blood advances

American Society of Hematology
2021 L Street NW, Suite 900,
Washington, DC 20036
Phone: 202-776-0544 | Fax 202-776-0545
bloodadvances@hematology.org

Meta-analysis of genome-wide association studies of stable warfarin dose in patients of African ancestry

Tracking no: ADV-2024-014227R1

Innocent Asiimwe (University of Liverpool, United Kingdom) Marc Blockman (University of Cape Town, South Africa) Larisa Cavallari (University of Florida, United States) Karen Cohen (University of Cape Town, South Africa) Clint Cupido (University of Cape Town, South Africa) Collet Dandara (Faculty of Health Sciences, University of Cape Town, South Africa) Brittney Davis (University of Alabama at Birmingham, United States) Barry Jacobson (University of the Witwatersrand, South Africa) Julie Johnson (Ohio State University, United States) Mohammed Lamorde (,) Nita Limdi (University of Alabama at Birmingham, United States) Jennie Morgan (Western Cape Department of Health, South Africa) Johannes Mouton (University of Cape Town, South Africa) Sarudzai Muyambo (University of Zimbabwe, Zimbabwe) Doreen Nakagaayi (Uganda Heart Institute, Uganda) Arinao Ndadza (University of Cape Town, South Africa) Emmy Okello (Uganda Heart Institute, Uganda) Minoli Perera (Northwestern University, United States) Elise Schapkaitz (University of the Witwatersrand, South Africa) Christine Sekaggya-Wiltshire (Infectious Diseases Institute, Makerere University /Mulago national referral hospital, Uganda) Jerome Semakula (Infectious Diseases Institute, College of Health Sciences, Makerere University, Kampala, Uganda., Uganda) Gayle Tatz (University of Cape Town, South Africa) Catriona Waitt (University of Liverpool, United Kingdom) Guang Yang (Northwestern University, United States) Eunice Zhang (University of Liverpool, United Kingdom) Andrea Jorgensen (University of Liverpool, United Kingdom) Munir Pirmohamed (The University of Liverpool, United Kingdom)

Abstract:
Warfarin dose requirements are highly variable due to clinical and genetic factors. While genetic variants influencing warfarin dose have been identified in European and East Asian populations, more work is needed to identify African-specific genetic variants to help optimize warfarin dosing. We performed genome-wide association studies (GWAS) in four African cohorts from Uganda, South Africa, and Zimbabwe, totalling 989 warfarin-treated participants who reached stable dose and had international normalized ratios within therapeutic ranges. We also included two African American cohorts recruited by the International Warfarin Pharmacogenetics Consortium (n=316) and the University of Alabama at Birmingham (n=199). Following the GWAS, we performed standard error-weighted meta-analyses and then conducted stepwise conditional analyses to account for known loci (the CYP2C cluster SNP rs12777823 and CYP2C9 in chromosome 10; VKORC1 in chromosome 16). The genome-wide significance threshold was set at $P<5\times10-8$. The meta-analysis, comprising 1,504 participants identified 242 significant SNPs across three genomic loci, with 99.6% of these located within known loci on chromosomes 10 (top SNP: rs58800757, $P=4.27\times10-13$) and 16 (top SNP: rs9925964, $P=9.97\times10-16$). Adjustment for the VKORC1 SNP -1639G>A revealed an additional locus on chromosome 2 (top SNPs rs116057875/rs115254730/rs115240773, $P=3.64\times10-8$), implicating the MALL gene, that could indirectly influence warfarin response through interactions with caveolin-1. In conclusion, our meta-analysis of six cohorts of warfarin-treated patients of African ancestry reaffirmed the importance of CYP2C9 and VKORC1 in influencing warfarin dose requirements. We also identified a new locus (MALL), that still requires direct evidence of biological plausibility.

Conflict of interest: COI declared - see note

COI notes: M.P. currently receives partnership funding for the following: MRC Clinical Pharmacology Training Scheme (co-funded by MRC and Roche, UCB, Eli Lilly and Novartis). He has developed an HLA genotyping panel with MC Diagnostics, but does not benefit financially from this. He is part of the IMI Consortium ARDAT (www.ardat.org). C.W. is supported by a Wellcome Clinical Research Career Development Fellowship 222075/Z/20/Z. None of these of funding sources have been used for the current paper. All other authors declared no competing interests for this work.

Preprint server: No;

Author contributions and disclosures: MB, KC, ML, JPM, CS-W, JR, CW, EJZ, ALJ, MP contributed to the design and/or funding of the research. All authors contributed to the acquisition of genotype and/or phenotype data. IGA analyzed the results and made the figures. IGA, ALP, and MP drafted the initial manuscript. All authors participated in critical review of the manuscript.

Non-author contributions and disclosures: No;

Agreement to Share Publication-Related Data and Data Sharing Statement: All relevant material is provided in the supplementary material.

Clinical trial registration information (if any):

Meta-analysis of genome-wide association studies of stable warfarin dose in patients of African ancestry.

Innocent G. Asiimwe[1], Marc Blockman[2], Larisa H. Cavallari[3], Karen Cohen[2], Clint Cupido[4], Collet Dandara[5], Brittney H. Davis[6], Barry Jacobson[7], Julie A. Johnson[8], Mohammed Lamorde[9], Nita A. Limdi[6], Jennie Morgan[10], Johannes P. Mouton[2], Sarudzai Muyambo[11], Doreen Nakagaayi[12], Arinao Ndadza[13], Emmy Okello[12], Minoli A. Perera[14], Elise Schapkaitz[15], Christine Sekaggya-Wiltshire[9], Jerome R. Semakula[9], Gayle Tatz[2], Catriona Waitt[1,9], Guang Yang[14,16], Eunice J. Zhang[1], Andrea L. Jorgensen[17,*], Munir Pirmohamed[1,*].

[1]Department of Pharmacology and Therapeutics, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, UK. [2]Division of Clinical Pharmacology, Department of Medicine, University of Cape Town, Cape Town, South Africa. [3]Center for Pharmacogenomics, Department of Pharmacotherapy and Translational Research, University of Florida College of Pharmacy, Gainesville, USA. [4]Victoria Hospital Internal Medicine Research Initiative, Victoria Hospital Wynberg and Department of Medicine, University of Cape Town, Cape Town, South Africa. [5]Pharmacogenomics and Drug Metabolism Research Group, Division of Human Genetics, Department of Pathology, Institute of Infectious Disease and Molecular Medicine (IDM), Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa. [6]Department of Neurology The University of Alabama at Birmingham, Birmingham, USA. [7]Department of Molecular Medicine and Haematology, University of the Witwatersrand, Johannesburg, South Africa. [8]Center for Clinical and Translational Science, College of Medicine, The Ohio State University, Columbus, USA. [9]Infectious Diseases Institute, Makerere University College of Health Sciences, Kampala, Uganda. [10]Metro District Health Services, Western Cape Department of Health, Cape Town, South Africa. [11]Department of Biological Sciences and Ecology, University of Zimbabwe, Harare, Zimbabwe. [12]Uganda Heart Institute, Kampala, Uganda. [13]Pharmacogenomics and Drug Metabolism Research Group, Division of Human Genetics, Department of Pathology, Institute of Infectious Disease and Molecular Medicine, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa. [14]Department of Pharmacology, Center for Pharmacogenomics, Northwestern University, Chicago, USA. [15]Department of Molecular Medicine and Hematology, Charlotte Maxeke Johannesburg Academic Hospital National Health Laboratory System Complex and University of Witwatersrand, Johannesburg, South Africa. [16]Center for Applied Bioinfomatics, St. Jude Children's Research Hospital, Memphis, USA. [17]Department of Health Data Science, Institute of Population Health Sciences, University of Liverpool, Liverpool, UK.

[*]Contributed equally to this study.

Correspondence:

Innocent G. Asiimwe

Department of Pharmacology and Therapeutics, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, UK.

Tel: +441517955387 Email: i.asiimwe@liverpool.ac.uk

Munir Pirmohamed

Department of Pharmacology and Therapeutics, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, UK.

Tel: +441517945549 Email: munirp@liverpool.ac.uk

Data availability statement

All relevant material is provided in the supplementary material.

Counts

Abstract word count: 246

Text word count: 4,568

Table count: 2

Figure count: 4

Reference count: 53

Key Points

- The *CYP2C* and *VKORC1* loci are the most important in determining warfarin dose requirements in individuals of African ancestry.
- We identified a biologically plausible locus (*MALL*) that could indirectly influence warfarin response through interactions with caveolin-1.

Abstract

Warfarin dose requirements are highly variable due to clinical and genetic factors. While genetic variants influencing warfarin dose have been identified in European and East Asian populations, more work is needed to identify African-specific genetic variants to help optimize warfarin dosing. We performed genome-wide association studies (GWAS) in four African cohorts from Uganda, South Africa, and Zimbabwe, totalling 989 warfarin-treated participants who reached stable dose and had international normalized ratios within therapeutic ranges. We also included two African American cohorts recruited by the International Warfarin Pharmacogenetics Consortium (*n*=316) and the University of Alabama at Birmingham (*n*=199). Following the GWAS, we performed standard error-weighted meta-analyses and then conducted stepwise conditional analyses to account for known loci (the *CYP2C* cluster SNP rs12777823 and *CYP2C9* in chromosome 10; *VKORC1* in chromosome 16). The genome-wide significance threshold was set at $P<5\times10^{-8}$. The meta-analysis, comprising 1,504 participants identified 242 significant SNPs across three genomic loci, with 99.6% of these located within known loci on chromosomes 10 (top SNP: rs58800757, $P=4.27\times10^{-13}$) and 16 (top SNP: rs9925964, $P=9.97\times10^{-16}$). Adjustment for the *VKORC1* SNP *-1639G>A* revealed an additional locus on chromosome 2 (*top SNPs rs116057875/rs115254730/rs115240773, P=3.64×10^{-8}*), implicating the *MALL* gene, that could indirectly influence warfarin response through interactions with caveolin-1. In conclusion, our meta-analysis of six cohorts of warfarin-treated patients of African ancestry reaffirmed the importance of *CYP2C9* and *VKORC1* in influencing warfarin dose requirements. We also identified a new locus (*MALL*), that still requires direct evidence of biological plausibility.

Keywords: admixed; African; genome-wide association study; personalized medicine; tractor; warfarin.

Introduction

Warfarin remains the primary choice for oral anticoagulation therapy in resource-limited settings like those in Africa, primarily due to its affordability.[1] However, the continent faces challenges in achieving optimal anticoagulation,[2] mainly due to difficulties in optimizing warfarin dosing. Its narrow therapeutic window, coupled with the significant risk of thrombotic or haemorrhagic events when doses are inadequate or excessive, respectively, contributes to warfarin being a leading cause of hospitalization due to preventable adverse drug reactions in South Africa.[3] Moreover, there exists considerable variability in dose requirements among patients, influenced by clinical and genetic factors. Despite the development of numerous dosing algorithms incorporating these factors, Africa significantly lags behind other regions globally, with less than 1% of these algorithms originating from this region.[4]

Whites and Asians have been extensively studied regarding genetic variants influencing warfarin dose requirements, particularly focusing on single nucleotide polymorphisms (SNPs) within the genes *CYP2C9* (cytochrome P450, family 2, subfamily C, polypeptide 9) and *VKORC1* (vitamin K epoxide reductase complex, subunit 1).[4] The most studied variants, including *CYP2C9*2* (rs1799853), *CYP2C9*3* (rs1057910), and *VKORC1 -1639G>A* (rs9923231), have lower allele frequencies in African populations, resulting in a lesser impact on warfarin dose requirements.[5] Therefore, more work is needed to identify additional variants that may play a more significant role in these populations. One approach to systematically explore the entire genome for such variants is through genome-wide association studies (GWAS). GWAS have been successfully conducted in a range of populations, including African Americans.[6-11] For instance, Perera and colleagues conducted a GWAS that identified a novel variant, rs12777823, within the *CYP2C* gene region in African Americans.[6] This variant, independent of *CYP2C9*2* and *CYP2C9*3*, was found to decrease the warfarin dose by 6.9 mg/week and 9.3 mg/week in African Americans who were heterozygous and homozygous for the A allele, respectively.

To identify rare variants or variants with small effect sizes, GWAS typically require sample sizes in the thousands.[12] However, achieving such large sample sizes in Africa, where pharmacogenomic-related clinical research is still in its early stages, is challenging. Establishing the necessary clinical research infrastructure and ensuring accurate data capture in resource-limited settings demands significant resources.[13] In the United States, obstacles including patient distrust of researchers and clinicians, limited knowledge about genomic research, inadequate culturally-appropriate educational materials, language barriers, and economic disadvantages have hindered the recruitment and retention of racial and ethnic minorities in clinical trials.[14] These challenges are likely to be present in

Africa as well. To help overcome sample-size related issues, we conducted several GWAS for the phenotype of stable warfarin dose in patients of African-ancestry (Africans and African Americans) and pooled the results using meta-analysis.

Methods

Reporting of the study follows the STrengthening the Reporting Of Pharmacogenetic Studies (STROPS) guideline (Supplemental Table 1).[15]

Study design, setting and participants

We used six cohorts including four African and two African American cohorts. The African cohorts were obtained through the WARfarin anticoagulation in PATients in Sub-Saharan Africa (War-PATH, http://warpath.info/) observational study.[16-18] The War-PATH study enrolled participants receiving warfarin treatment, who were on stable dose, defined as consistent prescribed doses for two consecutive clinic visits in the year prior to recruitment. Additionally, participants had to maintain an international normalized ratio (INR) within therapeutic range (2.0–3.0 for venous thromboembolism/atrial fibrillation and 2.5–3.5 for valvular heart disease) at both visits. The first cohort comprised 548 black Africans recruited from 12 Ugandan and South African outpatient clinics/hospital departments between 2018 and 2020.[16-18] The second cohort included an additional 214 black African participants recruited from Charlotte Maxeke Johannesburg Academic Hospital, South Africa (during 2019), INR clinics at Groote Schuur Hospital and Gugulethu Community Health Centre in Western Cape, South Africa (between 2016 and 2017) and Parirenyatwa Group of Hospitals in Harare, Zimbabwe (also between 2016 and 2017). Details of those recruited in South Africa and Zimbabwe between 2016 and 2017 have been previously reported.[19,20] The third cohort consisted of 133 participants of mixed ancestry recruited from the same South African sites between 2018 and 2020. The fourth cohort comprised 94 individuals of mixed ancestry and black Africans recruited from Uganda and South Africa between 2021 and 2022 as part of the War-PATH "bundle of care" study.[21] The above cohorts are referred to as War-PATH cohorts 1–4, respectively, throughout the manuscript.

African-American cohorts included the two cohorts analysed by Perera et al.[6] African Americans recruited via specific sites by members of the International Warfarin Pharmacogenetics Consortium (IWPC, *n* = 316), and those recruited by the University of Alabama at Birmingham (UAB, *n* = 199) as part of the Pharmacogenetic Optimization of Anticoagulation Therapy study.[22,23] Study participants in both cohorts were of self-reported African ancestry and maintained stable warfarin dosages (with

various definitions, but most requiring stability for at least three consecutive clinic visits).[6,22,23] The IWPC cohort was provided by IWPC researchers while the UAB cohort was downloaded from the Database of Genotype and Phenotype (dbGaP; dbGaP Study Accession: phs000708.v1.p1). All studies adhered to relevant ethical requirements, including obtaining institutional review board approvals and individual patient informed consent.[6,16-20,22,23]

Variables

We considered all SNPs (both genotyped and imputed, details provided under 'Data sources/measurement') as exposures, with the primary outcome being the stable warfarin dose, as defined in the 'Study design, setting, and participants' section. Additionally, we accounted for five non-genetic predictors of stable warfarin dose: age, sex, weight, target INR range, and simvastatin/amiodarone status. These predictors were chosen based on expert guidance, literature review and their availability across all analysed cohorts. Although we previously included country of recruitment (a proxy of underlying population structure) as a clinical predictor of stable warfarin dose,[16] it was omitted in this analysis. This decision was made because underlying population structure is better captured by principal components of genetic ancestry, which were available in the GWAS datasets. It is common practice in studies involving mixed ancestry participants to include the first ten principal components to adjust for population stratification.[24] Since some cohorts were admixed, and for consistency in analysis, we included ten principal components of genetic ancestry as additional predictor variables in all cohorts. For analysis on the deconvolved ancestral segments (details provided in the 'Statistical methods' section below), the proportion of specific ancestry was included as a covariate instead of the ten principal components.

Data sources/measurement

Details on the measurement of clinical data and DNA extraction and genotyping procedures are provided in previous reports for the included cohorts.[6,16-20,22,23] Genotyping quality control (QC) and imputation processes were also as previously reported,[18] except this time we did not exclude participants (e.g. mixed-ancestry participants) who did not cluster with the 1000 genomes African populations.[25] Briefly, participants were excluded if they had discordant sex information between clinically reported sex and sex inferred from genetic information, a genotype call rate below 95%, were related to other included participants (based on an identity-by-descent coefficient cut-off of 0.1875 in a pruned subset of uncorrelated SNPs) and had a higher amount of missingness compared to the related participants, and/or displayed extreme heterozygosity (identified from a plot of mean heterozygosity versus proportion of missing genotypes). SNPs were excluded if they had a minor allele frequency (MAF) below 0.01, a Hardy-Weinberg Equilibrium P-value below 0.000001, and/or a

genotype success rate below 95%. QC analysis was performed using PLINK v1.9[26] as previously documented.[18] Following QC, genotype imputation was conducted on the Michigan imputation server (https://imputationserver.sph.umich.edu/index.html#!), using the 1000 Genomes Project phase III reference panel (v5, GRCh37/hg19; 'AFR' population used for all cohorts, except War-PATH cohort 3 that used 'Other/Mixed'),[25] with pre-phasing and imputation carried out using SHAPEIT v2[27] and IMPUTE2[28] software, respectively. Post-imputation QC involved filtering out SNPs with an imputation quality (R^2) below 0.3 or a MAF below 0.01.[18]

Study power

Since we analyzed cohorts with known sample sizes, we computed study power using the power.calc.linear function ("Function to Calculate Power for Linear Models") within the R package "genpwr".[29] With a total sample size of 1,504 (all cohorts) and assuming an additive genetic mode of inheritance, a standard deviation of stable warfarin dose of 2.66 mg/day (taken from Perera et al.[6]), a minimal difference in stable dose requirements of 1 mg/day to be clinically important,[22] and a genome-wide significance threshold of 5×10^{-8}, study power was computed as 17.3%, 78.1%, 97.6%, and 99.8% for SNPs with MAFs of 5%, 10%, 15% and 20% respectively.

Statistical methods

Outcome transformation

We logarithmically transformed stable warfarin dose to achieve a normal distribution and to obtain a proportional/multiplicative scale that is clinically relevant and easy to interpret.[16,30,31]

Handling quantitative predictors

Quantitative predictor variables were neither transformed nor categorized.

Missing data

We used single imputation using the Michigan imputation server[32] (missing genotype data) and the Multivariate Imputation by Chained Equations (MICE) R package[33] (missing weight in War-PATH cohorts 1, 2 and 3) as previously described.[18]

Genome-wide association analysis

For all cohorts and following QC procedures, multivariable linear regression (using SNPTEST version 2)[34] was undertaken assuming an additive mode of inheritance and adjustment for five non-genetic covariates ('Variables' sub-section) and ten principal components of genetic ancestry. For the mixed ancestry cohorts and starting with imputed QC'd genotypes, we also conducted analysis on the

deconvolved ancestral segments to obtain ancestry-specific estimates (the proportion of the specific ancestry per chromosome was adjusted for instead of the principal components). These deconvolved ancestral segments were obtained using the Tractor pipeline[24] (available at https://github.com/eatkinson/Tractor) which includes a precursor step ('local ancestry deconvolution' using RFmix (version v2.03-r0), https://github.com/slowkoni/rfmix/blob/master/MANUAL.md)[35] that also outputs global ancestry estimates and a step ('extracting tracts and ancestral dosages') that produces Variant Call Format files containing ancestry-specific genotypes. We used the same parameters reported in the Tractor manuscript[24] except that, based on a previous report,[36] we used four (African, East Asian, European and South Asian) 1000 genome[25] reference ancestries (instead of two) for the sub-Saharan mixed-ancestry cohorts. To inspect technical artifacts, local ancestry calls were visualized using karyogram plots produced using publicly available code (https://github.com/armartin/ancestry_pipeline).[37]

Meta-analysis and step-wise conditional analysis

To pool the different GWAS, we undertook standard error-weighted meta-analyses with genomic control correction using METAL (version 2011-03-25).[38] Since conditioning on well recognised loci may help identify novel SNPs,[6] we also undertook stepwise conditional analyses on the meta-analysed cohorts in which we first adjusted for *VKORC1 -1639G>A* (rs9923231, number of *A* alleles), followed by the *CYP2C* cluster SNP rs12777823 (number of *A* alleles) and a composite *CYP2C9* genotype (number of the following alleles*: CYP2C9*2*, **3*, **5*, **6*, **8*, **9*, and **11*). METAL is capable of pooling cohorts of different ethnicities (https://genome.sph.umich.edu/wiki/METAL_Documentation). Given that newer methods like Meta-Regression of Multi-AncEstry Genetic Association (MR-MEGA, v0.2),[39] which uses a principal component approach, may better account for population stratification, we also conducted a secondary meta-analysis using MR-MEGA (using three principal components (the maximum permitted for six cohorts), genomic control correction on input files, and a second genomic control correction on output files). To ensure the generalizability of results, only SNPs present in all six cohorts were included in the meta-analysis.

Downstream analysis

Genes associated with SNPs passing the nominal significance threshold were obtained from the Single Nucleotide Polymorphism Database (dbSNP, https://www.ncbi.nlm.nih.gov/snp/).[40,41] Other downstream analysis (including heuristic fine-mapping to identify lead SNPs within each genomic region and expression quantitative trait loci (eQTL) analysis) were performed using the Functional Mapping and Annotation of Genome-Wide Association Studies (FUMA GWAS) platform

(https://fuma.ctglab.nl/, v1.5.2) that merges several databases and tools to facilitate genetic analysis.[42] Key input parameters were a maximum P-value of lead SNPs of 5×10^{-8}, a maximum P-value cutoff of 1×10^{-5}, r^2 threshold of 0.8 (used to define both independent significant and lead SNPs), 1000 genomes African populations as the reference panel, a minor allele frequency of 0.01, a maximum distance between Linkage disequilibrium (LD) blocks of 250 kb, and eQTL mapping using the Genotype-Tissue Expression (GTEx, v8) database). The statistical genome-wide significance threshold was set at $P < 5 \times 10^{-8}$. All statistical analysis codes are available in a published manuscript[18] or in public repositories (e.g. https://github.com/eatkinson/Tractor, https://github.com/slowkoni/rfmix/blob/master/MANUAL.md, https://github.com/armartin/ancestry_pipeline, https://genomics.ut.ee/en/tools).

Results

Participants

A total of 1,504 participants were included in the analysis, with their demographic and clinical characteristics, stratified by cohort, outlined in Table 1. The mean age of the participants was 52.1 ± 16.4 years, with 65.4% being females. The mean proportions of African ancestry were as follows: 97.1% ± 3.4% in War-PATH cohort 1, 98.6% ± 2.7% in War-PATH cohort 2, 35.1% ± 19.2% in War-PATH cohort 3, 83.1% ± 3.0% in War-PATH cohort 4, 82.6% ± 9.2% in the IWPC cohort, and 83.7% ± 7.8% in the UAB cohort. Principal component analysis plots for all six cohorts are displayed in Supplemental Figure 1, while Supplemental Figures 2 and 3, respectively, show karyograms from selected War-PATH cohort 3 (the most diverse cohort) and reference samples.

SNPs

After imputation and quality control, a total of 17268054 (War-PATH cohort 1), 16613526 (War-PATH cohort 2), 14860288 (War-PATH cohort 3), 16854326 (War-PATH cohort 4), 12605058 (IWPC cohort), and 16987173 (UAB cohort) SNPs were included in the GWAS analyses (Supplemental Tables 2–11).

Meta-analyses

In the first meta-analysis (Figure 1A, Supplemental Table 12) that pooled the results obtained from the standard GWAS (n = 1,504) using METAL, 242 SNPs across three loci passed the genome-wide significance threshold and included loci on chromosome 10 (215 SNPs, lead SNP rs58800757, $P = 4.27 \times 10^{-13}$, average MAF = 0.33), chromosome 12 (one SNP, rs3794303, $P = 2.66 \times 10^{-8}$, MAF =

0.12), and chromosome 16 (26 SNPs, lead SNP rs9925964, $P = 9.97 \times 10^{-16}$, MAF = 0.14). When only the African tracts from the mixed-ancestry cohorts were included in the meta-analysis (Figure 1B, Supplemental Table 13), the number of genome-wide significant SNPs slightly decreased to 181 with the number of loci remaining similar: chromosome 10 (173 SNPs, lead SNP rs58800757, $P = 1.84 \times 10^{-11}$, MAF = 0.37), chromosome 16 (7 SNPs, lead SNP rs9925964, $P = 4.69 \times 10^{-9}$, MAF = 0.08), and chromosome 18 (one SNP, rs192807508, $P = 3.02 \times 10^{-8}$, MAF = 0.04). In the secondary meta-analysis that used MR-MEGA (Figure 1C, Supplemental Table 14), 119 SNPs across four loci passed the genome-wide significance threshold and included: chromosome 6 (one SNP, rs55952617, $P = 1.90 \times 10^{-9}$, MAF = 0.12), chromosome 10 (96 SNPs, lead SNP rs34582766, $P = 2.10 \times 10^{-11}$, MAF = 0.27), chromosome 16 (21 SNPs, lead SNP rs9934438, $P = 1.25 \times 10^{-13}$, MAF = 0.09), and chromosome 18 (one SNP, rs17080365, $P = 7.29 \times 10^{-10}$, MAF = 0.04). Corresponding gene-based meta-analyses (Figure 2) revealed only two significant loci on chromosomes 10 and 16, which respectively include the well-established warfarin genes *CYP2C9* (involved in warfarin metabolism) and *VKORC1* (the molecular target of warfarin). Another observation in Figures 1 and 2 is that for a meta-analysis using the standard GWAS, the most statistically significant locus is in chromosome 16, but this changes to chromosome 10 when only African ancestry tracts are pooled. The Quantile-quantile (QQ) plots of both the SNP- and gene-based meta-analyses are shown in Supplemental Figure 4.

Supplemental Table 15 shows MAFs and p-values for the well-established *CYP2C* gene cluster (including *CYP2C9*), *CYP4F2* and *VKORC1* warfarin-related SNPs in the individual cohorts, while Table 2 shows the pooled MAFs, effect-sizes and P-values for the SNPs present in all the six cohorts. The SNPs that passed the genome-wide significance threshold include one in *CYP2C9* (*CYP2C9*8*, rs7900194, lowest $P = 1.63 \times 10^{-9}$), two in *VKORC1* (*VKORC1 1173C>T*, rs9934438, lowest $P = 1.21 \times 10^{-15}$; *VKORC1 -1639G>A*, rs9923231, lowest $P = 1.24 \times 10^{-15}$) and one in the *CYP2C* cluster (rs12777823, lowest $P = 7.94 \times 10^{-11}$). Figure 2 visually presents the MAFs and P-values for *CYP2C9*, highlighting the relative importance of the different star alleles across the six populations. For example, War-PATH cohort 3 with the lowest proportion of African ancestry (35%) and the highest proportion of European ancestry (30%) had the highest *CYP2C9**2 MAF (6%) and the most significant *CYP2C9**2 *P*-value ($P = 0.02$).

To identify potential novel loci, we adjusted our analysis for the well-established chromosome 10 and 16 loci that influence warfarin dose requirements. The results of these adjustments using METAL are shown in Figure 4, Supplemental Figure 5, and Supplemental Tables 16–18. Specifically, after adjusting for 1) *VKORC1 -1639G>A*, 2) *VKORC1 -1639G>A* and rs12777823, and 3) *VKORC1 -1639G>A*, rs12777823, and compound *CYP2C9* genotype [number of the following alleles*: CYP2C9*2*, **3*, **5*, **6*, **8*, **9*, and **11*] loci, we identified two (on chromosomes 2 and 10), one (chromosome 10), and

no significant loci, respectively. Of note, the adjustment for *VKORC1* alone unveiled a locus (Supplemental Figure 6) on chromosome 2 (top SNPs rs116057875, rs115254730, and rs115240773 [in perfect linkage disequilibrium/LD], $P = 3.64 \times 10^{-8}$, *MALL* [mal, T cell differentiation protein like] introns). These SNPs had an effect size (beta) of 0.4053 (standard error 0.0736), which means each mutant allele increased weekly warfarin dose by an average of 50% (using the formula: absolute(exp(beta) - 1) * 100). However, none of these SNPs had significant eQTLs (Supplemental Figure 6). Additionally, in a secondary meta-analysis using MR-MEGA (Supplemental Table 19), this locus was not detected. Instead, three new loci were identified: one on chromosome 13 (three SNPs in the *STK24* [serine/threonine kinase 24] gene), one on chromosome 14 (three intergenic SNPs), and one on chromosome 17 (one SNP in the *TSPOAP1* [TSPO Associated Protein 1] gene). Due to this discrepancy and because METAL does not directly adjust for population stratification, we repeated the METAL meta-analysis using only the five cohorts with at least 80% African ancestry (n = 1371, Supplemental Table 20). The chromosome 2 locus, represented by the intergenic SNP rs114960663 ($P = 4.96 \times 10^{-8}$), was one of the three detected loci. The *MALL* SNPs were an average of 111 kb from this SNP, placing them within the same locus (defined as a maximum distance of 250 kb between LD blocks).

Discussion

Patients of African ancestry are often underrepresented in pharmacogenomic research, primarily due to challenges in study recruitment.[13,18] Additionally, the considerable genetic diversity within these populations necessitates large sample sizes for well-powered studies.[13,43] To address these issues, we conducted a meta-analysis combining data from 1,504 warfarin-treated patients across four African and two African-American cohorts. In this pooled analysis, we identified 242 SNPs that reached genome-wide significance ($P < 5 \times 10^{-8}$). Of these, 99.6% (241/242) were within known loci on chromosomes 10 (*CYP2C9*, involved in warfarin metabolism) and 16 (*VKORC1*, warfarin's molecular target).

To identify potential new genetic loci, we adjusted our analysis for known loci, including *VKORC1 -1639G>A*, rs12777823, and compound *CYP2C9* genotype. Adjusting for *VKORC1* revealed a potential new locus on chromosome 2 within the *MALL* (mal, T cell differentiation protein like) gene. None of the three intronic variants (rs116057875, rs115254730, rs115240773) in this gene were associated with any literature in the National Library of Medicine's National Centre for Biotechnological Information (NCBI) SNP database as of 14th June 2024, nor were they implicated in gene regulation during expression quantitative trait loci (eQTL) analysis. Additionally, these variants were not

reported in the GWAS Catalog, as queried through the FUMA GWAS platform. According to NCBI's gene database, the protein encoded by *MALL* is involved in raft-mediated trafficking in endothelial cells and localizes in glycolipid- and cholesterol-enriched membrane rafts. The integrity of endothelial cells (known to control the haemostatic pathway),[44] is regulated by the caveolin-1 membrane protein.[45] Given that the *MALL* encoded protein interacts with caveolin-1,[46] an indirect influence of *MALL* on warfarin's anticoagulant action is biologically plausible. However, in a secondary meta-analysis using MR-MEGA instead of METAL, the *MALL* locus was not detected. Whereas MR-MEGA uses meta-regression incorporating ancestry principal components to account for population structure in multi-ethnic GWAS,[39] it can systematically detect weaker associations when risk loci confer relatively homogeneous risk across ancestries.[47] This could explain the increase in the P-value from $P = 3.64 \times 10^{-8}$ (METAL meta-analysis; heterogeneity $I^2 = 0\%$) to $P = 4.24 \times 10^{-6}$ (MR-MEGA meta-analysis). This observation was not unique to the *MALL* locus but was also observed for other SNPs such as *VKORC1* -1639G>A (heterogeneity $I^2 = 0\%$ in a standard GWAS METAL meta-analysis, Table 2). In cases where allelic effect heterogeneity exists between cohorts (e.g. rs12777823, which has been reported to influence warfarin dose only in patients of African ancestry),[6] MR-MEGA demonstrates the highest power.[39,47] This is consistent with the decrease in P-value for rs12777823 from $P = 4.86 \times 10^{-9}$ (METAL meta-analysis; heterogeneity $I^2 = 76\%$) to $P = 7.94 \times 10^{-11}$ (MR-MEGA meta-analysis). Given MR-MEGA's limitations and the fact that this locus was still significant when we included only the five cohorts with >80% African ancestry, the influence of the biologically plausible *MALL* SNPs cannot be discounted, warranting further examination. Another issue that questions *MALL*'s significance is why it only became significant after conditioning on the *VKORC1* locus. This highlights the importance of balancing covariate inclusion: adding covariates can reduce power by increasing model complexity, especially with small sample sizes, unless the covariates explain significant variability in the outcome, thus reducing residual variance. In the case of *MALL*, the *P*-value decreases from $P = 6.14 \times 10^{-8}$ to $P = 3.64 \times 10^{-8}$ when adjusting for *VKORC1*. However, stepwise additions of rs12777823 and *CYP2C9* increase the P-values to 7.01×10^{-8} and 2.21×10^{-7} respectively, which may indicate that the variance explained by these covariates is offset by the increased model complexity, resulting in less power. Larger studies or more direct evidence of *MALL*'s influence on warfarin dose are needed to clarify its importance.

Excluding admixed individuals from GWAS studies may be unacceptable from both ethical (may exacerbate health inequalities) and scientific (some cardiovascular disorders are more common in admixed individuals) perspectives.[24,48-50] We therefore included four admixed cohorts in our analysis. Among these, three cohorts (War-PATH cohort 4, IWPC cohort, and UAB cohort) had relatively high mean African ancestry proportions (around 83%), while the fourth cohort (War-PATH cohort 3) had a

more diverse ancestry profile, including African (35%), East Asian (13%), European (30%), and South Asian (22%) ancestries, which is expected in some South African populations.[36] To account for admixture, we used Tractor GWAS to deconvolute ancestral tracts/segments;[24] when only African tracts were pooled, the results were consistent with those of standard GWAS meta-analyses, identifying two main loci associated with *CYP2C9* and *VKORC1*. However, the relative importance of these loci shifted, with *CYP2C9* becoming more significant compared to *VKORC1*. This shift can be attributed to the higher influence of *VKORC1*'s key variant, *VKORC1 -1639G>A*, in European ancestries due to higher minor allele frequencies (MAFs, 0.39 vs 0.05 in the 1000 genomes populations).[25] An examination of Supplemental Table 15 indeed shows that local ancestry deconvolution prior to Tractor GWAS changed the *VKORC1 -1639G>A* MAF from 0.30 to 0.05 (War-PATH cohort 3), 0.12 to 0.08 (War-PATH cohort 4), 0.10 to 0.03 (IWPC cohort) and 0.10 to 0.06 (UAB cohort) which shifted *VKORC1*'s importance. Importantly, despite this shift, *VKORC1* remained clinically relevant in patients with exclusively African ancestry.

It's now recognized that the *CYP2C9* alleles **2* and **3*, more common in European populations (with MAFs of 0.12 and 0.07 respectively) but rarer in African populations (both <0.01),[25] may have less impact in Africans (at population level) compared to other *CYP2C9* alleles like **5*, **6*, **8*, and **11*.[5,51,52] In Figure 3, we illustrate how these alleles vary in importance across the six cohorts. For instance, *CYP2C9*2* and **3* were not assessed in the predominantly African War-PATH cohorts 1 and 2 due to their very low MAFs (<1%). Compared to other cohorts, War-PATH cohort 3 had the highest European ancestry (30%) and lowest African ancestry (35%); in this cohort, both alleles had the highest MAFs (around 0.06), resulting in relatively low P-values (0.02 and 0.01 respectively). The IWPC cohort (17% European ancestry) showed a more significant *CYP2C9*3* P-value (P = 0.003), likely due to its larger sample size (IWPC: 316, War-PATH cohort 3: 133). As recommended by the Association for Molecular Pathology and the College of American Pathologists, genotyping panels for all, including African, populations should at a minimum include the *CYP2C9* tier 1 alleles **2*, **3*, **5*, **6*, **8*, and **11*.[51] Lastly, observed MAFs align with weighted averages of reference/ancestral populations. For instance, the 1000 genomes data indicates *CYP2C9*2* MAFs of 0.008 for Africans, 0.001 for East Asians, 0.124 for Europeans, and 0.035 for South Asians.[25] When weighted by the mean proportion of each ancestry in War-PATH cohort 3 (35% African, 13% East Asian, 30% European, and 22% South Asian), the weighted MAF for *CYP2C9*2* is 0.048, aligning with the reported MAF of 0.06 in this study. Similarly, for *CYP2C9*3*, with MAFs of 0.002 (African), 0.034 (East Asian), 0.073 (European), and 0.109 (South Asian), the weighted average is 0.051, which is comparable to the MAF of 0.056 in our study.

Our study's main limitation lies in its limited statistical power and the lack of a replication cohort. Updated power calculations using the observed warfarin dose standard deviation of 17.8 mg/week (Table 1), a clinically-important effect size of 7 mg/week,[22] and a *P*-value of 5×10^{-8} revealed that with a sample size of 1504 participants, we had 23% power to detect SNPs with MAFs of 5% (86% power for SNPs with MAFs of 10%). The 7 mg/week effect size corresponds to 20% of a 35 mg/week dose, meaning a 50% change in dose (as seen for the *MALL* SNPs) would correspond to an effect size of 17.5 mg/week. This would result in 98% power to detect SNPs with MAFs as low as 1%, given a sample size of 1,504 participants. However, SNPs with large effect sizes are rare, which means our study was underpowered to detect SNPs with moderate effect sizes and low MAFs, which are common in diverse populations. Future studies aiming to identify such SNPs should prioritize larger sample sizes. To try addressing this limitation, we included mixed-ancestry cohorts and utilized the Tractor pipeline, which can enhance power by detecting ancestry-specific signals.[24] However, like MR-MEGA,[39,47] Tractor's power is optimized when allele effect sizes vary by ethnicity.[24,53] As a result, some Tractor analyses yielded less significant results, especially when the effect sizes were consistent across ancestries. Additionally, we did not apply corrections for multiple testing or use a more stringent significance threshold in our Tractor analyses. For instance, the appropriate *P*-value threshold for Tractor associations estimated at 1×10^{-8} for 2-way admixture,[24] which becomes even more conservative with 4-way admixture, potentially leading to some false-positive findings. Furthermore, the imputation quality may have been influenced by the imputation panels used, which included populations from the 1000 Genomes Project representing African or Mixed/Other groups. However, it is worth noting that two of the *MALL* SNPs were directly genotyped in the War-PATH cohorts, and had imputation quality R^2 values of 94% and 97% in the IWPC and UAB cohorts, respectively. Despite these limitations, our study represents one of the first GWAS investigating warfarin dose in sub-Saharan Africa, and the inclusion of diverse populations from both sub-Saharan Africa and North America increases the generalizability of our findings.

In conclusion, we report a meta-analysis of 1,504 warfarin-treated patients of African ancestry across four African and two African American cohorts. This study reinforced the significant roles of *CYP2C9* (involved in warfarin metabolism) and *VKORC1* (the molecular target of warfarin) in influencing warfarin dose requirements. We also identified a new locus (*MALL*), that still requires direct evidence of biological plausibility. This study emphasizes the need for larger warfarin-related pharmacogenetic studies of patients of African ancestry to identify African-specific genetic variants, and ultimately improve the quality of anticoagulation in this understudied patient population.

Acknowledgements

The findings presented in this publication build upon the research conducted in IGA's thesis, titled "Warfarin anticoagulation in patients with cardiovascular disease in sub-Saharan Africa" and available at https://livrepository.liverpool.ac.uk/3134701/. I.G.A. thanks the University of Liverpool for studentship funding support. The authors would like to acknowledge Alison Gummery and Claire Hutchinson for providing administrative support for the War-PATH projects.

This research was funded the Medical Research Council (MR/V033867/1) and the National Institute for Health Research (NIHR) (ref: 16/137/101) using UK aid from the UK Government to support global health research. The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR or the UK government.

Author Contributions

MB, KC, ML, JPM, CS-W, JR, CW, EJZ, ALJ, MP contributed to the design and/or funding of the research. All authors contributed to the acquisition of genotype and/or phenotype data. IGA analyzed the results and made the figures. IGA, ALP, and MP drafted the initial manuscript. All authors participated in critical review of the manuscript.

Conflict of Interest Disclosures

M.P. currently receives partnership funding for the following: MRC Clinical Pharmacology Training Scheme (co-funded by MRC and Roche, UCB, Eli Lilly and Novartis). He has developed an HLA genotyping panel with MC Diagnostics, but does not benefit financially from this. He is part of the IMI Consortium ARDAT (www.ardat.org). C.W. is supported by a Wellcome Clinical Research Career Development Fellowship 222075/Z/20/Z. None of these of funding sources have been used for the current paper. All other authors declared no competing interests for this work.

References

1. Semakula JR, Kisa G, Mouton JP, et al. Anticoagulation in sub-Saharan Africa: Are direct oral anticoagulants the answer? A review of lessons learnt from warfarin. *Br J Clin Pharmacol.* 2021.
2. Mouton JP, Blockman M, Sekaggya-Wiltshire C, et al. Improving anticoagulation in sub-Saharan Africa: What are the challenges and how can we overcome them? *Br J Clin Pharmacol.* 2021.

3. Mouton JP, Njuguna C, Kramer N, et al. Adverse Drug Reactions Causing Admission to Medical Wards: A Cross-Sectional Survey at 4 Hospitals in South Africa. *Medicine (Baltimore).* 2016;95(19):e3437.
4. Asiimwe IG, Zhang EJ, Osanlou R, Jorgensen AL, Pirmohamed M. Warfarin dosing algorithms: A systematic review. *Br J Clin Pharmacol.* 2020.
5. Cavallari LH, Perera MA. The future of warfarin pharmacogenetics in under-represented minority groups. *Future Cardiol.* 2012;8(4):563-576.
6. Perera MA, Cavallari LH, Limdi NA, et al. Genetic variants associated with warfarin dose in African-American individuals: a genome-wide association study. *Lancet.* 2013;382(9894):790-796.
7. Cha PC, Mushiroda T, Takahashi A, et al. Genome-wide association study identifies genetic determinants of warfarin responsiveness for Japanese. *Hum Mol Genet.* 2010;19(23):4735-4744.
8. Cooper GM, Johnson JA, Langaee TY, et al. A genome-wide scan for common genetic variants with a large influence on warfarin maintenance dose. *Blood.* 2008;112(4):1022-1027.
9. Takeuchi F, McGinnis R, Bourgeois S, et al. A genome-wide association study confirms VKORC1, CYP2C9, and CYP4F2 as principal genetic determinants of warfarin dose. *PLoS Genet.* 2009;5(3):e1000433.
10. Parra EJ, Botton MR, Perini JA, et al. Genome-wide association study of warfarin maintenance dose in a Brazilian sample. *Pharmacogenomics.* 2015;16(11):1253-1263.
11. El Rouby N, Shahin MH, Bader L, Khalifa SI, Elewa H. Genomewide association analysis of warfarin dose requirements in Middle Eastern and North African populations. *Clin Transl Sci.* 2021.
12. Spencer CC, Su Z, Donnelly P, Marchini J. Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip. *PLoS Genet.* 2009;5(5):e1000477.
13. Teo YY, Small KS, Kwiatkowski DP. Methodological challenges of genome-wide association analysis in Africa. *Nat Rev Genet.* 2010;11(2):149-160.
14. Kusnoor SV, Villalta-Gil V, Michaels M, et al. Design and implementation of a massive open online course on enhancing the recruitment of minorities in clinical trials - Faster Together. *BMC Med Res Methodol.* 2021;21(1):44.
15. Chaplin M, Kirkham JJ, Dwan K, Sloan DJ, Davies G, Jorgensen AL. STrengthening the Reporting Of Pharmacogenetic Studies: Development of the STROPS guideline. *PLoS Med.* 2020;17(9):e1003344.
16. Asiimwe IG, Waitt C, Sekaggya-Wiltshire C, et al. Developing and Validating a Clinical Warfarin Dose-Initiation Model for Black-African Patients in South Africa and Uganda. *Clin Pharmacol Ther.* 2020.
17. Asiimwe IG, Blockman M, Cohen K, et al. Stable warfarin dose prediction in sub-Saharan African patients: A machine-learning approach and external validation of a clinical dose-initiation algorithm. *CPT Pharmacometrics Syst Pharmacol.* 2021.
18. Asiimwe I, Blockman M, Cohen K, et al. A genome-wide association study of plasma concentrations of warfarin enantiomers and metabolites in sub-Saharan black-African patients. *Frontiers in Pharmacology.* 2022.
19. Ndadza A, Muyambo S, Mntla P, et al. Profiling of warfarin pharmacokinetics-associated genetic variants: Black Africans portray unique genetic markers important for an African specific warfarin pharmacogenetics-dosing algorithm. *Journal of thrombosis and haemostasis : JTH.* 2021;19(12):2957-2973.
20. Ndadza A, Cindi Z, Makambwa E, et al. Warfarin Dose and CYP2C Gene Cluster: An African Ancestral-Specific Variant Is a Strong Predictor of Dose in Black South African Patients. *OMICS: A Journal of Integrative Biology.* 2019;23(1):36-44.

21. Jorgensen AL, Orrell C, Waitt C, et al. A "Bundle of Care" to Improve Anticoagulation Control in Patients Receiving Warfarin in Uganda and South Africa: Protocol for an Implementation Study. *JMIR Res Protoc.* 2023;12:e46710.
22. International Warfarin Pharmacogenetics C, Klein TE, Altman RB, et al. Estimation of the warfarin dose with clinical and pharmacogenetic data. *N Engl J Med.* 2009;360(8):753-764.
23. Limdi NA, Arnett DK, Goldstein JA, et al. Influence of CYP2C9 and VKORC1 on warfarin dose, anticoagulation attainment and maintenance among European-Americans and African-Americans. *Pharmacogenomics.* 2008;9(5):511-526.
24. Atkinson EG, Maihofer AX, Kanai M, et al. Tractor uses local ancestry to enable the inclusion of admixed individuals in GWAS and to boost power. *Nat Genet.* 2021;53(2):195-204.
25. Genomes Project C, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature.* 2015;526(7571):68-74.
26. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007;81(3):559-575.
27. Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. *Nat Methods.* 2011;9(2):179-181.
28. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* 2009;5(6):e1000529.
29. Moore CM, Jacobson SA, Fingerlin TE. Power and Sample Size Calculations for Genetic Association Studies in the Presence of Genetic Model Misspecification. *Hum Hered.* 2019;84(6):256-271.
30. Vittinghoff E, Glidden D, Shiboski S, McCulloch C. *Regression methods in biostatistics.* New York: Springer; 2012.
31. Keene ON. The log transformation is special. *Stat Med.* 1995;14(8):811-819.
32. Das S, Forer L, Schonherr S, et al. Next-generation genotype imputation service and methods. *Nat Genet.* 2016;48(10):1284-1287.
33. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. . *Journal of Statistical Software.* 2011;45(3):1-67.
34. Marchini J, Howie B. Genotype imputation for genome-wide association studies. *Nat Rev Genet.* 2010;11(7):499-511.
35. Maples BK, Gravel S, Kenny EE, Bustamante CD. RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference. *Am J Hum Genet.* 2013;93(2):278-288.
36. Patterson N, Petersen DC, van der Ross RE, et al. Genetic structure of a unique admixed population: implications for medical research. *Hum Mol Genet.* 2010;19(3):411-419.
37. Martin AR, Gignoux CR, Walters RK, et al. Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am J Hum Genet.* 2017;100(4):635-649.
38. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics.* 2010;26(17):2190-2191.
39. Magi R, Horikoshi M, Sofer T, et al. Trans-ethnic meta-regression of genome-wide association studies accounting for ancestry increases power for discovery and improves fine-mapping resolution. *Hum Mol Genet.* 2017;26(18):3639-3650.
40. Kitts A, Sherry S. The Single Nucleotide Polymorphism Database (dbSNP) of Nucleotide Sequence Variation. In: McEntyre J OJ, editors. , ed. *The NCBI Handbook [Internet].* Bethesda (MD): National Center for Biotechnology Information (US); 2011.
41. Smigielski EM, Sirotkin K, Ward M, Sherry ST. dbSNP: a database of single nucleotide polymorphisms. *Nucleic Acids Res.* 2000;28(1):352-355.
42. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun.* 2017;8(1):1826.
43. Tishkoff SA, Reed FA, Friedlaender FR, et al. The genetic structure and history of Africans and African Americans. *Science.* 2009;324(5930):1035-1044.

44. van Hinsbergh VW. Endothelium--role in regulation of coagulation and inflammation. *Semin Immunopathol.* 2012;34(1):93-106.
45. Xu L, Guo R, Xie Y, Ma M, Ye R, Liu X. Caveolae: molecular insights and therapeutic targets for stroke. *Expert Opin Ther Targets.* 2015;19(5):633-650.
46. de Marco MC, Kremer L, Albar JP, et al. BENE, a novel raft-associated protein of the MAL proteolipid family, interacts with caveolin-1 in human endothelial-like ECV304 cells. *J Biol Chem.* 2001;276(25):23009-23017.
47. Ishigaki K, Sakaue S, Terao C, et al. Multi-ancestry genome-wide association analyses identify novel genetic mechanisms in rheumatoid arthritis. *Nat Genet.* 2022;54(11):1640-1651.
48. Benjamin EJ, Muntner P, Alonso A, et al. Heart Disease and Stroke Statistics-2019 Update: A Report From the American Heart Association. *Circulation.* 2019;139(10):e56-e528.
49. Benetos A, Aviv A. Ancestry, Telomere Length, and Atherosclerosis Risk. *Circ Cardiovasc Genet.* 2017;10(3).
50. Asiimwe IG, Pirmohamed M. Ethnic diversity and warfarin pharmacogenomics. *Front Pharmacol.* 2022.
51. Pratt VM, Cavallari LH, Del Tredici AL, et al. Recommendations for Clinical CYP2C9 Genotyping Allele Selection: A Joint Recommendation of the Association for Molecular Pathology and College of American Pathologists. *J Mol Diagn.* 2019.
52. Asiimwe IG, Zhang EJ, Osanlou R, et al. Genetic Factors Influencing Warfarin Dose in Black-African Patients: A Systematic Review and Meta-Analysis. *Clinical Pharmacology and Therapeutics.* 2020;107(6):1420-1433.
53. Hou K, Bhattacharya A, Mester R, Burch KS, Pasaniuc B. On powerful GWAS in admixed populations. *Nat Genet.* 2021;53(12):1631-1633.

Table 1. Clinical/demographic characteristics of the participants included in the analysis.

	War-PATH cohort 1 (N=548)	War-PATH cohort 2 (N=214)	War-PATH cohort 3 (N=133)	War-PATH cohort 4 (N=94)	IWPC cohort (N=316)	UAB cohort (N=199)	Overall (N=1504)
Age (years)							
Mean (SD)	46.4 (15.7)	49.0 (16.0)	60.1 (14.9)	49.3 (17.0)	57.0 (14.7)	58.9 (15.3)	52.1 (16.4)
Median [Q1, Q2]	45.6 [34.8, 57.0]	48.0 [37.0, 60.8]	62.0 [53.0, 71.0]	49.0 [36.0, 63.0]	58.0 [48.0, 67.0]	60.0 [48.0, 72.0]	52.0 [40.0, 64.0]
Weight (kg)*							
Mean (SD)	75.1 (20.0)	73.3 (18.0)	77.1 (18.2)	74.8 (18.6)	93.9 (26.6)	67.3 (4.08)	77.9 (21.6)
Median [Q1, Q2]	71.4 [60.0, 86.0]	70.0 [60.0, 82.0]	74.0 [62.8, 92.0]	72.0 [61.0, 83.8]	91.8 [75.0, 107.0]	67.0 [64.0, 71.0]	72.0 [63.0, 90.0]
Sex							
Female	390 (71.2%)	160 (74.8%)	87 (65.4%)	39 (41.5%)	199 (63.0%)	109 (54.8%)	984 (65.4%)
Male	158 (28.8%)	54 (25.2%)	46 (34.6%)	55 (58.5%)	117 (37.0%)	90 (45.2%)	520 (34.6%)
Ancestry proportions,† mean (SD)							
African	97.1% (3.4%)	98.6% (2.7%)	35.1% (19.2%)	83.1% (3.0%)	82.6% (9.2%)	83.7% (7.8%)	86.1% (19.1%)
East Asian	0.2% (0.2%)	0.1% (0.1%)	13.2% (8.4%)	2.6% (0.8%)	-	-	1.4% (4.5%)
European	2.2% (2.7%)	1.1% (2%)	30.2% (14.2%)	8.8% (1.9%)	17.4% (9.2%)	16.3% (7.8%)	10.0% (11.5%)
South Asian	0.5% (0.8%)	0.3% (0.9%)	21.5% (10.9%)	5.4% (1.8%)	-	-	2.5% (6.9%)
Target INR range‡							
2.0–3.0	357 (65.1%)	163 (76.2%)	82 (61.7%)	87 (92.6%)	285 (90.2%)	199 (100%)	1173 (78.0%)
2.5–3.5	191 (34.9%)	51 (23.8%)	51 (38.3%)	7 (7.4%)	31 (9.8%)	0 (0%)	331 (22.0%)
On Simvastatin/amiodarone							
No	497 (90.7%)	208 (97.2%)	86 (64.7%)	84 (89.4%)	238 (75.3%)	192 (96.5%)	1305 (86.8%)
Yes	51 (9.3%)	6 (2.8%)	47 (35.3%)	10 (10.6%)	78 (24.7%)	7 (3.5%)	199 (13.2%)
Weekly dose (mg)							
Mean (SD)	41.2 (18.5)	36.4 (14.7)	35.2 (15.4)	34.3 (11.9)	45.5 (18.7)	42.4 (18.6)	40.6 (17.8)
Median [Q1, Q2]	35.0 [30.0, 50.0]	35.0 [27.5, 42.5]	32.5 [25.0, 45.0]	32.5 [27.5, 37.5]	42.0 [32.5, 55.1]	40.0 [27.5, 50.0]	35.0 [30.0, 50.0]

INR indicates international normalized ratio; IWPC cohort, International Warfarin Pharmacogenetics Consortium cohort (African Americans recruited from USA); MICE, Multivariate Imputation by Chained Equations; N, sample size; Q1, first quartile; Q2, second quartile; SD, standard deviation; UAB cohort, University of Alabama at Birmingham cohort (African Americans recruited from USA); War-PATH, WARfarin anticoagulation in PATients in Sub-Saharan Africa; War-PATH cohort 4, mixed-ancestry and black-Africans recruited from Uganda and South Africa.

*11 (2.0%), 1 (0.5%), and 7 (5.3%) participants were missing weight information in War-PATH cohorts 1 (black-Africans recruited from Uganda and South Africa), 2 (black-Africans from South Africa and Zimbabwe) and 3 (mixed-ancestry South African participants) respectively, which was singly imputed using the MICE R package.

†Two and four reference populations respectively used for the African American and sub-Saharan African cohorts.

‡Those with heart valve disorders have a higher target range (2.5–3.5) than the rest (2.0–3.0) who mostly include those with atrial fibrillation and venous thromboembolism.

Table 2. Effect-sizes and P-values for the well-established warfarin-related SNPs (*n* = 1,504).

rsID (reference/ alternative alleles)	Common name	CHR	Position*	All cohorts (METAL)						All cohorts (African tracts, METAL)						All cohorts (MR-MEGA)†		
				Common MAF (%)	Beta‡ (SE)	P	Direction§	Het I^2	Change in warfarin dose‖ (%)	Common MAF (%)	Beta‡ (SE)	P	Direction§	Het I^2	Change in warfarin dose‖ (%)	Common MAF (%)	P	Direction‡
rs7900194 (G/A)	*CYP2C9*8*	10	96702066	0.075	-0.421 (0.070)	***1.63E-09***	--+---	36.8	34.4	0.081	-0.365 (0.064)	***9.19E-09***	------	0.0	30.6	0.070	8.62E-08	--+---
rs2256871 (A/G)	*CYP2C9*9*	10	96708974	0.134	-0.141 (0.054)	9.29E-03	------	0.0	13.1	0.137	-0.119 (0.051)	1.93E-02	----+-	0.0	11.2	0.117	1.18E-01	------
rs28371685 (C/T)	*CYP2C9*11*	10	96740981	0.023	-0.393 (0.126)	1.77E-03	------	0.0	32.5	0.028	-0.286 (0.110)	9.64E-03	------	0.0	24.9	0.020	1.16E-02	------
rs12777823 (G/A)	*rs12777823 (G>A)*	10	96405502	0.274	-0.229 (0.039)	***4.86E-09***	--++--	75.7	20.5	0.284	-0.230 (0.037)	***7.75E-10***	--+---	38.7	20.5	0.274	***7.94E-11***	--++--
rs7294 (G/A)	*VKORC1 3730G>A*	16	31102321	0.470	0.148 (0.036)	3.60E-05	++++++	68.8	15.9	0.473	0.089 (0.035)	9.56E-03	+++-++	0.4	9.4	0.471	1.96E-06	++++++
rs2359612 (C/T)	*VKORC1 2255C>T*	16	31103796	0.240	0.191 (0.041)	3.01E-06	+++-++	62.6	21.0	0.207	0.115 (0.042)	6.00E-03	+++-++	0.0	12.2	0.233	8.21E-07	+++-++
rs8050894 (G/C)	*VKORC1 1542G>C*	16	31104509	0.256	-0.211 (0.041)	2.63E-07	------	60.4	19.0	0.229	-0.106 (0.041)	1.03E-02	----+-	0.0	10.0	0.248	1.52E-07	------
rs9934438 (C/T)	*VKORC1 1173C>T*	16	31104878	0.141	-0.495 (0.062)	***1.21E-15***	------	0.0	39.0	0.052	-0.449 (0.080)	***1.56E-08***	------	9.7	36.2	0.093	***1.25E-13***	------
rs9923231 (G/A)	*VKORC1 -1639G>A*	16	31107689	0.141	-0.496 (0.062)	***1.24E-15***	------	0.0	39.1	0.052	-0.450 (0.080)	***1.60E-08***	------	9.4	36.2	0.093	***1.29E-13***	------

Only SNPs present in all six cohorts were pooled during meta-analysis.

Gene names are italicized while p-values passing the genome-wide significance threshold $p < 5 \times 10^{-8}$ are italicized and bolded.

CYP indicates cytochrome P450; Het I^2 = heterogeneity I^2 statistic, MAF, minor allele frequency; MR-MEGA, Meta-Regression of Multi-AncEstry Genetic Association; NA, not applicable; rsID, reference SNP cluster ID; SE, standard error; SNP, single nucleotide polymorphism; *VKORC1*, vitamin K epoxide reductase complex, subunit 1.

*Human assembly GRCh37 (hg19).

†No Beta (SE) and Het I^2 values reported for MR-MEGA.

‡Coefficient of the alternative allele relative to the reference allele.

§Each symbol represents a cohort in the order War-PATH cohort 1, War-PATH cohort 2, War-PATH cohort 3, War-PATH cohort 4, IWPC cohort, and UAB cohort (see Table 1 and Supplemental Table 15 for details). The "+" and "-" directions respectively indicate increased and decreased weekly warfarin doses in the individual included cohorts.

‖Computed from the log-transformed betas using the formula: absolute(exp(beta) - 1) * 100. A corresponding negative (positive) beta means the variant will decrease (increase) weekly warfarin dose by the computed percentage.

Figure Legends

Figure 1. Manhattan plots of the association between SNPs and stable warfarin dose after pooling six African-ancestry cohorts in a meta-analysis (n = 1,504 participants). Individual GWAS analyses were undertaken using logarithm transformed stable warfarin dose, adjusted for age, sex, weight, target INR range, simvastatin/amiodarone status, and either the first 10 principal components of genetic ancestry or the proportion of the specific ancestry per chromosome by frequentist association testing assuming an additive model of inheritance before being pooled using METAL or MR-MEGA. **A.** METAL meta-analysis with standard GWAS (242 GWAS-significant SNPs [three genomic loci] and a genomic inflation factor of 1.023). **B.** METAL meta-analysis with GWAS using African-ancestry tracts (181 GWAS-significant SNPs [three genomic loci] and a genomic inflation factor of 1.027). **C.** MR-MEGA meta-analysis with standard GWAS (119 GWAS-significant SNPs [four genomic loci] and a genomic inflation factor of 1.000). The red horizontal lines represent the genome-wide (5×10^{-8}) significance threshold. GWAS = genome-wide association study, INR = international normalized ratio, MR-MEGA = Meta-Regression of Multi-AncEstry Genetic Association.

Figure 2. Manhattan plots of the association between SNPs and stable warfarin dose after pooling six African-ancestry cohorts in a meta-analysis, gene-based analysis (n = 1,504 participants). Individual GWAS analyses were undertaken using logarithm transformed stable warfarin dose, adjusted for age, sex, weight, target INR range, simvastatin/amiodarone status, and either the first 10 principal components of genetic ancestry or the proportion of the specific ancestry per chromosome by frequentist association testing assuming an additive model of inheritance before being pooled using METAL or MR-MEGA. The red horizontal lines represent the Bonferroni corrected significance thresholds (0.05 divided by the number of protein coding genes) and included 2.7×10^{-6} (18,244 genes), 2.7×10^{-6} (18,231 genes), and 2.7×10^{-6} (18,244 genes) for **A.** METAL meta-analysis with standard GWAS, **B.** METAL meta-analysis with GWAS using African-ancestry tracts, and **C.** MR-MEGA meta-analysis with standard GWAS, respectively. Some top genes are annotated. *CYP2C9/18* = cytochrome P450, family 2, subfamily C, polypeptide 9/18, GWAS = genome-wide association study, INR = international normalized ratio, *KAT8* = lysine acetyltransferase 8, *VKORC1* = vitamin K epoxide reductase complex, subunit 1.

Figure 3. MAFs and P-values for the *CYP2C9* star variants across the six cohorts. The left and right panels respectively show the MAFs and negative log-transformed P-values for the seven *CYP2C9* star variants. SNPs with MAFs < 1% were not analyzed in the respective cohorts. The cohorts are arranged by the proportion of African ancestry: War-PATH cohort 3 comprising 133 mixed-ancestry South African participants (35.1% African ancestry); IWPC cohort comprising 316 African Americans recruited from USA (82.6% African ancestry); War-PATH cohort 4 comprising 94 mixed-ancestry and black-Africans recruited from Uganda and South Africa (83.1% African ancestry); UAB cohort comprising 199 African Americans recruited from USA (83.7% African ancestry); War-PATH cohort 1 comprising 548 black-Africans from Uganda and South Africa (97.1% African ancestry); and, War-PATH cohort 2 comprising 214 black-Africans recruited from South Africa and Zimbabwe (98.6% African ancestry). *CYP2C9* = cytochrome P450, family 2, subfamily C, polypeptide 9, IWPC = International Warfarin Pharmacogenetics Consortium, MAF = minor allele frequency, UAB = University of Alabama at Birmingham, War-PATH = WARfarin anticoagulation in PATients in Sub-Saharan Africa.

Figure 4. Manhattan plots of the association between SNPs and stable warfarin dose after pooling six African-ancestry cohorts in a meta-analysis (n = 1,504) and conditioning for well-established loci. Individual GWAS analyses were undertaken using logarithm transformed stable warfarin dose, adjusted for age, sex, weight, target INR range, simvastatin/amiodarone status, and either the first 10 principal components of genetic ancestry or the proportion of the specific ancestry per chromosome by frequentist association testing assuming an additive model of inheritance before being pooled using METAL. **A.** Adjustment for *VKORC1 -1639 G>A* (genomic inflation factor = 1.023). **B.** Adjustment for *VKORC1* and rs12777823 (genomic inflation factor = 1.022). **C.** Adjustment for *VKORC1*, rs12777823 and *CYP2C9* (genomic inflation factor = 1.020). The red horizontal lines represent the genome-wide (5×10^{-8}) significance thresholds. The top genes (obtained from a FUMA-GWAS [platform] gene-based Manhattan plot) per main loci are annotated. *CYP2C9/18* = cytochrome P450, family 2, subfamily C, polypeptide 9/18, FUMA-GWAS = Functional Mapping and Annotation of Genome-Wide Association Studies, GWAS = genome-wide association study, INR = international normalized ratio, MALL

= mal, T cell differentiation protein like, SNP = single nucleotide polymorphism, *VKORC1* = vitamin K epoxide reductase complex, subunit 1.

Figure 1

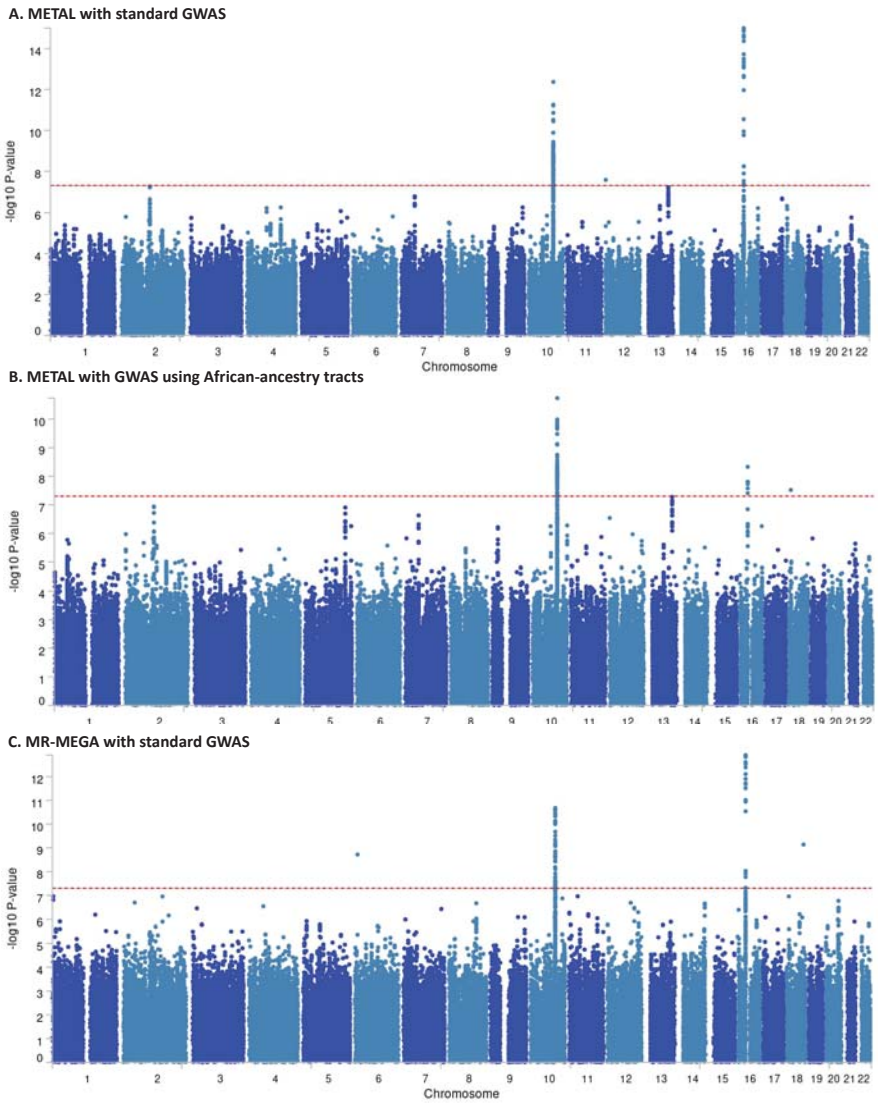

Figure 2

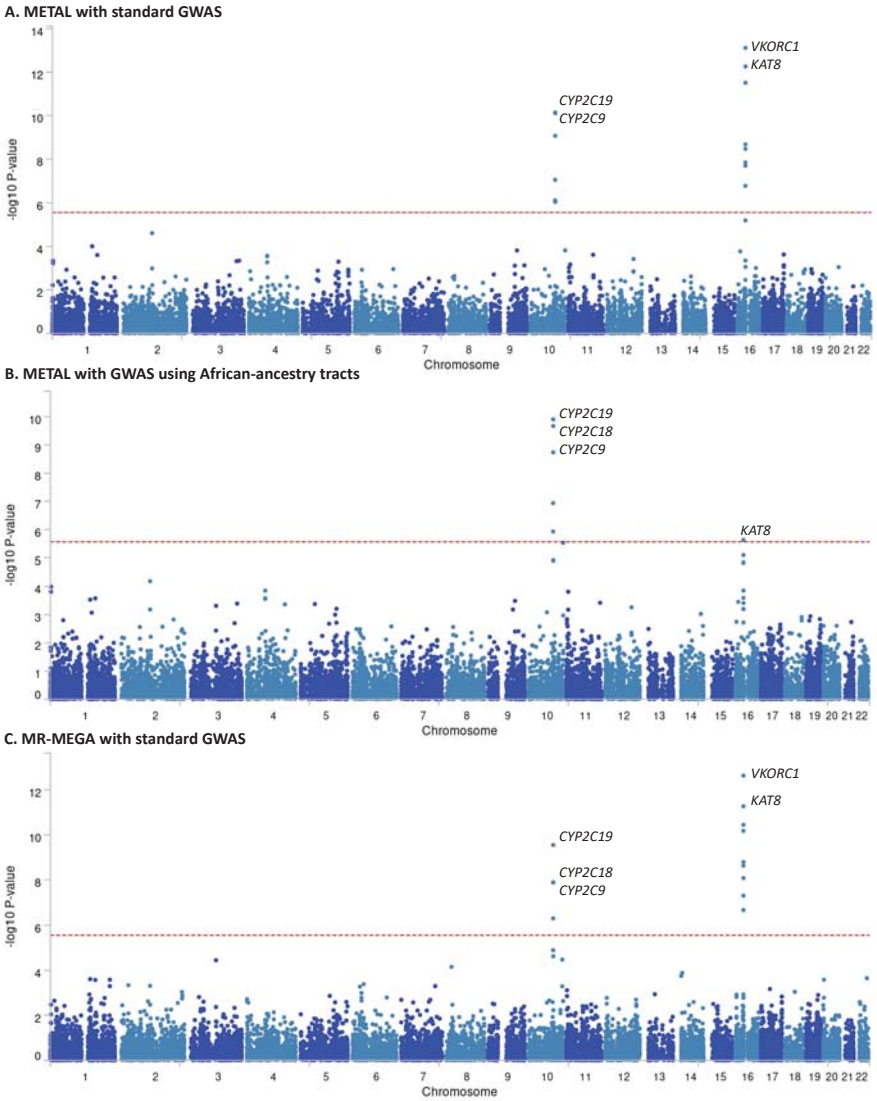

Figure 3

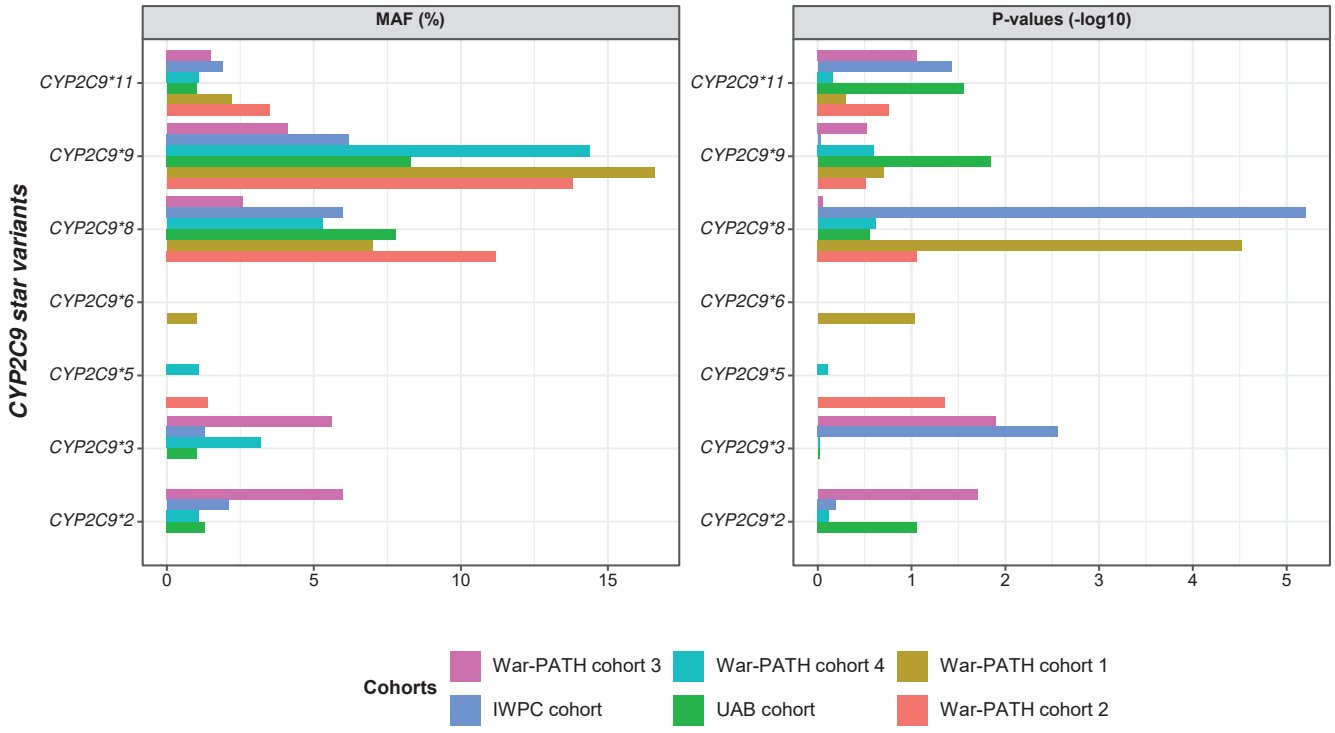

Figure 4