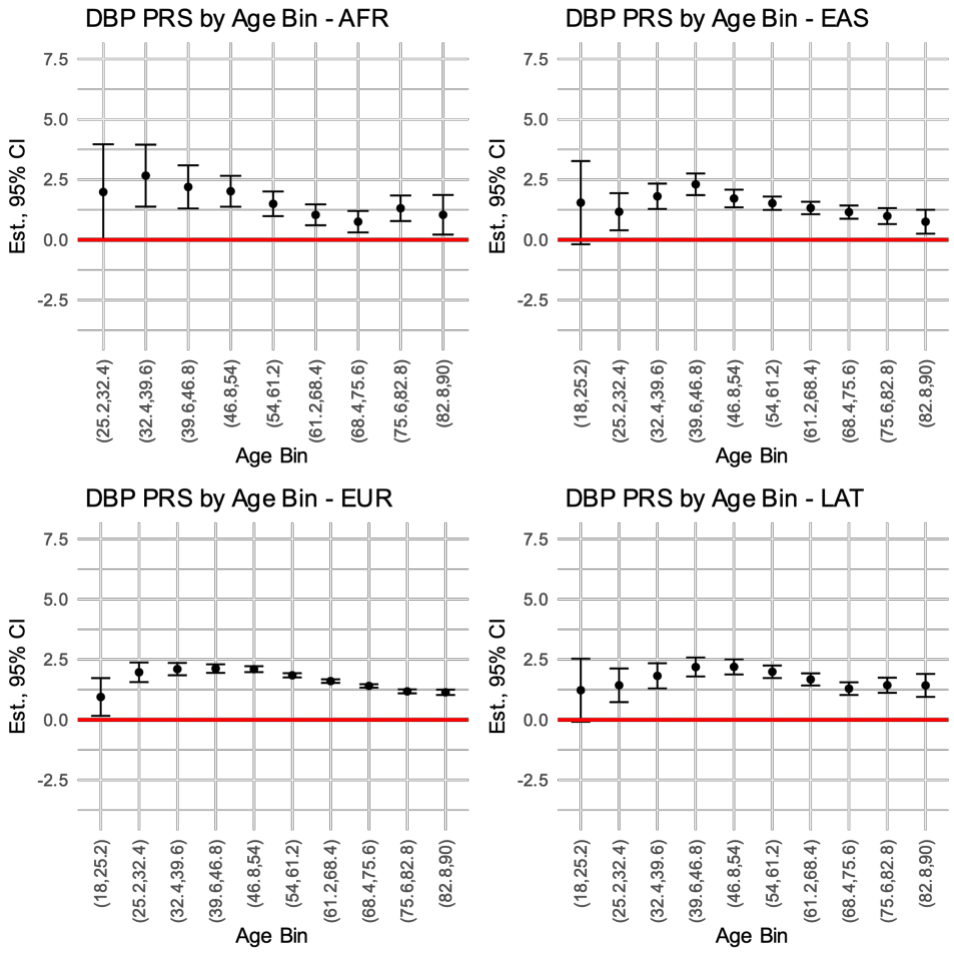

Figure 1.12: DBP PRS estimates by ancestry, across 10 age bins

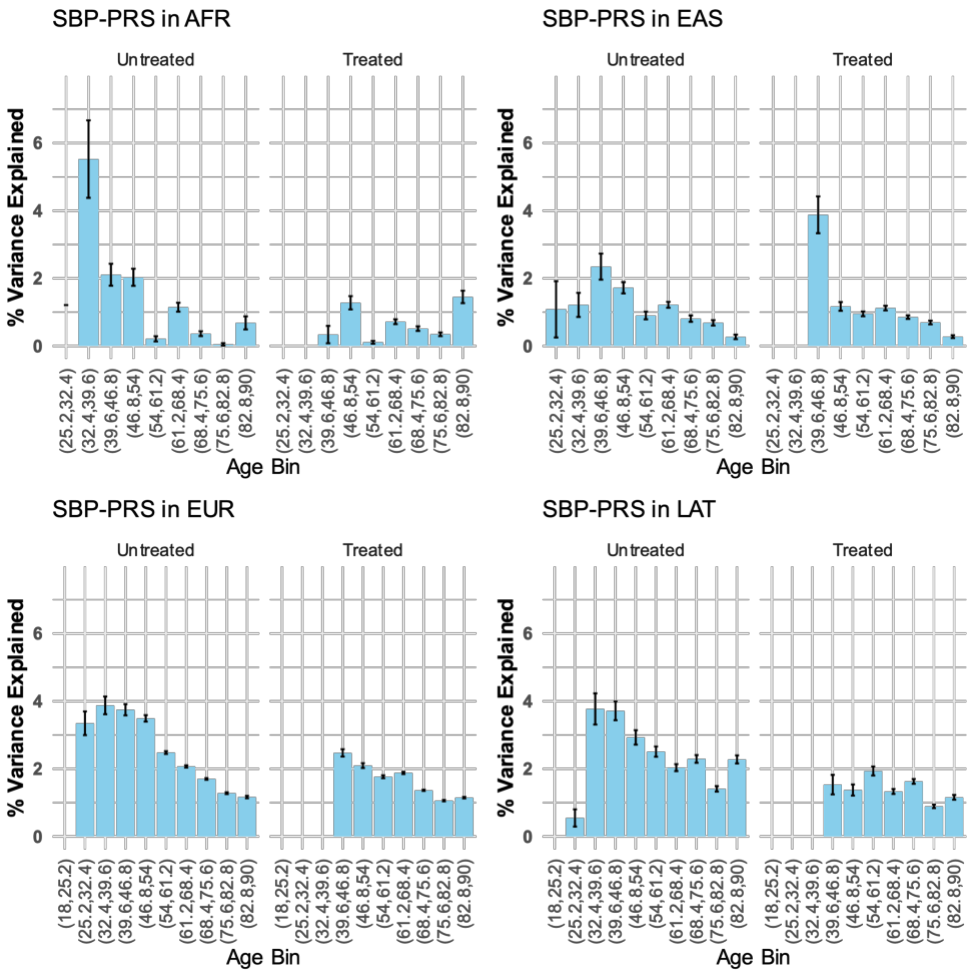

Figure 1.13: SBP PRS percent variance explained within each of the age bins, by ancestry and treatment status, among those who had both treated and untreated measures.

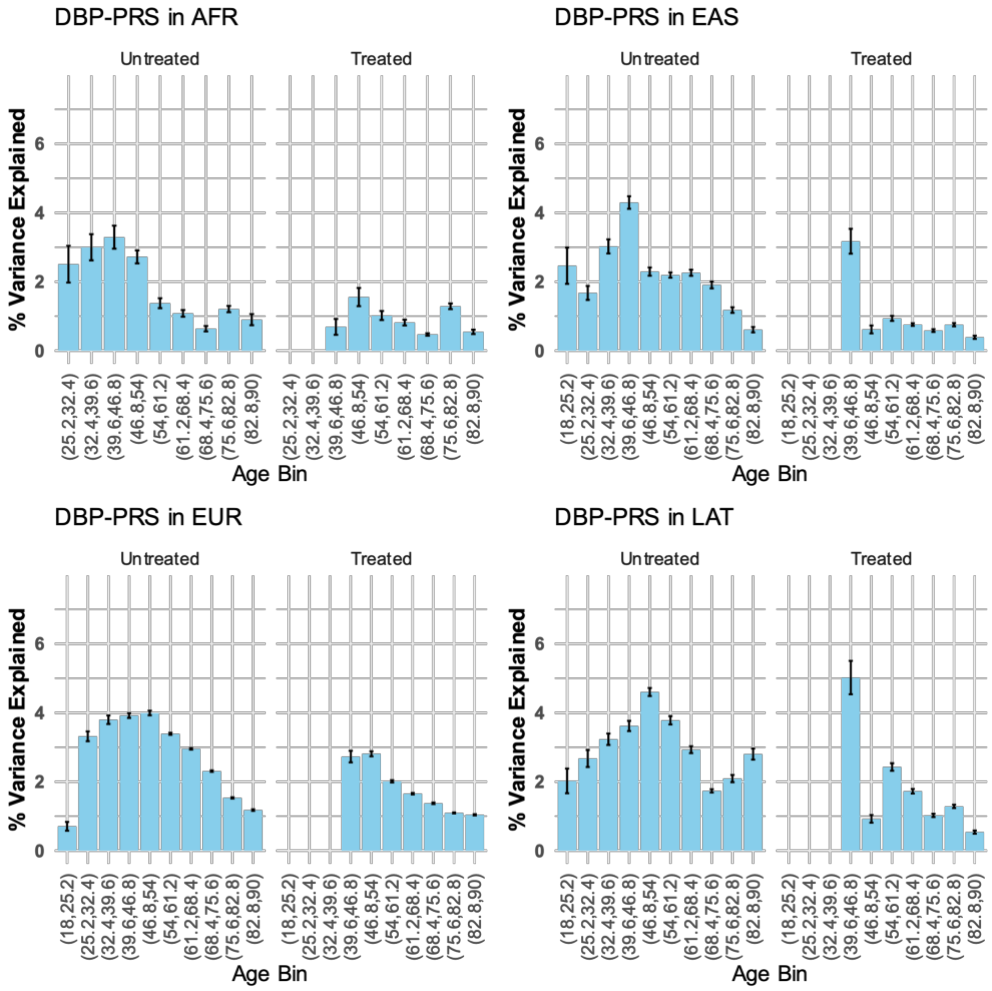

Figure 1.14: DBP PRS percent variance explained within each of the age bins, by ancestry and treatment status, among those who had both treated and untreated measures.

2. **Chapter 2: Association of multi-ancestry polygenic risk scores for blood pressure with the incidence of target organ damage in a real-world clinical cohort.**

Introduction

Negative outcomes related to high blood pressure that damage organs (target organ damage-TOD)[22] such as chronic kidney disease, heart failure, ischemic stroke, and other cardiovascular disease outcomes (CVD) are often hard to treat and manage. TOD outcomes often progressively worsen health over time and associated with poor quality of life and ultimately mortality[23–26]. Preventing and minimizing risk of TOD may involve early identification of individuals with a high risk of incidence.

Polygenic risk scores (PRS) have emerged as tool for disease risk stratification in populations in research studies[27]. PRSs, using genome-wide genetic variants, are weighted sums of risk alleles[28]. Ultimately, including PRS of predictors of TOD such as BP in risk models may aid TOD early risk identification in clinical settings, however the extent to which BP PRS is associated with TOD incidence is not known.

Prior studies have shown that BP PRS is associated with prevalence of higher hypertension-related disease risk (CAD, type 2 diabetes, atrial fibrillation, in a research cohort[1] and cardiovascular disease incidence in a cohort of Finnish participants [29]. Here, we would like to extend this and investigate the association of BP PRS with TOD incidence, in real world clinical cohorts with ancestrally diverse population, not just those of European ancestry. Ultimately, prediction for risk of incidence of TOD outcomes using BP PRS will be improved by evaluating the association of BP PRS with TOD as an independent predictor in the context of using real-world, clinical cohorts with individuals from diverse ancestral backgrounds.

The objective of this study is to investigate the association of multi-ancestry BP PRS with incident TOD in individuals stratified by ancestry groups from the ancestrally diverse real-world cohort, GERA. A key strength/ improvement in this study is the availability of individuals from Afr, Eas, Lat ancestries in addition to Eur ancestry.

Methods

GERA population

GERA dataset, with longitudinal EHR-linked clinical data, was used in this study as the target dataset. GERA has 102,859 individuals with prevalent, incident, or censored data on hypertension-related adverse outcomes, and have imputed genetic array data[30]. Individuals, genotyped in ancestry groups: afr, eas, lat, eur, sas groups[12] were included in the study.

Exposure and outcomes

The primary outcomes/ dependent variables were age at incident target organ damage outcomes. These included chronic kidney disease (CKD), Ischemic stroke (IS), heart failure (HF), atrial fibrillation (Afib), cardiomyopathy (CMP), coronary artery disease (CAD) and myocardial infarction (MI). See Appendix below for the list of International Classification of Diseases (ICD) codes used for the outcomes.

In each of the above outcomes, the earliest age of any of the ICD-9/10 codes for each of the individual was taken as the outcome date if incidence of outcome occurred for the individual. If incidence of outcome did not occur, then the individual was censored. The age at censoring was taken as the age at latest office visit BP measurement taken if individuals did not have incidence of outcome. Individuals were censored if death was recorded before incidence of outcome.
The exposure/ independent variables were multi-ancestry BP PRS (DBP and SBP).

Multi-Ancestry PRS construction

Briefly, we used GWAS summary statistics from three sources: 1) a meta-analysis of UKB, ICBP, BioVU, 2) BBJ, 3) COGENT for both SBP and DBP. The base data for the construction

of the multi-ancestry PRSs consisted of medication adjusted GWAS summary statistics (Beta coefficients and P-values) from 1) 1,028,980 European ancestry individuals in the meta-analysis study that included UKB, ICBP, BioVU, 2) 31,969 African ancestry individuals from COGENT, 3) 145,515 East Asian ancestry individuals from BBJ.

SNPs used in construction were intersected between base and test sets, as well as SNPs available in the UKBB LD reference panel used to control for SNP-SNP correlations. We used European, East Asian and African specific UKBB LD reference panel to match the ancestry specific BP GWAS summary statistics in the base data described above. The final number of SNPs used was 980,722 across all ancestry groups in the base and test data (Figure 1.1).

The PRSs were developed using the PRS-CSx[14] tool, a Bayesian regression method for PRS generation which considers MAF and LD information in its priors. PRS CSx is designed for cross population prediction and performs best compared to other contemporary PRS generating tools using other algorithms such as clumping and thresholding which are limited in settings where genetic architecture of populations in base datasets notably differ (such as African and European populations)[14].

The workflow for PRS generation using PRS-CSx is shown in figure 1.3 (Figure 1.3). PRS-CSx is a Bayesian approach to computing PRS that derives each genetic variant's posterior mean effect size (weight) from the prior GWAS base data while accounting for the LD using the 1000 Genomes or UKBB as the reference population. We used the UKBB LD panels since the reference panel used data from more samples than 1000 Genomes. We used the global shrinkage

parameter 'phi' values set to auto and generated combined SNP effect sizes across populations using an inverse-variance-weighted meta-analysis of the population-specific posterior effect size estimates using the flag '-META'. After computing posterior effect sizes, we used Plink 1.9[15] for allelic scoring, to compute the PRS for each of the test set's ancestry analysis group Afr, Eas, Eur, Lat.

To facilitate comparison of PRS effect sizes across all ancestry/ analysis groups, we first scored and standardized the BP PRS within each ancestry group (Figure 2.1).

Target organ damage PRS were downloaded from the PGS catalogue[31], these included, CKD[32], HF[33], Ischemic stroke[34], Afib[35], CMP[36], CAD[37] and MI[38]. These PRSs were scored and standardized in the GERA dataset by ancestry group (figure 2.2).

Inclusion / Exclusion criteria

Individuals with prevalent outcomes at age of entry into the study/EHR were excluded from the dataset. Sas ancestry group was dropped from the final analyses due to low numbers which resulted in big standard errors and wide confidence intervals for estimates and hazard ratios respectively.

Statistical Analyses

We tested the hazard of incidence of the outcomes with increase in 1 standard deviation of SBP or DBP PRS using Cox hazards regression models. The primary outcomes tested were age at incident CKD, IS, HF, Afib, CMP, CAD and MI. We used age as the timescale in the cox models with baseline hazard being a function of age. The ancestry group-stratified Cox models were

adjusted for covariates including sex, the first 10 or 6 genetic principal components[9]. This was the base model (adjusted for covariates + BP PRS). We also added respective PRS of the target organ damage phenotype obtained from PGS catalogue (model adjusted for covariates + BP PRS + TOD PRS) and diabetes diagnosis status at follow-up start (model adjusted for covariates + BP PRS + TOD PRS + diabetes), in additional analyses.

The association tests were two-sided. The alpha threshold of significance of the HR_{adj} was taken to be $P< 0.05$, then corrected for multiple testing using Bonferroni correction. We corrected for multiple testing across 4 ancestries within each outcome/analysis, and the additional 2 models accounting for TOD PRS and diabetes in primary outcomes. The final level of statistical significance was $P<0.0017$ for the outcomes. The analyses were conducted in R.

Results

At baseline, which was the age at entry into KP EHR records in the GERA cohort, 95,199 were free of prevalent IS, 93,759 were free of prevalent CKD, 93,345 were free of prevalent HF, 91,821 were free of prevalent Afib, 87,978 were free of prevalent CAD, 95,192 were free of prevalent CMP, 95,625 were free of prevalent MI in GERA. Baseline characteristics stratified by ancestry group in Kaiser Permanente's GERA are shown in table 1, for each of the target organ damage endpoints.

PRS Association with incidence of target organ damage across ancestry groups

Generally, the estimates and adjusted HRs (HR_{adj}) from SBP PRS (Figure 2.3) were slightly higher compared to estimates and HR_{adj} for DBP (Figure 2.4). Here, we report the SBP PRS HR_{adj}.

From survey date, there were 23,440 incident CKD events which varied between ancestry groups (highest in Afr 30.3%, and lowest in Eas 19.9% Afr 15.1%, and lowest in Eas 10.1%). The adjusted hazard ratio for incident CKD (HR_{adj}) at any age for an increase in 1 SD in SBP PRS was 1.21, 1.11, 1.13, 1.22 for Afr, Eas, Eur, Lat (Figure 2.3). This corresponds to 21%, 11%, 13 % and 22% increase in risk of incidence of CKD at any age for Afr, Eas, Eur, and Lat respectively, and 32.61, 11.05, 16.98, 22.72 extra cases per 10,000 person years for Afr, Eas, Eur, Lat respectively (Figure 2.5). In the models adjusted for covariates + BP PRS + CKD PRS, there was a slight attenuation of the HR_{adj} of BP PRS across groups, although statistical significance remained. Further adjustment for diabetes status in the model, in addition to covariates + BP PRS + CKD PRS generally attenuates the association for incident CKD (HR_{adj}) across all groups, more than addition of CKD PRS, although the HR_{adj} for BP-PRS still remain

associated with incidence of CKD (Figure 2.3). We observed variable positive correlation via calculation of Pearson correlation coefficients between SBP PRS and the CKD-PRS by ancestry groups; 0.61, 0.23, 0.12, -0.07 for Afr, Lat, Eas, Eur respectively (Table 2.2). Despite the high correlation in Afr, the association between SBP-PRS and CKD incidence remained robust. From survey date, there were 8,339 incident IS events which varied between ancestry groups (highest in Afr 10.2%, and lowest in Eas 5.7%). The adjusted hazard ratio for incident ischemic stroke (HR_{adj}) at any age for an increase in 1 SD in SBP PRS was 1.21, 1.11, 1.13, 1.22 for Afr, Eas, Eur, Lat (Figure 2.3). This corresponds to 21%, 11%, 13 % and 22% increase in risk of incidence of CKD at any age for Afr, Eas, Eur, and Lat respectively, and 32.61, 11.05, 16.98, 22.72 extra cases per 10,000 person years for Afr, Eas, Eur, Lat respectively (Figure 2.5). In the models adjusted for covariates + BP PRS + IS PRS, there was a significant attenuation of the HR_{adj} of BP PRS across groups, and the HRs were no longer statistically significant. Further adjustment for diabetes status in the model, in addition to covariates + BP PRS + CKD PRS also attenuates the association for incident CKD (HR_{adj}) across all groups, however not as much as than addition of IS PRS (Figure 2.3). We observed a high positive correlation via calculation of Pearson correlation coefficients between SBP PRS and the IS-PRS across all ancestries (Table 2.2).

We also tested associations between BP PRS and cardiovascular disease (CVD) related TOD outcomes other than ischemic stroke. These outcomes were Atrial fibrillation (Afib), Cardiomyopathy (CMP), coronary artery disease (CAD), heart failure (HF), myocardial infarction (MI). Similar to Ischemic stroke, increase in 1 SD SBP PRS was associated with incident CVD TOD outcomes, although the associations were generally weaker, and did not

reach statistical significance across all ancestry groups (Figure 2.3). Eur ancestry was the group where we observed consistent statistically significantly increased HR_{adj} in the CVD-TOD ranging from 1.14 CAD HR_{adj} to 1.1 Afib HR_{adj}. For Lat ancestry, the HR_{adj} was higher than in Eur, although the confidence intervals were wider, thus weaker statistical significance. The CVD-TOD ranged from 1.25 CMP HR_{adj} to 1.11 CAD HR_{adj}, the latter of which not statistically significant. Although there was a positive association between SBP-PRS and incidence of CVD-TOD, none of the CVD-TOD HR_{adj} were statistically significant including ischemic stroke in Afr ancestry group (1.16 HR_{adj} for CAD to 1.01 HR_{adj} for MI). This was similar to Eas group where the HR_{adj} ranged from 1.16 HR_{adj} for HF to 1.08 HR_{adj} for CMP (Figure 2.3). Addition of respective TOD PRS and diabetes status in the models did not significantly attenuate the SBP PRS HR_{adj}.

Discussion

In this study, we performed association analyses between multi-ancestry blood pressure polygenic risk scores (BP PRS), constructed using ~ 1 million variants from European, East Asian, and African BP GWAS summary statistics, and target organ damage (TOD) outcomes. We factored in sex and genetic components, in the GERA cohort, which comprises ~100,000 individuals from African, East Asian, European, and Latino(a) ancestries.

The multi-ancestry BP PRS predicted the hazard of incident TOD outcomes at any age, while controlling for sex, and ancestry stratification. We noted significant association between SBP PRS and CKD incidence, with and without controlling for diabetes and CKD PRS, except in Eas. Notably, systolic BP PRS (SBP PRS) had marginally higher estimates and hazard ratios compared to diastolic BP PRS (DBP PRS) across ancestries, although this was less evident in the East Asian ancestry group. Even though the PRS was constructed from majority Eur ancestry summary statistics, the highest associations seen were in Lat and Afr ancestries respectively.

However, the association with IS, HF, atrial fibrillation (Afib), cardiomyopathy, myocardial infarction (MI), and coronary artery disease (CAD) was not as strong, statistically significant or consistent across ancestries. This may be due to other factors, such as LDL cholesterol, which could explain more variation in CVD incidence compared to genetic determinants of blood pressure[39].

Interestingly, even though the association of BP PRS and TOD CVD outcomes was weakest in the Afr strata, the association with CKD outcomes was significant, where Afr ancestry had the second highest hazard ratio for incidence, after Lat, which was statistically significant.

As noted by Parcha et al. (2022)[40], the utility of SBP PRS in CVD risk and incidence prediction may be limited, where the American College of Cardiology/American Heart Association (ACC/AHA) Pooled Cohort Equations (PCE), which are a guideline-recommended risk estimator for atherosclerotic cardiovascular disease (ASCVD) provides similar risk quantification. Vaura et al (2021)[29] found that SBP PRS independently predicted CVD incidence but failed to improve the American College of Cardiology/American Heart Association clinical risk prediction model in a cohort of Finnish Europeans. Our findings extend this observation beyond European populations to African, East Asian, and Latino(a) ancestries, in a clinical cohort. In an analysis of association of BP PRS and prevalent CKD in All of Us, Kurniansyah et al 2023[1] found that both SBP and DBP PRS and their weighted sum were only weakly associated with CKD and only among non-antihypertension medication users, while a hypertension PRS was not. Furthermore, the study did not investigate association between PRS and incidence of CKD, only prevalence.

One of the major limitations of our study is the heterogeneity in sample sizes across ancestry populations, with most of the PRS being constructed from European samples. Additionally, there is a lack of large multi-ancestry GWAS for BP, which limits the applicability of BP PRS in certain populations.

Despite the limitations, we have shown that there are significant associations between BP PRS and TOD, particularly CKD incidence, independent of TOD PRS and diabetes. This could potentially be useful in risk stratification, particularly in risk stratification for incidence of CKD, which is strongly linked to high BP. BP PRS could help identify individuals at higher risk of CKD TOD at any age. Those with the highest genetic risk for high BP had the highest hazard for particularly CKD, except for Eas, suggesting that BP PRS might have clinical utility in specific contexts.

Further research using larger sample sizes and the inclusion of higher proportion non-European ancestry samples in GWAS would improve the accuracy and generalizability of BP PRS in clinical settings.

Conclusion

This study provides evidence of association between BP PRS and TOD incidence outcomes specifically ischemic stroke, heart failure, chronic kidney disease, independent of TOD and diabetes in a clinical cohort. The association was strongest in CKD incidence, particularly in African, European, and Latino(a) but not East Asian ancestry populations. However, the association between BP-PRS and CVD related TOD outcomes including IS, HF, Afib, CAD, MI, Cardiomyopathy was weaker than CKD, and non-consistent across ancestries.

Tables

Table 2.1: Characteristics of GERA cohort at time of start of follow, and primary outcomes

	Ancestry group			
	AFR	**EAS**	**EUR**	**LAT**
Atrial fibrillation (Afib)				
n	2928	6882	74231	7780
female = 1 (%)	1745 (59.6)	3996 (58.1)	43628 (58.8)	4692 (60.3)
Age at entry in years (mean (SD))	49.48 (12.60)	47.58 (13.58)	52.74 (12.26)	47.13 (13.78)
Age at outcome or censoring in years (mean (SD))	70.23 (12.71)	68.24 (14.03)	72.85 (12.07)	67.65 (14.10)
Afib = 1 (%)	312 (10.7)	607 (8.8)	12754 (17.2)	798 (10.3)
Diabetes = 1 (%)	144 (4.9)	135 (2.0)	1410 (1.9)	230 (3.0)
Coronary Artery Disease (CAD)				
n	2763	6622	71183	7410
female = 1 (%)	1661 (60.1)	3927 (59.3)	43010 (60.5)	4585 (61.9)
Age at entry in years (mean (SD))	48.92 (12.60)	47.05 (13.44)	52.34 (12.23)	46.51 (13.69)
Age at outcome or censoring in years (mean (SD))	69.44 (12.77)	67.66 (14.00)	72.60 (12.20)	66.97 (14.07)
CAD = 1 (%)	357 (12.9)	531 (8.0)	8398 (11.8)	677 (9.1)
Diabetes = 1 (%)	115 (4.2)	118 (1.8)	1068 (1.5)	174 (2.3)
Chronic Kidney Disease (CKD)				
n	2820	6861	76286	7792
female = 1 (%)	1677 (59.5)	3986 (58.1)	44416 (58.3)	4675 (60.0)
Age at entry in years (mean (SD))	49.19 (12.52)	47.49 (13.50)	53.04 (12.31)	47.15 (13.78)
Age at outcome or censoring in years (mean (SD))	68.59 (12.19)	67.47 (13.61)	72.48 (11.84)	66.94 (13.68)
CKD = 1 (%)	855 (30.3)	1368 (19.9)	19625 (25.7)	1592 (20.4)
Diabetes = 1 (%)	123 (4.4)	127 (1.9)	1344 (1.8)	207 (2.7)
Cardiomyopathy (CMP)				
n	2923	6958	77421	7890

	Ancestry group			
	AFR	**EAS**	**EUR**	**LAT**
female = 1 (%)	1739 (59.5)	4017 (57.8)	45001 (58.2)	4734 (60.0)
Age at entry in years (mean (SD))	49.54 (12.65)	47.71 (13.61)	53.23 (12.36)	47.36 (13.87)
Age at outcome or censoring in years (mean (SD))	70.45 (12.87)	68.63 (14.18)	73.75 (12.23)	68.14 (14.30)
CMP = 1 (%)	115 (3.9)	120 (1.7)	2455 (3.2)	183 (2.3)
Diabetes = 1 (%)	136 (4.7)	144 (2.1)	1544 (2.0)	229 (2.9)
Heart Failure (HF)				
n	2873	6902	75796	7774
female = 1 (%)	1707 (59.4)	3993 (57.9)	44197 (58.3)	4668 (60.1)
Age at entry in years (mean (SD))	49.37 (12.59)	47.58 (13.57)	52.95 (12.28)	47.17 (13.81)
Age at outcome or censoring in years (mean (SD))	70.04 (12.73)	68.36 (14.06)	73.31 (12.13)	67.77 (14.14)
HF = 1 (%)	391 (13.6)	530 (7.7)	10489 (13.8)	784 (10.1)
Diabetes = 1 (%)	131 (4.6)	135 (2.0)	1341 (1.8)	202 (2.6)
Myocardial Infarction (MI)				
n	2974	6981	77772	7898
female = 1 (%)	1761 (59.2)	4023 (57.7)	45124 (58.1)	4741 (60.1)
Age at entry in years (mean (SD))	49.68 (12.66)	47.77 (13.63)	53.28 (12.37)	47.38 (13.87)
Age at outcome or censoring in years (mean (SD))	70.54 (12.81)	68.62 (14.15)	73.74 (12.22)	68.10 (14.28)
MI = 1 (%)	157 (5.3)	248 (3.6)	3664 (4.7)	320 (4.1)
Diabetes = 1 (%)	150 (5.0)	147 (2.1)	1586 (2.0)	231 (2.9)
Ischemic Stroke (IS)				
n	2956	6962	77402	7879
female = 1 (%)	1751 (59.2)	4001 (57.5)	44869 (58.0)	4713 (59.8)
Age at entry in years (mean (SD))	49.61 (12.65)	47.70 (13.59)	53.20 (12.34)	47.34 (13.85)
Age at outcome or censoring in years (mean (SD))	70.33 (12.76)	68.52 (14.09)	73.57 (12.16)	67.99 (14.23)
stroke = 1 (%)	302	396	7055	586

	Ancestry group			
	AFR	**EAS**	**EUR**	**LAT**
	(10.2)	(5.7)	(9.1)	(7.4)
Diabetes = 1 (%)	149	149	1547	228
	(5.0)	(2.1)	(2.0)	(2.9)

Table 2.2: Correlation between SBP PRS and TOD PRS in GERA cohort, estimated using Pearson correlation coefficients.

	Pearson correlation coefficient and (P value)			
	AFR	**EAS**	**LAT**	**EUR**
TOD Outcome PRS				
Chronic Kidney Disease	0.61 (2.29E-318)	0.12 (8.65E-26)	0.23 (2.42E-96)	-0.07 (9.52E-97)
Atrial Fibrillation	0.48 (2.22E-180)	0.02 (0.069)	0.21 (3.86E-81)	0.02 (8.10E-08)
Coronary artery disease	-0.26 (6.64E-47)	0.10 (7.10E-17)	-0.02 (0.073)	0.19 (0)
Cardiomyopathy	0.17 (4.78E-21)	0.03 (0.003)	0.10 (1.00E-18)	0.02 (4.73E-08)
Heart Failure	0.02 (0.38)	0.05 (3.35E-05)	0.05 (1.01E-06)	0.11 (1.42E-218)
Ischemic stroke	0.65 (0)	0.32 (7.80E-166)	0.46 (0)	0.49 (0)

Figures

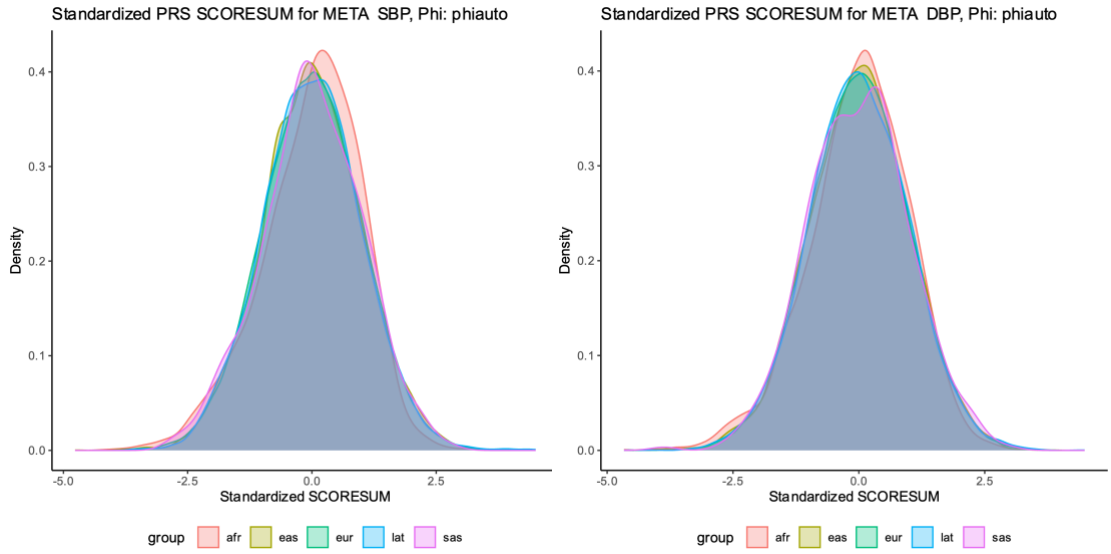

Figure 2.1: SBP and DBP PRS standardized within ancestry groups in GERA cohort

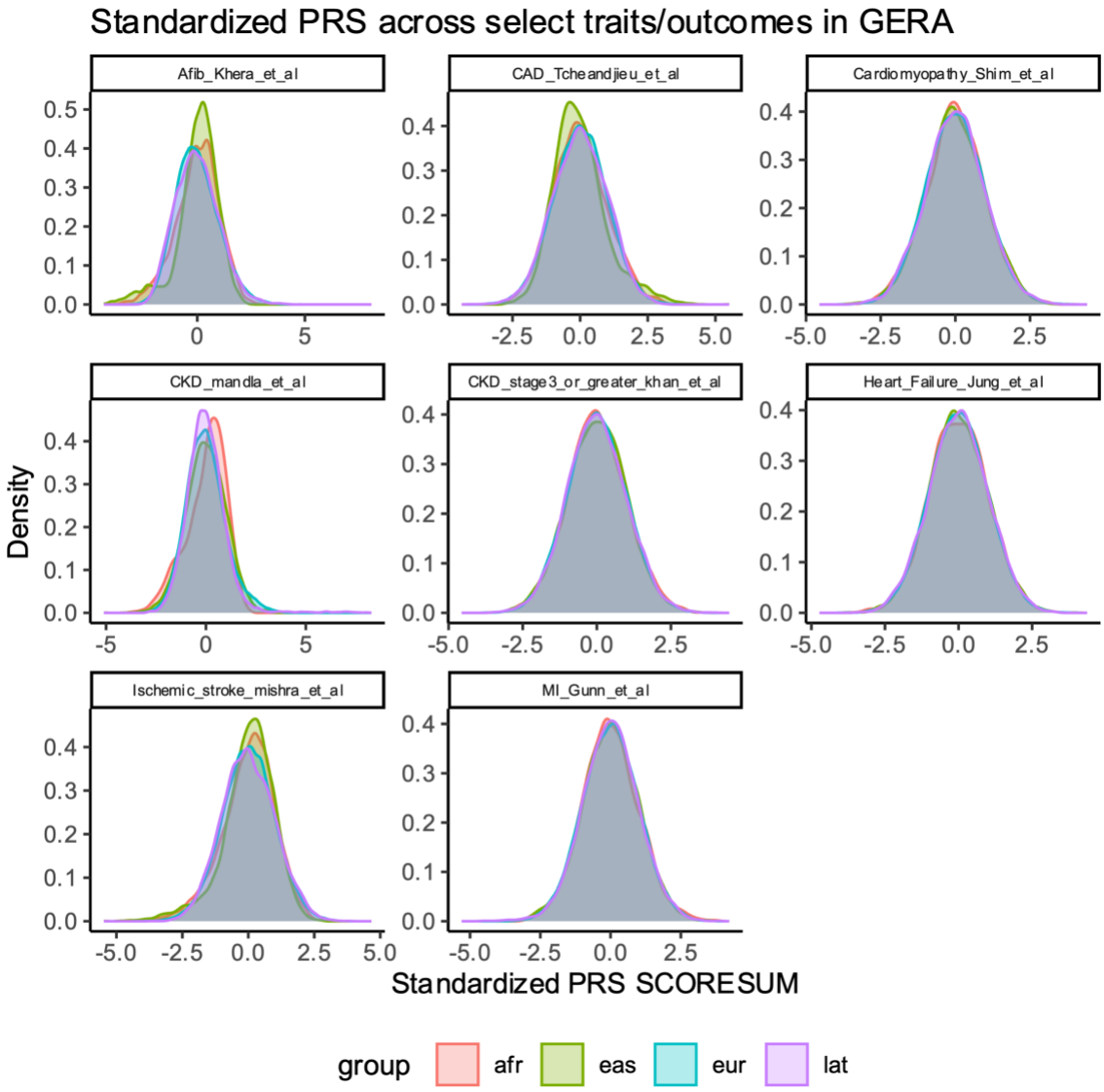

Figure 2.2: Standardized scored PRS of select target organ damage and BP phenotypes in GERA. The PRSs were downloaded from the PGS catalogue, an online database with published PRS

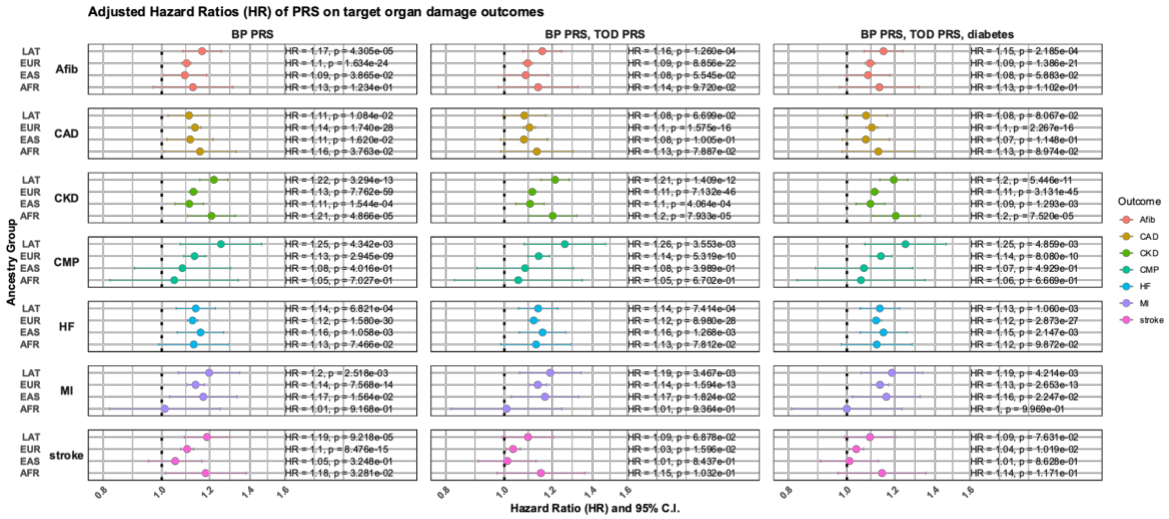

Figure 2.3: HR_{adj} for incident TOD outcomes with 1 SD increase in SBP PRS stratified by ancestry

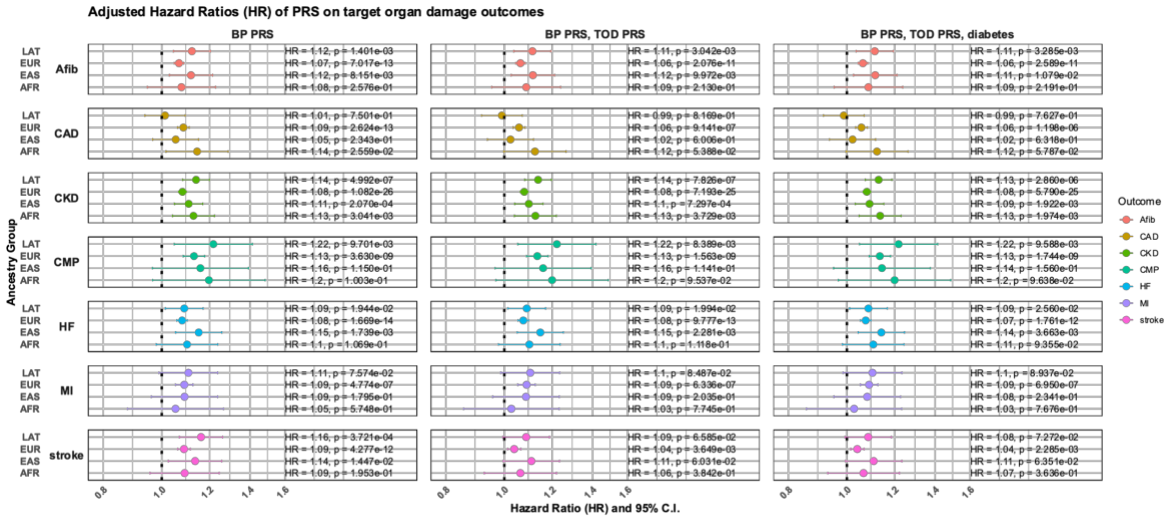

Figure 2.4: HR_{adj} for incident TOD outcomes with 1 SD increase in DBP PRS stratified by ancestry

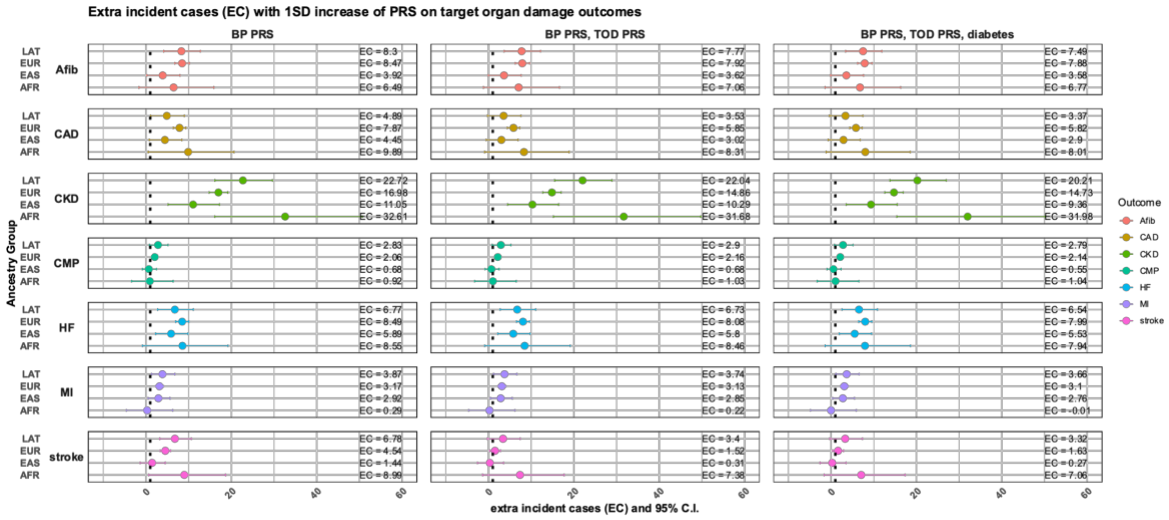

Figure 2.5: Extra incident cases per 10,000 person years with 1 SD increase of SBP PRS by ancestry groups in GERA cohort

3. **Chapter 3: Estimated changes in blood pressure from monotherapy antihypertension medication treatment in a real-world cohort.**

Background

The use of antihypertensive (AHT) medications is a common and essential approach to managing high blood pressure, a key risk factor for cardiovascular disease and kidney disease worldwide[41]. While clinical trials have shown that different classes of AHT drugs can reduce blood pressure, responses to these medications often vary based on factors like ancestry and age[42]. However, most of the data comes from controlled trial settings, which may not be generalizable to real-world clinical populations.

The observed effect of AHT medication on blood pressure in real-world clinical settings, particularly among different ancestral population groups, is still underexplored. Characterization and understanding of AHT treatment effects in real world clinical cohorts is important for improving approaches to hypertension treatment. Additionally, genome-wide association studies (GWAS), the effect of AHT treatment is often assumed to be uniform—commonly modeled as a 15-mmHg reduction in systolic blood pressure (SBP) and a 10-mmHg reduction in diastolic blood pressure (DBP) across all groups.[43] Such assumptions overlook important variations in drug response related to ancestry or specific drug classes and doses.

This study aims to fill these gaps by investigating BP response to first-line AHT medications in a real-world clinical cohort, using longitudinal electronic health record (EHR) data. In addition, an objective of this paper is to investigate the response differences by ancestry across individuals of African, East Asian, European, and Latino(a) ancestry with hypertension. Furthermore, we sought to investigate dose effect of the BP response to monotherapy treatment status, by drug

classes. Findings could help improve hypertension AHT treatment strategies and improve strategies for controlling AHT use in the GWAS studies using real-world clinical cohort data.

Methods

Dataset

We used data from the GERA cohort, which includes longitudinal electronic health records (EHR) along with multiple measurements of systolic (SBP) and diastolic blood pressure (DBP) per individual across time. The GERA dataset links pharmacy prescription data to individual prescriptions, including the start and end dates of antihypertensive (AHT) drug use. This study included individuals of African, East Asian, European, and Latino(a) ancestry, while South Asian individuals were excluded due to insufficient sample size.

Pharmacy prescription records were merged with BP measurement data from the internal medicine department using unique subject IDs. Blood pressure measurements were classified as "treated" if they occurred between the start and end dates of an AHT prescription. Measurements outside these periods were classified as "untreated." We restricted the analysis to cases where only one AHT drug (monotherapy) was prescribed during the treatment period. The daily dose (DD) was calculated as the product of the dose in milligrams and the strength of the prescribed drug. To be eligible for the study, individuals had to have both treated and untreated BP measurements, in addition to a confirmed diagnosis of hypertension (Figure 3.1).

Main outcomes and measures

The dependent variable/ outcome of interest was the delta between the average BP measurement while on monotherapy treatment vs average BP measurement while not on treatment at the individual level. We examined both SBP and DBP measurements, in mm HG, taken at the Internal Medicine department office visits. For monotherapy treatment, the AHT drug classes considered in this study included ACE inhibitors[44] (Lisinopril, Enalapril, Captopril, Benazepril,

Fosinopril, Quinapril, Ramipril), hydrochlorothiazide diuretics[45], and dihydropyridine calcium channel blockers[46] (Amlodipine, Felodipine, and Nifedipine). These calcium channel blockers primarily act as peripheral vasodilators and have minimal direct effects on the myocardium at therapeutic doses.

We stratified treatment into daily dose categories of daily doses (see Appendix for dose equivalency). We assumed adherence to the dose once prescribed.

Statistical Analysis

To examine ancestry-based differences in monotherapy response, we performed Analysis of Variance (ANOVA) followed by Tukey's post-hoc tests to test whether the average changes in BP between ancestry groups were statistically significant. Analyses were stratified by antihypertensive drug classes: ACE inhibitors, hydrochlorothiazide (HCTZ), and dihydropyridine calcium channel blockers (CCBs). This approach allowed for the identification of significant differences in average drug response among ancestral groups for each antihypertensive class.

We evaluated the association between monotherapy AHT drug treatment response with important covariates including dose (DD), sex at birth using linear models. The association between response and dose was modelled as log linear, meaning a unit increase in daily dose (DD) leads to a multiplicative change in response by a factor of e^ β (coefficient of dose in model). A 100% increase or doubling in dose is associated with a β * log (2) (or 0.693β) unit change in the BP response (dependent variable).

Other covariates in the model included average age at BP measurement during treatment, sex, genetic principal components (first 10 for Eur/ 6 for Afr, Eas, Lat), diabetes, heart failure and chronic kidney disease diagnosis status at time of treatment and BP measurement. The formula for linear models for any drug class, and ancestry group strata is specified as:

$[_{mean}\text{Treated\_BP}] - [_{mean}\text{Untreated\_BP}] = \beta_0 + \beta_1 \times \log(\text{Dose}+1) + \beta_2 \times (PC_1) + \beta_3 \times (PC_2) + \ldots + \beta_p X_p$

Where: X_p: covariates (female sex, body mass index (BMI), chronic kidney disease diagnosis status, heart failure diagnosis status, diabetes diagnosis status, the average age difference (at treated vs untreated) and β_p: coefficients for these covariates.

In the statistical analyses above, the threshold for statistical significance was determined by an original alpha level set at $P< 0.05$, followed by Bonferroni correction due to multiple testing across SBP and DBP, as well as 3 drug classes for the Anova test, and additional 4 ancestry strata per drug class for the linear models. For the Anova test and for the Tukey's Post Hoc test, the statistical significance threshold was $P< 0.0083$. For the linear models, the statistical significance threshold for covariates was $P<0.0021$ (0.05 (original alpha level) divided by 24 tests to account for multiple hypothesis testing explained above).

Results

Characteristics of GERA cohort (those diagnosed with essential hypertension)

A total of 11, 901 participants with diagnosed hypertension met the inclusion criteria for study entry in the ACE-inhibitor analysis, 8,001 in the hydrochlorothiazide analysis, and 7,616 in the CCBs analysis (Table 3.1). A total of 190,366 BP measures were included in the ACE-inhibitor analysis. A total of 124,351 BP measures were included in the HCTZ analysis. A total of 110,392 BP measures were included in the CCB analysis. The % treated differed by ancestry group, and drug class type (Table 1). Majority of individuals included were female, with low of 50.9% among Eas ancestry ACE-inhibitors analysis, and high of 75.4% among Afr ancestry in the Hydrochlorothiazide analysis (Table 3.1). The average number of treated measures per individual ranged from 3.87 in Eur in CCB analysis to 6.09 in Lat in the ACE-inhibitor analysis. The average number of untreated measures per individual ranged from 9.76 in Eas in the ACE-inhibitor analysis to 11.84 in Lat in the CCB analysis (table 3.2). Generally, average treated BP (SBP and DBP) was lower than average untreated SBP and DBP across the ancestry groups (table 3.2).

Since different drugs within each of the ACE-inhibitor drug class have different recommended effective doses, we generated a daily dose (DD) value for each type such that 1.0 DD was equal to 10 mg of Lisinopril daily. We also generated equivalencies for CCBs, 1.0 DD was equal to 5 mg of Amlodipine daily (See Appendix).

Average SBP and DBP changes due to monotherapy treatment by ancestry

In the ACE-inhibitor analysis, the average change in SBP for treated compared to untreated was -0.74 mm Hg [SD 14.52] for African, -3.53 mm Hg [SD 13.43] for East Asian, -2.6 mm Hg [SD 13.65] for European, and -3.15 mm Hg [SD 13.35] for Latino(a) groups (Figure 3.2). The Anova F-value was 4.04, which was statistically significant with P-value 0.007, indicating statistically significant differences in average of at least one group. The average reduction in SBP was lower in the African ancestry group compared to the others, although only statistically significant with East Asian group (Table 3.3). For DBP the estimated change was 1.07 mm Hg [SD 9.25] for African, 0.21 mm Hg [SD 8.19] for East Asian, 0.26 mm Hg [SD 8.36] for European, and 0.51 mm Hg [SD 8.2] for Latino(a) groups (Figure 3.3). The average change in DBP was not statistically significantly different between any of the ancestry groups (Table 3.4). In all ancestry groups, there was no statistically significant estimated effect of ACE-inhibitor dose doubling (100% increase in dose) in both SBP (Table 3.5) and DBP (Table 3.6).

In the hydrochlorothiazide (HCTZ) analysis, the average change in SBP for treated compared to untreated was -3.97 mm Hg [SD 13.23] for African, -3.17 mm Hg [SD 14.18] for East Asian, -1.72 mm Hg [SD 13.58] for European, and -0.95 mm Hg [SD 13.58] for Latino(a) groups (Figure 3.4). The Anova F-value was 6.12, which was statistically significant with P-value 4e-04, indicating statistically significant differences in average of at least one group. The average reduction in SBP was higher in the African ancestry group compared to the other ancestry groups, although only statistically significant with European and Latino(a) groups (Table 3.3).

For DBP the estimated change was 0.43 mm Hg [SD 8.32] for African, 0.86 mm Hg [SD 8.64] for East Asian, 1.14 mm Hg [SD 8.33] for European, and 1.63 mm Hg [SD 8.64] for Latino(a) groups (Figure 3.5). The average change in DBP was not statistically significantly different between any of the ancestry groups (Table 3.4).

In European and Latino ancestry groups, there was statistically significant estimated change in SBP reduction if HCTZ dose was doubled. A 100% increase in HCTZ dose was shown to lead to a change of -1.99 mmHG (95% C.I.: -0.78 mmHG- -3.2 mmHG) in European ancestry group and -6.08 mmHG (95% C.I.: -2.51 mmHG- -9.66 mmHG) in Latino ancestry group. (Table 3.5). However, there was no observed statistically significant estimated effect of HCTZ change with dose doubling (100% increase in dose) on DBP (Table 3.6).

For dihydropyridine calcium channel blockers (CCBs) analysis, the average change in SBP for treated compared to untreated was -3.03 mm Hg [SD 13.23] for African, -3.07 mm Hg [SD 13.08] for East Asian, -1.99 mm Hg [SD 14.06] for European, and -2.28 mm Hg [SD 12.84] for Latino(a) groups (Figure 3.6). The Anova F-value was 1.36, which was not statistically significant with P-value 0.25, indicating no statistically significant differences in average of ancestry groups. The average reduction in SBP was not statistically higher in any ancestry group compared to the other ancestry groups (Table 3.3). For DBP the estimated change was 0.28 mm Hg [SD 8.68] for African, 0.66 mm Hg [SD 7.76] for East Asian, 0.21 mm Hg [SD 8.33] for European, and 0.27 mm Hg [SD 8.25] for Latino(a) groups (Figure 3.7). The average change in DBP was not statistically significantly different between any of the ancestry groups (Table 3.4).

In European ancestry group, there was a statistically significant estimated change in SBP reduction if CCB dose was doubled. A 100% increase in CCB dose was shown to lead to a change of -4.00 mmHG (95% C.I.: -3.03 mmHG- -4.97 mmHG) in European ancestry group (Table 3.5). We observed statistically significant estimated effect of CCB change with dose doubling. A 100% increase in CCB dose on DBP in European, and East Asian ancestry groups resulted in change of -2.05 mmHG (95% C.I.: -1.48 mmHG- -2.61 mmHG), -2.69 mmHG (95% C.I.: -1.10 mmHG- -4.27 mmHG) in DBP respectively (Table 3.6).

Discussion

This study examined the relationship between monotherapy antihypertensive treatment (AHT) and estimated average changes in systolic (SBP) and diastolic blood pressure (DBP) across African (Afr), East Asian (Eas), European (Eur), and Latino(a) (Lat) ancestry groups using real-world clinical data. Overall, SBP showed a greater change and overall reduction compared to DBP (where most of the outcomes were an average positive/ higher DBP value) across drug types and ancestries. Notably, the SBP reduction from treatment with hydrochlorothiazide (HCTZ) was highest in individuals of African ancestry, while the estimated SBP reduction with ACE inhibitors was significantly lower for Afr ancestry group compared to Eas, Eur, Lat while there was no difference between ancestries with dihydropyridine calcium channel blockers (CCBs).

Previous research, particularly from clinical trials, has explored blood pressure response differences particularly between individuals from Black and White race-ethnicities[47–49]. However, few studies have investigated these differences across multiple ancestry groups (Afr, Eas, Eur, Lat) or across multiple drug classes in a real-world clinical setting. Interestingly, we found lower reductions in BP in our cohort compared to clinical trials, where the average BP response to ACE inhibitors, HCTZ, and CCBs tends to be higher. Differences across drug classes and ancestries may be explained by other studies that have examined the mechanisms of action for these drugs.

Interpretation of Results and Mechanisms

ACE inhibitors reduce blood pressure by inhibiting the renin-angiotensin-aldosterone system (RAAS), thereby decreasing the production of angiotensin II, a potent vasoconstrictor[50]. They are also frequently prescribed for comorbid conditions such as diabetes, heart failure, and chronic kidney disease[50]. Individuals of African ancestry have been shown to have a lower response to ACE inhibitors compared to individuals of European ancestry, with a systematic review of clinical trials reporting no significant difference between ACE inhibitors and placebo in Black individuals[48]. Additionally, African ancestry individuals are more likely to experience side effects such as angioedema[51–53], which may contribute to lower adherence and thus lower observed efficacy. The combination of a higher prevalence of reduced plasma renin levels (common in hypertensive African Americans and Caribbean Hispanics), salt sensitivity, and older age (which affects response to renin) [54] may explain the lower response to ACE inhibitors in the African ancestry group in this study.

For CCBs, which work by blocking calcium channels and primarily act as peripheral vasodilators,[46] systematic reviews have shown that individuals of African ancestry generally respond better compared to those of European ancestry[47]. However, few studies have compared CCB effectiveness across multiple ancestry groups. Our findings suggest no notable differences across ancestries, though SBP reductions appeared to be slightly lower in the European group. HCTZ, a diuretic, reduces blood pressure by promoting salt and water excretion through the kidneys and by causing vasodilation over time[45,55,56]. Previous studies have indicated that HCTZ is more effective in African ancestry individuals[47], particularly those with salt-sensitive hypertension, a condition common in both African ancestry and older individuals. We observed

similar results, with the highest SBP reductions in African ancestry individuals compared to European and Latino(a) groups. The combination of diuretics with RAAS inhibitors has been shown to improve blood pressure control, particularly in Black individuals[57], suggesting that much of the effect in this population may be attributable to the use of diuretics like HCTZ.

One limitation of this study is the relatively small sample size, which may affect the generalizability of our findings. Another limitation is the absence of a washout period for drug use, which may affect the accuracy of our estimates.

Our study provides important insights into the variation in average blood pressure (BP) change associated with antihypertensive medications across different ancestry groups, drug classes, and doses in a real-world clinical cohort. We found that BP responses are not uniform across ancestry populations, contrary to the common practice in research studies of applying generalized assumptions, such as a 15 mm Hg reduction in systolic blood pressure (SBP)[43], regardless of ancestry or medication type. Specifically, individuals of African ancestry showed a higher estimated reduction in SBP with hydrochlorothiazide (HCTZ) but a significantly lower response to ACE inhibitors compared to other groups. The lack of a consistent dose-response relationship across some ancestry groups further emphasizes that BP medication effects are complex and context-dependent.

These findings challenge the conventional approach used in genetic and epidemiological research, where uniform BP medication adjustments are often applied without considering the diverse responses across ancestry populations. This highlights the need for more personalized

approaches in both clinical care and research, particularly in GWAS that correct for BP medication use. Tailoring medication adjustments to specific ancestry and drug type could improve the accuracy of BP-related genetic findings and lead to better treatment strategies for hypertension across diverse populations.

Further studies should explore the association between drug response and genetic ancestry proportions rather than ancestry group categories, while controlling for factors like diet and salt intake. Future research should assess how these differences affect BP correction factors in GWAS studies.

Conclusion

This study provides insights into the relationship between monotherapy antihypertensive treatment (AHT) and changes in systolic (SBP) and diastolic blood pressure (DBP) across diverse ancestry groups, including African, East Asian, European, and Latino(a) populations, using real-world data. Hydrochlorothiazide (HCTZ) showed the highest SBP reduction among individuals of African ancestry, while ACE inhibitors were less effective for this group compared to others, with no ancestry-related differences observed for calcium channel blockers (CCBs). Notably, the blood pressure reductions observed in this real-world cohort were smaller than those reported in clinical trials. The findings show that estimated BP changes from monotherapy antihypertensive medications vary by ancestry, drug class, in real-world clinical cohort settings.

Tables

Table 3.1: Characteristics of dataset by analysis group

	Ancestry group			
	AFR	EAS	EUR	LAT
ACE-Inhibitor analysis				
Number of individuals	380	758	9,677	1,086
proportion female	59.1%	50.9%	55.5%	56.2%
Total BP Measures	6,342	11,925	153,413	18,686
% treated measures	34.8%	36.8%	37.6%	35.4%
Dihydropyridine Calcium Channel Blockers analysis				
Number of individuals	455	533	6,053	575
proportion female	64.3%	64.0%	66.7%	65.3%
Total BP Measures	6,754	7,510	86,974	9,154
% treated measures	28.9%	29.7%	26.9%	24.0%
Hydrochlorothiazide (HCTZ) analysis				
Number of individuals	397	572	6,368	664
proportion female	75.4%	73.9%	73.1%	73.5%
Total BP Measures	6,561	8,450	98,193	11,147
% treated measures	35.5%	33.9%	32.5%	30.2%

Table 3.2: Extended characteristics of dataset by analysis group

	Ancestry group			
	AFR	**EAS**	**EUR**	**LAT**
ACE-Inhibitors analysis				
Mean (SD) number of treated measures per individual	5.81 (6.62)	5.79 (6.12)	5.96 (6.48)	6.09 (6.61)
Mean (SD) number of untreated measures per individual	10.88 (10.86)	9.94 (9.91)	9.90 (10.59)	11.11 (10.71)
Mean (SD) age at treated BP measure	66.38 (10.71)	67.29 (10.74)	70.42 (10.23)	65.55 (11.35)
Mean (SD) age at untreated BP measure	67.93 (11.15)	69.70 (11.46)	72.42 (10.78)	67.61 (12.11)
Mean (SD) BMI	30.94 (6.55)	25.62 (4.45)	28.16 (5.97)	29.92 (6.48)
Mean (SD) systolic BP treated	130.34 (15.12)	126.41 (14.57)	128.73 (15.60)	128.06 (15.24)
Mean (SD) diastolic BP treated	74.38 (10.33)	72.59 (9.93)	72.22 (10.08)	72.99 (9.47)
Mean (SD) systolic BP untreated	133.72 (16.23)	131.39 (16.97)	132.98 (16.78)	132.53 (16.73)
Mean (SD) diastolic BP untreated	75.30 (11.15)	72.93 (10.96)	72.84 (11.08)	73.20 (11.17)
% with Heart Failure	15.3	5.8	11.3	6.4
% with diabetes	54.8	56.0	40.8	59.1
% with CKD	42.4	40.3	42.1	36.5
Dihydropyridine Calcium Channel Blockers analysis				
Mean (SD)_number of treated measures per individual	4.29 (5.64)	4.18 (5.49)	3.87 (5.60)	4.08 (5.44)
Mean (SD)_number of untreated measures per individual	10.55 (10.64)	9.91 (9.83)	10.50 (11.18)	11.84 (11.52)
Mean (SD) age at treated BP measurement	68.42 (10.69)	70.59 (11.80)	73.61 (9.69)	70.71 (10.65)
Mean (SD) age at untreated BP measurement	69.34 (11.54)	71.11 (12.52)	73.65 (10.79)	70.85 (12.25)
Mean (SD) BMI	29.14 (6.04)	24.88 (4.43)	27.12 (5.38)	28.64 (6.09)
Mean (SD) systolic treated	133.65 (15.28)	132.39 (14.41)	133.23 (14.94)	133.33 (14.69)
Mean (SD) diastolic treated	75.13 (9.98)	73.65 (10.54)	73.17 (9.93)	72.65 (9.86)
Mean (SD) systolic untreated	136.89 (17.50)	134.80 (17.14)	133.78 (17.51)	134.61 (17.13)
Mean (SD) diastolic untreated	75.77 (11.25)	73.51 (11.08)	73.09 (10.99)	72.97 (11.44)
% with Heart Failure	11.2	5.3	10.9	11.6
% with diabetes	38.2	35.9	22.8	36.7
% with CKD	43.1	34.6	41.4	36.5

	Ancestry group			
	AFR	**EAS**	**EUR**	**LAT**
Hydrochlorothiazide (HCTZ) analysis				
Mean (SD) number of treated measures per individual	5.87 (6.55)	5.01 (5.33)	5.02 (5.70)	5.08 (5.35)
Mean (SD) number of untreated measures per individual	10.65 (11.91)	9.76 (9.58)	10.40 (10.85)	11.71 (12.04)
Mean (SD) age at treated BP measurement	63.61 (11.49)	65.49 (10.92)	69.49 (10.03)	64.91 (12.14)
Mean (SD) age at untreated BP measurement	66.16 (12.48)	68.01 (11.80)	71.66 (11.06)	66.17 (13.36)
Mean (SD)_BMI	30.39 (6.14)	25.41 (4.35)	27.69 (5.73)	29.47 (6.27)
Mean (SD) systolic treated	131.31 (13.44)	130.06 (13.71)	131.72 (13.67)	132.10 (14.05)
Mean (SD) diastolic treated	76.72 (9.60)	76.05 (9.67)	75.16 (9.37)	76.19 (9.55)
Mean (SD) systolic untreated	136.49 (16.09)	134.15 (16.03)	134.39 (16.36)	134.37 (16.11)
Mean (SD) diastolic untreated	77.07 (10.74)	75.03 (10.88)	74.17 (10.81)	75.12 (10.62)
% with Heart Failure	5.2	3.4	5.8	6.1
% with diabetes	32.1	33.2	19.6	34.5
% with CKD	25.1	23.2	31.7	26.6

Table 3.3: Tukey's post hoc test results comparing average systolic blood pressure responses between ancestry groups for ACE inhibitors, CCBs, and HCTZ monotherapies.

	SBP difference	lower	upper	P value adj
ACE inhibitor				
EAS-AFR	-2.78	-4.99	-0.58	0.0064
EUR-AFR	-1.85	-3.69	-0.02	0.0460
LAT-AFR	-2.40	-4.49	-0.31	0.0166
EUR-EAS	0.93	-0.39	2.25	0.2708
LAT-EAS	0.38	-1.28	2.04	0.9343
LAT-EUR	-0.55	-1.67	0.57	0.5928
CCBs				
EAS-AFR	-0.03	-2.52	2.46	1.0000
EUR-AFR	1.04	-0.88	2.96	0.5049
LAT-AFR	0.75	-1.72	3.22	0.8618
EUR-EAS	1.07	-0.68	2.83	0.3975
LAT-EAS	0.78	-1.56	3.13	0.8254
LAT-EUR	-0.29	-2.02	1.44	0.9742
HCTZ				
EAS-AFR	0.80	-1.49	3.08	0.8069
EUR-AFR	2.24	0.43	4.05	0.0080
LAT-AFR	3.02	0.80	5.24	0.0027
EUR-EAS	1.44	-0.08	2.97	0.0717
LAT-EAS	2.22	0.23	4.22	0.0221
LAT-EUR	0.78	-0.65	2.20	0.4993

Table 3.4: Tukey's post hoc test results comparing average diastolic blood pressure responses between ancestry groups for ACE inhibitors, CCBs, and HCTZ monotherapies.

	DBP difference	lower	upper	P value adj
ACE inhibitors				
EAS-AFR	-0.86	-2.22	0.49	0.35
EUR-AFR	-0.81	-1.94	0.31	0.25
LAT-AFR	-0.56	-1.84	0.72	0.67
EUR-EAS	0.05	-0.76	0.87	1.00
LAT-EAS	0.30	-0.72	1.32	0.87
LAT-EUR	0.25	-0.44	0.94	0.79
CCBs				
EAS-AFR	0.38	-1.11	1.87	0.92
EUR-AFR	-0.08	-1.23	1.07	1.00
LAT-AFR	-0.01	-1.49	1.47	1.00
EUR-EAS	-0.46	-1.51	0.60	0.68
LAT-EAS	-0.39	-1.79	1.02	0.89
LAT-EUR	0.07	-0.97	1.10	1.00
HCTZ				
EAS-AFR	0.43	-0.98	1.83	0.86
EUR-AFR	0.71	-0.40	1.82	0.36
LAT-AFR	1.20	-0.17	2.57	0.11
EUR-EAS	0.28	-0.66	1.22	0.87
LAT-EAS	0.77	-0.46	2.00	0.37
LAT-EUR	0.49	-0.39	1.37	0.48

Table 3.5: Estimated effect of doubling dose of monotherapy on SBP in mm HG by different drug classes, controlled for sex, genetic PCs, age, diabetes, chronic kidney disease and heart failure status.

Drug	Ancestry	N	Effect of 100% dose increase	95% C.I. lower	95% C.I. upper	P value
CCBs						
	EUR	4357	-4.00	-4.97	-3.03	p = 0.000
	LAT	469	-0.33	-3.24	2.57	p = 0.823
	AFR	373	-3.85	-7.11	-0.59	p = 0.021
	EAS	454	-3.13	-5.87	-0.40	p = 0.025
ACE Inhibitors						
	EUR	9660	0.28	-0.15	0.72	p = 0.205
	LAT	1086	0.18	-1.08	1.44	p = 0.781
	AFR	380	0.64	-1.49	2.77	p = 0.556
	EAS	756	-0.04	-1.56	1.47	p = 0.954
HCTZ						
	EAS	571	-3.83	-8.66	1.01	p = 0.121
	EUR	6361	-1.99	-3.20	-0.78	p = 0.001
	LAT	664	-6.08	-9.66	-2.51	p = 0.001
	AFR	396	-2.67	-7.40	2.06	p = 0.270

Table 3.6: Estimated effect on DBP in mm HG of doubling dose of monotherapy by different drug classes, controlled for sex, genetic PCs, age, diabetes, chronic kidney disease and heart failure status.

Drug	Ancestry	N	Effect of 100% dose increase	C.I. Lower	C.I. Upper	P value
CCBs						
	EUR	4357	-2.05	-2.61	-1.48	p = 0.000
	LAT	469	-1.31	-3.15	0.54	p = 0.166
	AFR	373	-2.86	-4.72	-0.99	p = 0.003
	EAS	454	-2.69	-4.27	-1.10	p = 0.001
ACE Inhibitors						
	EUR	9660	0.20	-0.06	0.45	p = 0.136
	LAT	1086	-0.14	-0.88	0.61	p = 0.719
	AFR	380	0.32	-0.97	1.62	p = 0.625
	EAS	756	-0.18	-1.10	0.74	p = 0.702
HCTZ						
	EAS	571	-0.60	-3.46	2.26	p = 0.681
	EUR	6361	-0.42	-1.14	0.30	p = 0.252
	LAT	664	-2.24	-4.46	-0.03	p = 0.048
	AFR	396	-0.75	-3.69	2.19	p = 0.618

Figures

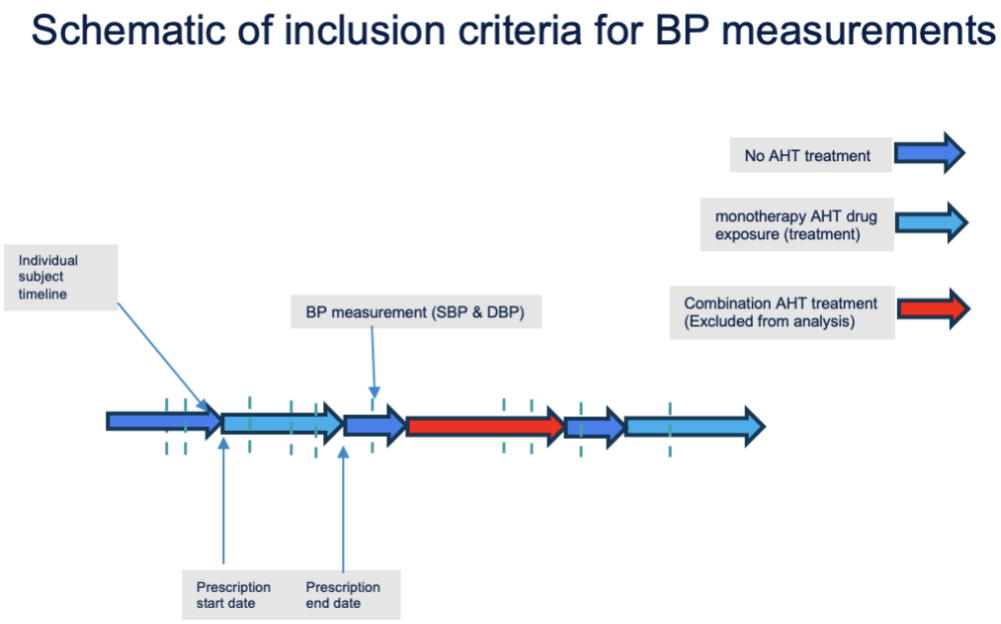

Figure 3.1: Schematic of blood pressure measurements and AHT treatments included in this study from the GERA cohort.

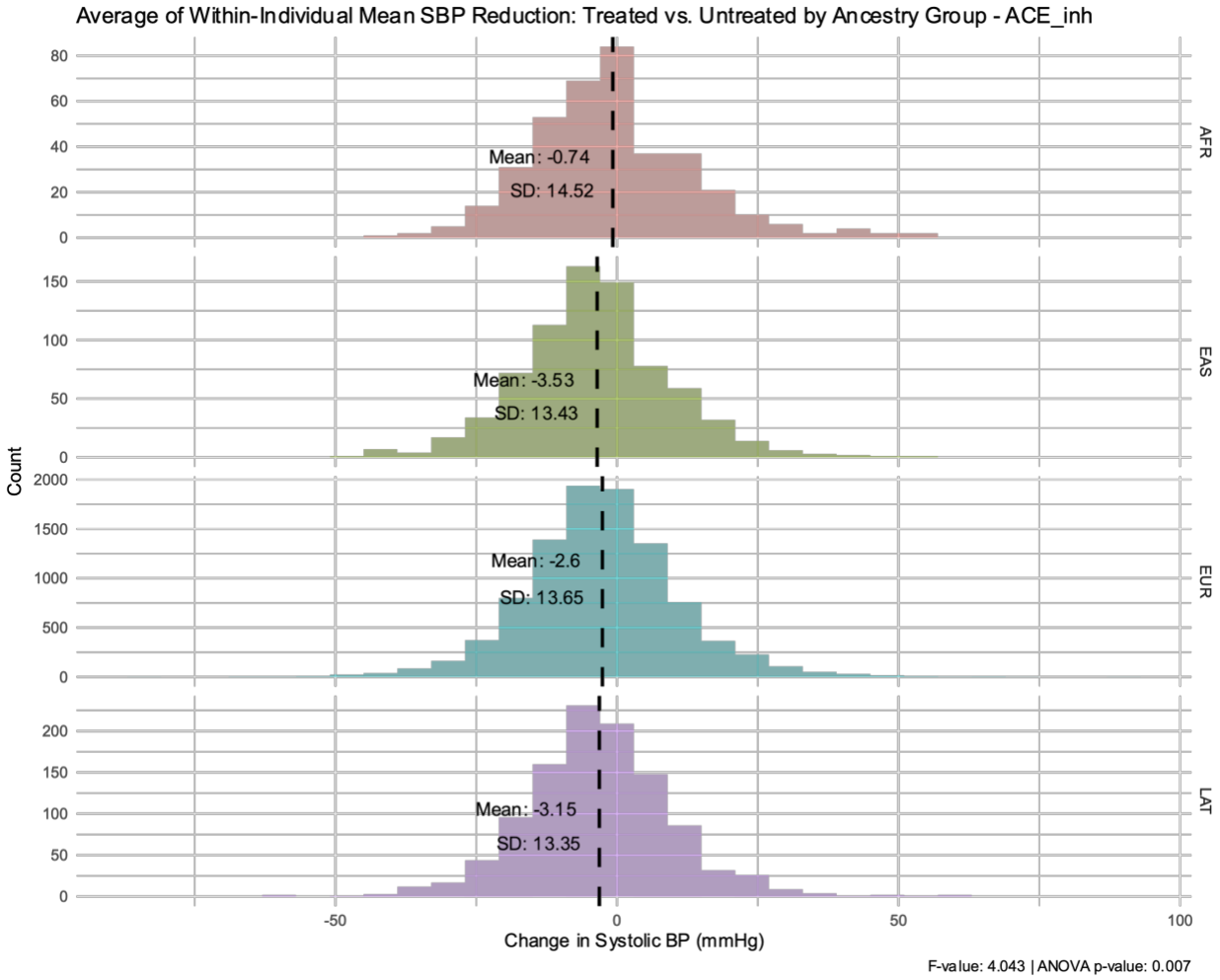

Figure 3.2: Histogram of average within-individual systolic blood pressure (SBP) changes (treated vs. untreated) for ACE-inhibitors across ancestries, and P-value for Anova.

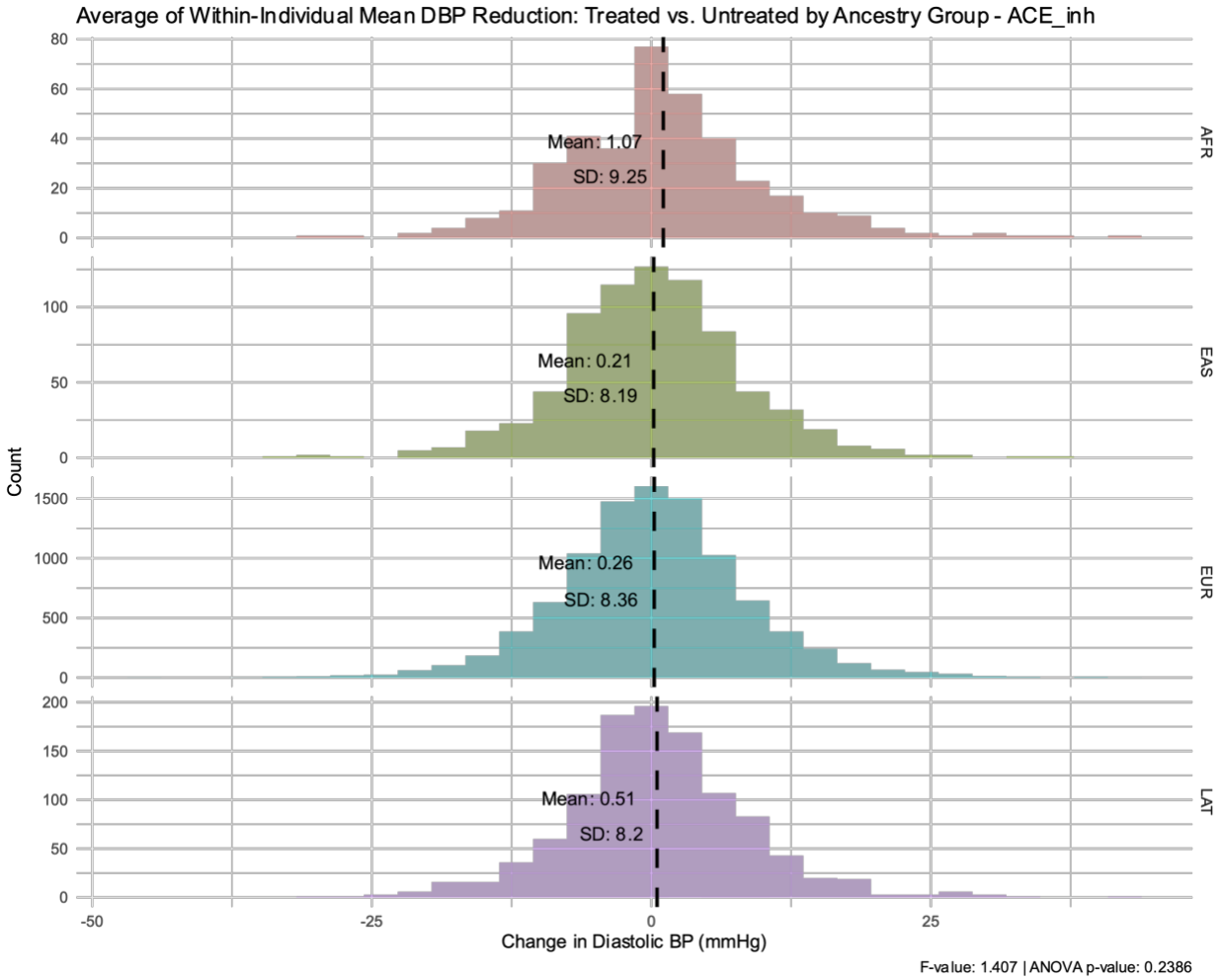

Figure 3.3: Histogram of average within-individual diastolic blood pressure (DBP) changes (treated vs. untreated) for ACE-inhibitors across ancestries, and P-value for Anova.

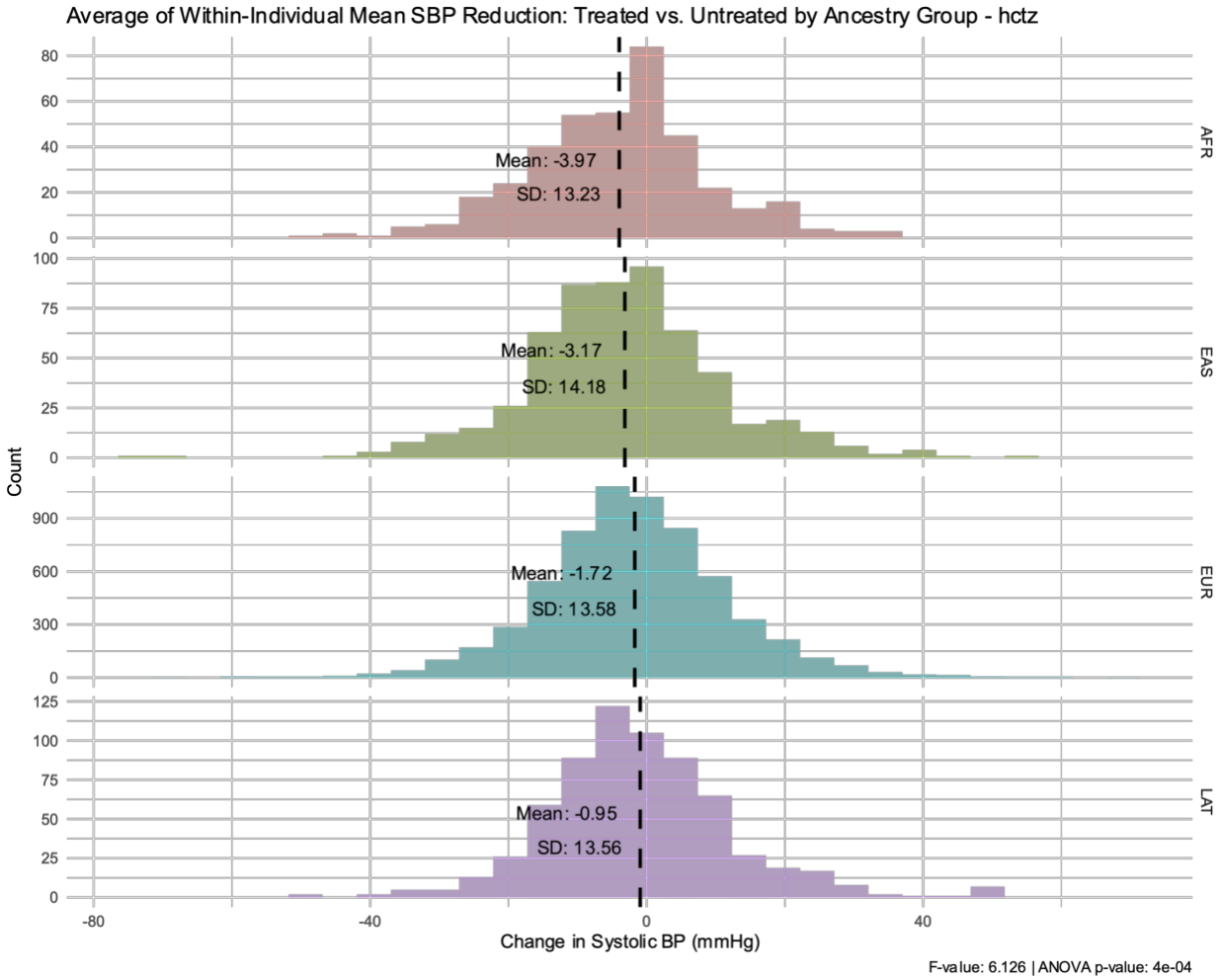

Figure 3.4: Histogram of average within-individual systolic blood pressure (SBP) changes (treated vs. untreated) for HCTZs across ancestries, and P-value for Anova.

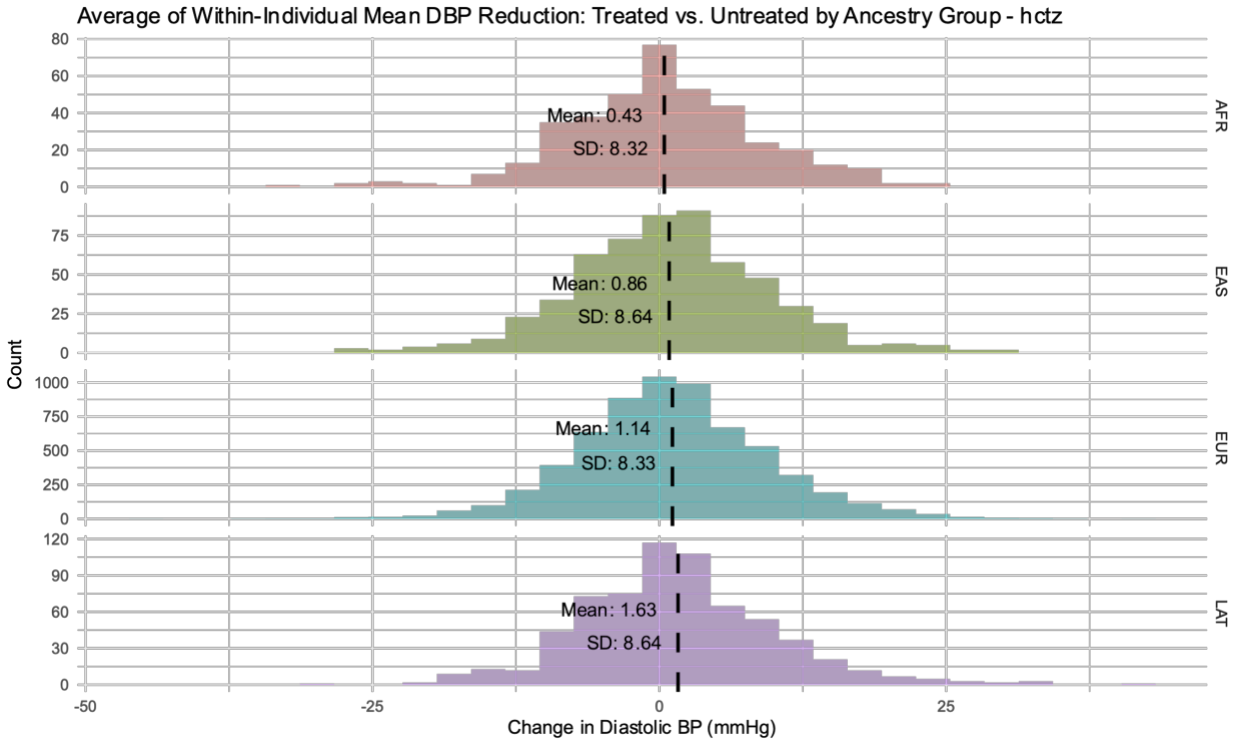

Figure 3.5: Histogram of average within-individual diastolic blood pressure (DBP) changes (treated vs. untreated) for HCTZs across ancestries, and P-value for Anova.

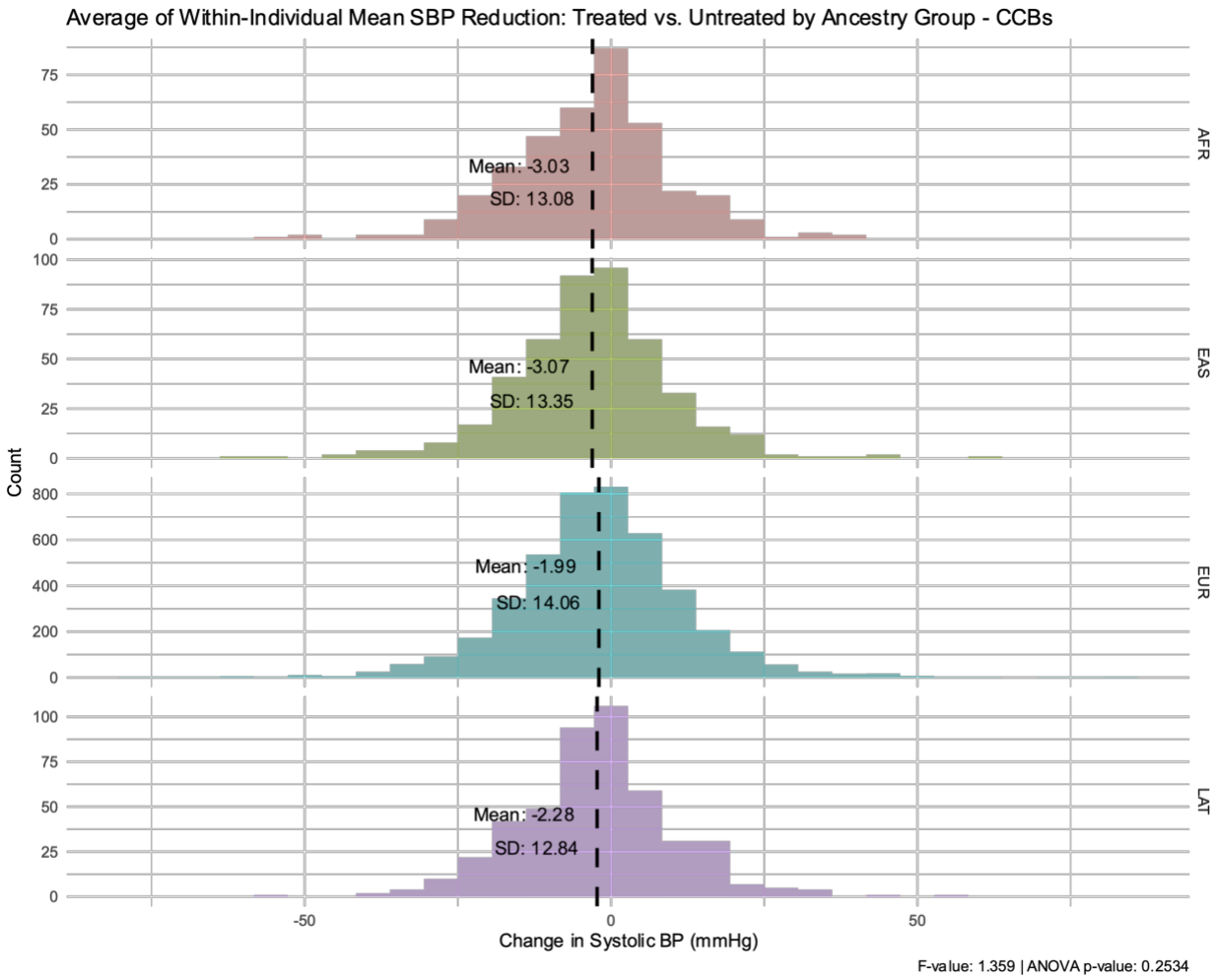

Figure 3.6: Histogram of average within-individual systolic blood pressure (SBP) changes (treated vs. untreated) for CCBs across ancestries, and P-value for Anova.

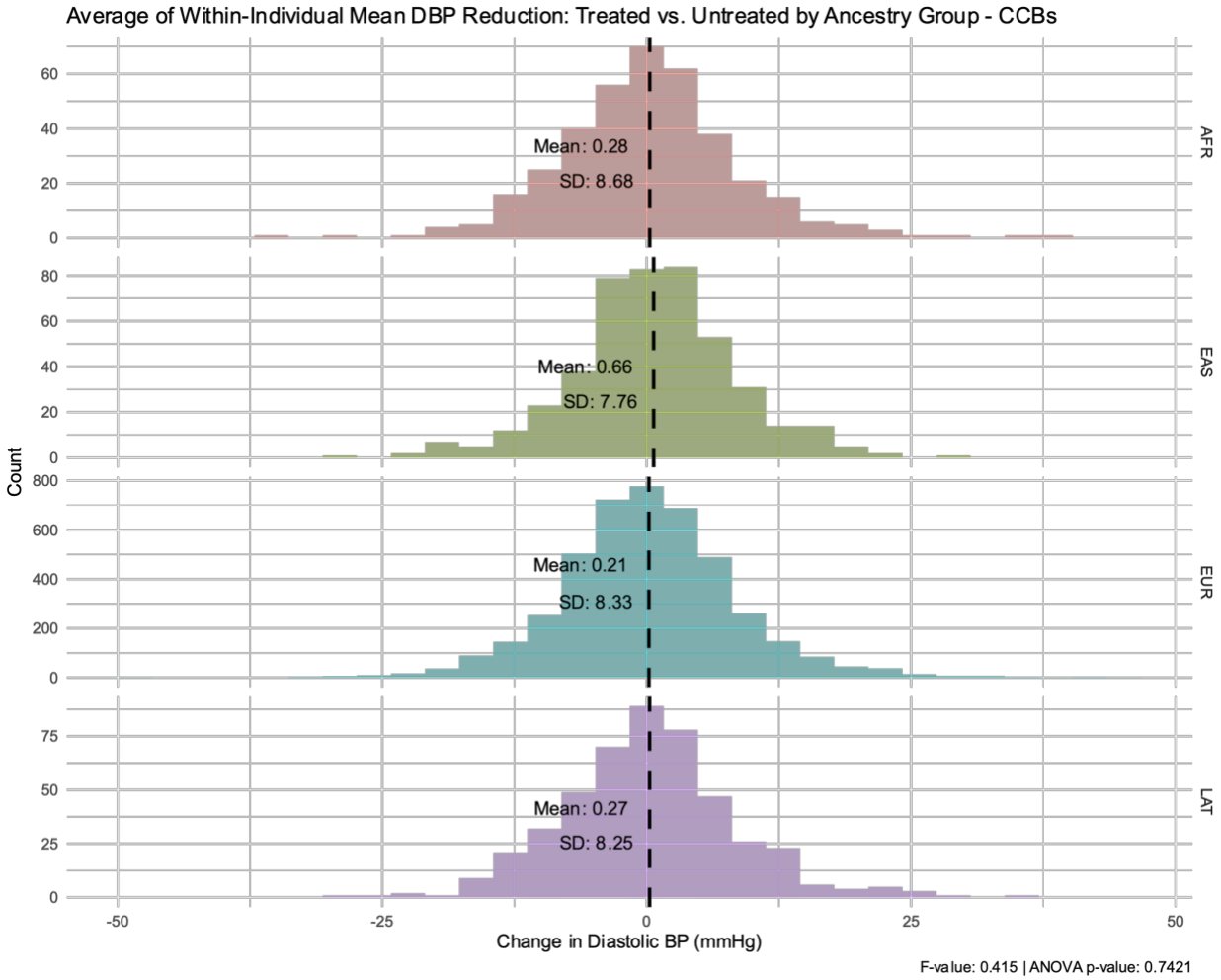

Figure 3.7: Histogram of average within-individual diastolic blood pressure (DBP) changes (treated vs. untreated) for CCBs across ancestries, and P-value for Anova.

References

1. Kurniansyah, N. *et al.* Evaluating the use of blood pressure polygenic risk scores across race/ethnic background groups. *Nat. Commun.* **14**, 3202 (2023).
2. Koch, S., Schmidtke, J., Krawczak, M. & Caliebe, A. Clinical utility of polygenic risk scores: a critical 2023 appraisal. *J. Community Genet.* **14**, 471–487 (2023).
3. Staerk, C., Klinkhammer, H., Wistuba, T., Maj, C. & Mayr, A. Generalizability of polygenic prediction models: how is the R2 defined on test data? *BMC Med. Genomics* **17**, 132 (2024).
4. Dudbridge, F. Power and Predictive Accuracy of Polygenic Risk Scores. *PLoS Genet.* **9**, e1003348 (2013).
5. Kullo, I. J. *et al.* Polygenic scores in biomedical research. *Nat. Rev. Genet.* **23**, 524–532 (2022).
6. Sakaue, S. *et al.* A cross-population atlas of genetic associations for 220 human phenotypes. *Nat. Genet.* **53**, 1415–1424 (2021).
7. Liang, J. *et al.* Single-trait and multi-trait genome-wide association analyses identify novel loci for blood pressure in African-ancestry populations. *PLoS Genet.* **13**, e1006728 (2017).
8. Keaton, J. M. *et al.* Genome-wide analysis in over 1 million individuals of European ancestry yields improved polygenic risk scores for blood pressure traits. *Nat. Genet.* **56**, 778–791 (2024).
9. Banda, Y. *et al.* Characterizing Race/Ethnicity and Genetic Ancestry for 100,000 Subjects in the Genetic Epidemiology Research on Adult Health and Aging (GERA) Cohort. *Genetics* **200**, 1285–1295 (2015).

10. Hoffmann, T. J. *et al.* Next generation genome-wide association tool: design and coverage of a high-throughput European-optimized SNP array. *Genomics* **98**, 79–89 (2011).

11. Hoffmann, T. J. *et al.* Design and coverage of high throughput genotyping arrays optimized for individuals of East Asian, African American, and Latino race/ethnicity using imputation and a novel hybrid SNP selection algorithm. *Genomics* **98**, 422–430 (2011).

12. Kvale, M. N. *et al.* Genotyping Informatics and Quality Control for 100,000 Subjects in the Genetic Epidemiology Research on Adult Health and Aging (GERA) Cohort. *Genetics* **200**, 1051–1060 (2015).

13. Hoffmann, T. J. *et al.* Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. *Nat. Genet.* **49**, 54–64 (2017).

14. Ruan, Y. *et al.* Improving polygenic prediction in ancestrally diverse populations. *Nat. Genet.* **54**, 573–580 (2022).

15. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, 7 (2015).

16. Nakagawa, S., Johnson, P. C. D. & Schielzeth, H. The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *J. R. Soc. Interface* **14**, 20170213 (2017).

17. Mostafavi, H. *et al.* Variable prediction accuracy of polygenic scores within an ancestry group. *eLife* **9**, e48376 (2020).

18. Price, A. L. *et al.* A Genomewide Admixture Map for Latino Populations. *Am. J. Hum. Genet.* **80**, 1024–1036 (2007).

19. Conrad, D. F. *et al.* A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. *Nat. Genet.* **38**, 1251–1260 (2006).

20. Yang, G., Mishra, M. & Perera, M. A. Multi-Omics Studies in Historically Excluded Populations: The Road to Equity. *Clin. Pharmacol. Ther.* **113**, 541–556 (2023).

21. Brown, W. M. *et al.* Age-stratified heritability estimation in the Framingham Heart Study families. *BMC Genet.* **4**, S32 (2003).

22. Mensah, G. A. Hypertension and Target Organ Damage: Don't Believe Everything You Think! *Ethn. Dis.* **26**, 275 (2016).

23. Krishnan, A. *et al.* Health-Related Quality of Life in People Across the Spectrum of CKD. *Kidney Int. Rep.* **5**, 2264 (2020).

24. Dhamoon, M. S. *et al.* Quality of life declines after first ischemic stroke: The Northern Manhattan Study[image][image]. *Neurology* **75**, 328 (2010).

25. Nieminen, M. S. *et al.* The patient perspective: Quality of life in advanced heart failure with frequent hospitalisations. *Int. J. Cardiol.* **191**, 256–264 (2015).

26. Lui, J. N. M. *et al.* Impact of New Cardiovascular Events on Quality of Life and Hospital Costs in People With Cardiovascular Disease in the United Kingdom and United States. *J. Am. Heart Assoc.* **12**, e030766 (2023).

27. Lewis, C. M. & Vassos, E. Polygenic risk scores: from research tools to clinical instruments. *Genome Med.* **12**, 44 (2020).

28. Polygenic Risk Score Task Force of the International Common Disease Alliance. Responsible use of polygenic risk scores in the clinic: potential benefits, risks and gaps. *Nat. Med.* **27**, 1876–1884 (2021).

29. Vaura, F. *et al.* Polygenic Risk Scores Predict Hypertension Onset and Cardiovascular Risk. *Hypertension* **77**, 1119–1127 (2021).

30. Hoffmann, T. J. *et al.* Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. *Nat. Genet.* **49**, 54–64 (2017).

31. Lambert, S. A. *et al.* Enhancing the Polygenic Score Catalog with tools for score calculation and ancestry normalization. *Nat. Genet.* 1–6 (2024) doi:10.1038/s41588-024-01937-x.

32. Mandla, R. *et al.* Polygenic scores for longitudinal prediction of incident type 2 diabetes in an ancestrally and medically diverse primary care physician network: a patient cohort study. *Genome Med.* **16**, 63 (2024).

33. Jung, H. *et al.* Integration of risk factor polygenic risk score with disease polygenic risk score for disease prediction. *Commun. Biol.* **7**, 1–13 (2024).

34. A, M. *et al.* Stroke genetics informs drug discovery and risk prediction across ancestries. - Abstract - Europe PMC.

35. Khera, A. V. *et al.* Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat. Genet.* **50**, 1219–1224 (2018).

36. Shim, I. *et al.* Clinical utility of polygenic scores for cardiometabolic disease in Arabs. *Nat. Commun.* **14**, 6535 (2023).

37. Tcheandjieu, C. *et al.* Large-scale genome-wide association study of coronary artery disease in genetically diverse populations. *Nat. Med.* **28**, 1679–1692 (2022).

38. Gunn, S. *et al.* Comparison of methods for building polygenic scores for diverse populations. *Hum. Genet. Genomics Adv.* **6**, (2025).

39. Jung, E., Kong, S. Y., Ro, Y. S., Ryu, H. H. & Shin, S. D. Serum Cholesterol Levels and Risk of Cardiovascular Death: A Systematic Review and a Dose-Response Meta-Analysis of Prospective Cohort Studies. *Int. J. Environ. Res. Public. Health* **19**, 8272 (2022).

40. Parcha, V. *et al.* Association of a Multiancestry Genome-Wide Blood Pressure Polygenic Risk Score With Adverse Cardiovascular Events. *Circ. Genomic Precis. Med.* **0**, e003946.

41. ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group. The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial. Major outcomes in high-risk hypertensive patients randomized to angiotensin-converting enzyme inhibitor or calcium channel blocker vs diuretic: The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). *JAMA* **288**, 2981–2997 (2002).

42. Cooper-DeHoff, R. M. & Johnson, J. A. Hypertension pharmacogenomics: in search of personalized treatment approaches. *Nat. Rev. Nephrol.* **12**, 110 (2015).

43. Tobin, M. D., Sheehan, N. A., Scurrah, K. J. & Burton, P. R. Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. *Stat. Med.* **24**, 2911–2935 (2005).

44. Heran, B. S., Wong, M. M., Heran, I. K. & Wright, J. M. Blood pressure lowering efficacy of angiotensin converting enzyme (ACE) inhibitors for primary hypertension. *Cochrane Database Syst. Rev.* **2008**, CD003823 (2008).

45. Herman, L. L., Weber, P. & Bashir, K. Hydrochlorothiazide. in *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2024).

46. McKeever, R. G., Patel, P. & Hamilton, R. J. Calcium Channel Blockers. in *StatPearls [Internet]* (StatPearls Publishing, 2024).

47. Brewster, L. M. & Seedat, Y. K. Why do hypertensive patients of African ancestry respond better to calciumblockers and diuretics than to ACE inhibitors and β-adrenergic blockers? Asystematic review. *BMC Med.* **11**, 141 (2013).

48. Brewster, L. M., van Montfrans, G. A. & Kleijnen, J. Systematic Review: Antihypertensive Drug Therapy in Black Patients. *Ann. Intern. Med.* **141**, 614–627 (2004).

49. Mokwe, E. *et al.* Determinants of Blood Pressure Response to Quinapril in Black and White Hypertensive Patients. *Hypertension* **43**, 1202–1207 (2004).

50. Herman, L. L., Padala, S. A., Ahmed, I. & Bashir, K. Angiotensin-Converting Enzyme Inhibitors (ACEI). in *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2024).

51. Hinton, T. C. *et al.* Investigation and Treatment of High Blood Pressure in Young People. *Hypertension* **75**, 16–22 (2020).

52. Deary, A. J. *et al.* Double-blind, placebo-controlled crossover comparison of five classes of antihypertensive drugs. *J. Hypertens.* **20**, 771 (2002).

53. Dickerson, J. E., Hingorani, A. D., Ashby, M. J., Palmer, C. R. & Brown, M. J. Optimisation of antihypertensive treatment by crossover rotation of four major classes. *Lancet Lond. Engl.* **353**, 2008–2013 (1999).

54. Williams, S. F., Nicholas, S. B., Vaziri, N. D. & Norris, K. C. African Americans, hypertension and the renin angiotensin system. *World J. Cardiol.* **6**, 878 (2014).

55. Duarte, J. D. & Cooper-DeHoff, R. M. Mechanisms for blood pressure lowering and metabolic effects of thiazide and thiazide-like diuretics. *Expert Rev. Cardiovasc. Ther.* **8**, 793 (2010).

56. Rapoport, R. M. & Soleimani, M. Mechanism of Thiazide Diuretic Arterial Pressure Reduction: The Search Continues. *Front. Pharmacol.* **10**, 815 (2019).

57. Papademetriou, V., Narayan, P. & Kokkinos, P. Angiotensin-Converting Enzyme Inhibitors and Angiotensin Receptor Blockers in African-American Patients With Hypertension. *J. Clin. Hypertens.* **6**, 310 (2007).

Appendix

Table A.1. List of ICD codes used to define outcome incidence:

Chronic kidney disease (CKD) — ICD codes:

ICD-9 Codes:

- **585**: Chronic kidney disease (CKD), unspecified
- **585.1**: Chronic kidney disease, stage 1
- **585.2**: Chronic kidney disease, stage 2 (mild)
- **585.3**: Chronic kidney disease, stage 3 (moderate)
- **585.4**: Chronic kidney disease, stage 4 (severe)
- **585.5**: Chronic kidney disease, stage 5
- **585.6**: End-stage renal disease (ESRD)
- **585.9**: Chronic kidney disease, unspecified

ICD-10 Codes:

- **N18.1**: Chronic kidney disease, stage 1
- **N18.2**: Chronic kidney disease, stage 2 (mild)
- **N18.3**: Chronic kidney disease, stage 3 (unspecified)
- **N18.30**: Chronic kidney disease, stage 3 unspecified
- **N18.31**: Chronic kidney disease, stage 3a (mild to moderate)
- **N18.32**: Chronic kidney disease, stage 3b (moderate to severe)
- **N18.4**: Chronic kidney disease, stage 4 (severe)
- **N18.5**: Chronic kidney disease, stage 5
- **N18.6**: End-stage renal disease (ESRD)

- **N18.9**: Chronic kidney disease, unspecified

Heart failure (HF) — ICD codes:

ICD-9 Codes:

- **428.0**: Congestive heart failure, unspecified
- **428.1**: Left heart failure
- **428.20**: Systolic heart failure, unspecified
- **428.21**: Acute systolic heart failure
- **428.22**: Chronic systolic heart failure
- **428.23**: Acute on chronic systolic heart failure
- **428.30**: Diastolic heart failure, unspecified
- **428.31**: Acute diastolic heart failure
- **428.32**: Chronic diastolic heart failure
- **428.33**: Acute on chronic diastolic heart failure
- **428.40**: Combined systolic and diastolic heart failure, unspecified
- **428.41**: Acute combined systolic and diastolic heart failure
- **428.42**: Chronic combined systolic and diastolic heart failure
- **428.43**: Acute on chronic combined systolic and diastolic heart failure

ICD-10 Codes:

- **I50.9**: Heart failure, unspecified
- **I50.1**: Left ventricular heart failure
- **I50.20**: Unspecified systolic (congestive) heart failure
- **I50.21**: Acute systolic (congestive) heart failure
- **I50.22**: Chronic systolic (congestive) heart failure

- **I50.23**: Acute on chronic systolic (congestive) heart failure
- **I50.30**: Unspecified diastolic (congestive) heart failure
- **I50.31**: Acute diastolic (congestive) heart failure
- **I50.32**: Chronic diastolic (congestive) heart failure
- **I50.33**: Acute on chronic diastolic (congestive) heart failure
- **I50.40**: Unspecified combined systolic and diastolic heart failure
- **I50.41**: Acute combined systolic and diastolic heart failure
- **I50.42**: Chronic combined systolic and diastolic heart failure
- **I50.43**: Acute on chronic combined systolic and diastolic heart failure

Ischemic stroke (IS) — ICD codes:

ICD-9 Codes:

- 434.00 – Cerebral thrombosis without mention of cerebral infarction
- 434.01 – Cerebral thrombosis with cerebral infarction
- 434.10 – Cerebral embolism without mention of cerebral infarction
- 434.11 – Cerebral embolism with cerebral infarction
- 434.90 – Cerebral artery occlusion without mention of cerebral infarction
- 434.91 – Cerebral artery occlusion with cerebral infarction

ICD-10 Codes:

- I63.30 – Cerebral infarction due to thrombosis of unspecified cerebral artery
- I63.311 – Cerebral infarction due to thrombosis of right middle cerebral artery
- I63.40 – Cerebral infarction due to embolism of unspecified cerebral artery
- I63.411 – Cerebral infarction due to embolism of right middle cerebral artery
- I63.9 – Cerebral infarction, unspecified

Atrial fibrillation (Afib):

ICD-9 Code:

- **427.31**: Atrial fibrillation

ICD-10 Codes:

- **I48.0**: Paroxysmal atrial fibrillation
- **I48.1**: Persistent atrial fibrillation
- **I48.2**: Chronic atrial fibrillation
- **I48.91**: Unspecified atrial fibrillation

Cardiomyopathy (CMP):

ICD-9 Codes:

- **425.4**: Other primary cardiomyopathies
- **425.8**: Other specified cardiomyopathies
- **425.9**: Cardiomyopathy, unspecified

ICD-10 Codes:

- **I42.0**: Dilated cardiomyopathy
- **I42.1**: Obstructive hypertrophic cardiomyopathy
- **I42.2**: Other hypertrophic cardiomyopathy
- **I42.9**: Cardiomyopathy, unspecified

Coronary artery disease (CAD):

ICD-9 Codes:

- **414.00**: Coronary atherosclerosis of unspecified type of vessel, unspecified
- **414.01**: Coronary atherosclerosis of unspecified type of vessel, with angina pectoris

- **414.2**: Other forms of chronic ischemic heart disease

  (This includes various forms of CAD not specified in other categories.)

ICD-10 Codes:

- **I25.10**: Atherosclerotic heart disease of native coronary artery without angina pectoris
- **I25.11**: Atherosclerotic heart disease of native coronary artery with angina pectoris
- **I25.82**: Atherosclerosis of coronary artery bypass graft(s)

Myocardial infarction (MI):

ICD-9 Codes:

- **410.00**: Acute myocardial infarction, unspecified, without mention of acute myocardial infarction
- **410.90**: Acute myocardial infarction, unspecified, without mention of acute myocardial infarction

ICD-10 Codes:

- **I21.9**: Acute myocardial infarction, unspecified
- **I21.3**: ST elevation (STEMI) myocardial infarction of the inferior wall
- **I21.4**: ST elevation (STEMI) myocardial infarction of the other wall

Table A.2. Drug conversions for 1 daily dose (DD) by different drug classes.

Each of these drug doses equal 1 DD within each of the drug classes.
ACE-Inhibitors BENAZEPRIL 10 MG CAPTOPRIL 12.5 MG ENALAPRIL 5 MG FOSINOPRIL 10 MG LISINOPRIL 10 MG MOEXIPRIL 7.5 MG QUINAPRIL 10 MG RAMIPRIL 2.5 MG (https://hospitalhandbook.ucsf.edu/02-ace-inhibitor-conversion-table/02-ace-inhibitor-conversion-table)
Hydrochlorothiazide HYDROCHLOROTHIAZIDE 25 MG
dihydropyridine Calcium Channel blockers AMLODIPINE 5 MG FELODIPINE 5 MG NIFEDIPINE 30 MG

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:
Yusuph Mavura
1CB7AD4CB19D43E...
Author Signature

3/21/2025
Date