University of Glasgow

Le, Nhu Ngoc (2024) *Prioritising new antihypertensive drug targets and understanding the genetic basis of disease modulation by antihypertensive drugs using Mendelian randomisation.* PhD thesis.

https://theses.gla.ac.uk/84627/

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses
https://theses.gla.ac.uk/
research-enlighten@glasgow.ac.uk

Prioritising New Antihypertensive Drug Targets and Understanding the Genetic Basis of Disease Modulation by Antihypertensive Drugs Using Mendelian Randomisation

Nhu Ngoc Le

BSc and MSc

Thesis submitted in fulfilment for the degree of Doctor of Philosophy (Ph.D)

BHF Cardiovascular Research Centre

School of Cardiovascular and Metabolic Health

University of Glasgow

June 2024

Abstract

Background

Several critical gaps remain in understanding the detailed effects of antihypertensive drugs on cardiovascular, diabetic, and renal outcomes. This thesis aims to address these gaps using the Mendelian randomisation (MR) framework. The first objective is to systematically evaluate the causal effects of genetically predicted systolic blood pressure (SBP) and various antihypertensive drug classes on a range of cardiovascular, diabetic, and renal outcomes. The second objective is to discover if any detected associations are mediated through the transcriptions of genes encoding proteins targeted by the corresponding antihypertensive drug classes.

Methods

Both established and newly developed drugs, as well as first-line and non-first-line therapies, were included in this study. Two-sample MR was conducted to investigate the causal effects of genetically proxied SBP reduction and antihypertensive drug classes on seven outcomes: coronary artery disease (CAD), myocardial infarction (MI), atrial fibrillation (AF), heart failure (HF), ischemic stroke, chronic kidney disease (CKD), and type 2 diabetes (T2D). Summary statistics were obtained from the largest European ancestry genome-wide association studies (GWAS). Sensitivity analyses (MR Egger, weighted median, simple median, negative control outcome) were performed to evaluate the robustness of the MR findings to potential biases. Summary-based MR (SMR), which integrates data from GWAS and eQTL studies, was performed to investigate if the detected drug-outcome associations were mediated through gene transcription. eQTL data were obtained from GTEx v8 for the main analysis, and from the eQTLGen consortium for the sensitivity analysis. To evaluate whether the observed associations were due to a shared causal variant or linkage scenario, the Heterogeneity in Dependent Instruments (HEIDI) test and colocalisation were performed.

Results

Genetically proxied SBP reduction was associated with a lower risk of all the outcomes. The observed associations of genetically proxied calcium channel blockers, angiotensin-converting enzyme inhibitors, and beta-blockers with cardiovascular diseases were consistent with the findings from previous randomised controlled trials and observational studies. New associations were observed between the following genetically predicted drug classes and the outcomes: angiotensinogen inhibition and a decreased risk of CAD and ischemic stroke; loop diuretic and a decreased risk of CAD; endothelin receptor antagonists

and a decreased risk of CAD and ischemic stroke; PDE5 inhibition and a lower risk of CAD, ischemic stroke, and CKD; sGC stimulation and a decreased risk of CAD, MI, and CKD.

SMR analysis found significant associations of genetically increased *GUCY1A3* expression in tibial artery with lower SBP and with a reduced risk of CAD ($p_{SMR} = 1.74\times10^{-06}$; $p_{HEIDI} = 0.664$, $H_4 = 0.99$), genetically increased *PDE5A* expression in aorta with higher SBP and with increased risk of CAD ($p_{SMR} = 9.12\times10^{-06}$, $p_{HEIDI} = 0.338$, $H_4 = 0.48$), genetically increased *KCNH2* expression in the brain cerebellum was associated with higher SBP and increased risk of AF ($p_{SMR} = 6.02\times10^{-05}$, $p_{HEIDI} = 0.195$, $H_3 = 0.99$).

Discussion

The consistency of MR estimates across various methods enhances confidence in the results, contributing to the credibility of genetic approaches in pharmacological research. The SMR method applied in the studies to identify if the effect of a genetic variant on a phenotype is mediated by gene expression provides insights into the biological pathways involved in drug actions and potential targets for therapeutic interventions. Causal variants in *GUCY1A3* and *PDE5A* suggest that these genes play a central role in BP regulation across different tissues, indicating that modulation of these genes' expression could influence cardiovascular outcomes.

Conclusion

This thesis makes a significant contribution to the field by offering new insights and perspectives through the use of Mendelian randomisation and the integration of GWAS and eQTL data using the SMR method providing a more detailed understanding of the genetic mechanisms underlying BP regulation and cardiovascular outcomes. The results from this thesis support the need for further research to validate and functionally characterise these findings followed by clinical effectiveness trials for clinical translation.

Publications arising during this PhD

Papers published

1. **Le NN**, Tran TQB, Lip S, McCallum L, et al. Unravelling the Distinct Effects of Systolic and Diastolic Blood Pressure Using Mendelian Randomisation. Genes. 2022; 13(7):1226. doi:10.3390/genes13071226.
2. Garimella PS, du Toit C, **Le NN**, Padmanabhan S. A genomic deep field view of hypertension. Kidney Int. 2023 Jan;103(1):42-52. doi: 10.1016/j.kint.2022.09.029. Epub 2022 Oct 29.
3. **Le NN**, Tran TQB, du Toit C, Gill D, Padmanabhan S. Establishing plausibility of cardiovascular adverse effects of immunotherapies using Mendelian randomisation. Front Cardiovasc Med. 2023 May 19;10:1116799. doi: 10.3389/fcvm.2023.1116799.
4. du Toit C, Tran TQB, Deo N, Aryal S,…, **Le NN**, et al. Survey and Evaluation of Hypertension Machine Learning Research. Journal of the American Heart Association. 2023;12:e027896.

Published Conference Abstracts

1. **Le NN**, Tran TQB, du Toit C, Gill D, Padmanabhan S. Establishing plausibility of cardiovascular adverse effects of immunotherapies using Mendelian randomisation. Basic Clin Pharmacol Toxicol. 2023;132(S1):3–26. doi: 10.1111/bcpt.13883.
2. **Le NN**, Tran TQB; Bhaskar, R; Du Toit, C; Padmanabhan, S. Combining natural language processing method and Mendelian randomization to establish causal effect of blood pressure on outcomes. Journal of Hypertension 40(Suppl 1):p e77, June 2022. | DOI: 10.1097/01.hjh.0000835948.10184.1a.

Table of Contents

Chapter 1 INTRODUCTION 15

1.1 Hypertension 15

1.1.1 Definition of blood pressure and hypertension 15
1.1.2 Burden of Hypertension 17
1.1.3 Classification of Hypertension 17
1.1.4 Blood pressure measurement 18
1.1.5 Management approaches 20
1.1.6 Pharmacological Classes of Antihypertensive Drugs 22

1.2 Genetics and pharmacogenomics of hypertension 34

1.2.1 Genome-wide association studies 34
1.2.2 Genetics architecture of hypertension 35
1.2.3 Pharmacogenomics 36
1.2.4 Limitations of post-GWAS 37

1.3 Mendelian randomisation studies 40

1.3.1 Limitations of Current Research Methods 40
1.3.2 Mendelian randomisation 42
1.3.3 Data sources for MR studies 46
1.3.4 Selection of genetic instruments 49
1.3.5 Estimation of the causal effect 50
1.3.6 Assumptions underlying MR 51
1.3.7 Sensitivity analysis 56
1.3.8 Interpretation 57
1.3.9 Future of MR 58
1.3.10 Mendelian randomisation analysis in drug development 59
1.3.11 Limitations of MR 66

1.4 Summary-based Mendelian Randomization Analysis 69

1.4.1 About Summary-based Mendelian Randomization 69
1.4.2 Heterogeneity in Dependent Instruments Test 70
1.4.3 Applications of SMR 72
1.4.4 Limitations of SMR 73

1.5 Colocalisation analysis 73

1.5.1 Colocalisation analysis 73
1.5.2 Applications of colocalisation 75
1.5.3 Limitations of colocalisation 76

1.6 Gaps in Knowledge in Hypertension Research that Can Potentially Be Answered by Mendelian Randomisation 77
1.7 Aim of this thesis 78
Chapter 2 DATA AND METHODS 79
2.1 Two-sample Mendelian randomisation 79
2.1.1 Genetic instrument selection 79
2.1.2 Exposure data 80
2.1.3 Outcome data 80
2.1.4 Instrumental variable assumptions 81
2.1.5 Data harmonisation 81
2.1.6 Perform MR 82
2.1.7 Sensitivity analysis 83
2.1.8 Example 1: C reactive protein and coronary heart disease 84
2.1.9 Example 2: C reactive protein and 32 complex somatic and psychiatric outcomes 87
2.2 Summary-based Mendelian randomisation 89
2.2.1 Summary-based Mendelian randomisation 89
2.2.2 Heterogeneity in Dependent Instruments Test 90
2.2.3 Genetic instrument selection 91
2.2.4 Exposure data 92
2.2.5 Outcome data 93
2.2.6 Format dataset for SMR 93
2.2.7 Perform SMR 93
2.2.8 Interpret result 94
2.2.9 Replication analysis 94
2.2.10 Example: Antihypertensive drug classes and psychiatric disorders 94
2.3 Colocalisation analysis 96
2.3.1 Selecting SNPs for the colocalisation analysis 97
2.3.2 Coloc dataset 97
2.3.3 Colocalisation analysis 98
2.3.4 Interpret the result 98
Chapter 3 ANTIHYPERTENSIVE GENES AND CARDIOVASCULAR OUTCOMES 99
3.1 Introduction 99
3.2 Method 102

3.2.1 Study design 102
3.2.2 Identification of Drug Target Genes 103
3.2.3 Genetic Variants Selection 107
3.2.4 Outcome Data 108
3.2.5 Statistical Analysis 109
3.3 Results 111
3.3.1 Mendelian randomization analysis of genetically proxied systolic blood pressure 111
3.3.2 MR analysis of genetically proxied antihypertensive drug classes 115
3.4 Discussion 134
3.4.1 Genetic proxies for BP-lowering 135
3.4.2 Genetic proxies for calcium channel blockers 136
3.4.3 Genetic proxies for beta-blockers 147
3.4.4 Genetically proxies for ACE inhibitors 147
3.4.5 Genetic proxies for AGT inhibition 147
3.4.6 Genetic proxies for loop diuretic 148
3.4.7 Genetically proxies for phosphodiesterase-5 inhibitors 148
3.4.8 Genetic proxies for endothelin receptor antagonists 150
3.4.9 Genetic proxies for sGC stimulators 150
3.4.10 Strengths and limitations 151
3.5 Conclusion 152
Chapter 4 MR integrating GWAS and eQTL data to prioritise genes encoding BP-lowering drug targets for cardiovascular, diabetic, and renal diseases 153
4.1 Introduction 153
4.2 Methods 156
4.2.1 Study design 156
4.2.2 Identification of genes encoding BP-lowering drug targets 158
4.2.3 Identification of genetic instruments for gene expression 158
4.2.4 Outcomes 159
4.2.5 Summary-based Mendelian randomization (SMR) analyses 160
4.2.6 Colocalisation analysis 161
4.3 Results 161
4.3.1 Genetic instrument selection 161
4.3.2 Coronary artery disease 169
4.3.3 Myocardial Infarction 174

4.3.4 Atrial Fibrillation 176
4.3.5 Heart Failure 179
4.3.6 Ischemic stroke 179
4.3.7 Chronic kidney disease 184
4.3.8 Type 2 diabetes 184
4.3.9 Replication analyses 184
4.4 Discussions 190
4.4.1 *GUCY1A3* and CAD risk 193
4.4.2 *PDE5A* and CAD risk 194
4.4.3 *KCNH2* and *AF* risk 195
4.4.4 Suggestive evidence for *CACNB3*, *CACNA1B* and *ACE* gene expression 196
4.4.5 Strengths 198
4.4.6 Limitations 198
4.5 Conclusions 200
Chapter 5 Overall Discussion 201
Chapter 6 Final Conclusions 204
Chapter 7 List of References 205

List of Figures

Figure 1.1 Factors introducing bias to the association between exposure and outcome in observational study. 42
Figure 1.2 Comparison of study designs for observational cohort study, Mendelian randomisation study, and randomised controlled trial. 44
Figure 1.3 Instrumental variable assumptions. 52
Figure 1.4 Figures illustrating different types of pleiotropy. 55
Figure 1.5 Figure illustrating the rationale behind the summary-based Mendelian randomisation analysis. 70
Figure 1.6 Possible explanations for an observed association between the gene expression level of a gene and a phenotype. 71
Figure 1.7 Diagrams illustrating possible situations underlying an overlap between association signals for two different traits at a genetic locus. 74
Figure 3.1 Flowchart illustrating the study design. 104
Figure 3.2 MR estimates for the effect of genetically lower blood pressure through the drug target genes of five drug classes on the risk of the outcomes. 119
Figure 3.3 Drug target Mendelian randomization estimates estimating the association of genetically lower blood pressure through drug target perturbation with the risk of the outcomes. 120
Figure 3.4 MR estimates of the effect of genetically lower SBP generally and genetically lower SBP through drug target perturbation on outcomes in the sensitivity analysis. 121
Figure 3.5 MR estimates for the effect of genetically lower blood pressure through the drug target genes of five drug classes on the risk of the outcomes. 124
Figure 4.1 Study design 157
Figure 4.2 SMR associations of antihypertensive target gene-expression levels with SBP. 162
Figure 4.3 SMR associations of antihypertensive target gene-expression levels with CAD. 170
Figure 4.4 SMR associations of antihypertensive target gene-expression levels with MI... 174
Figure 4.5 SMR associations of antihypertensive target gene-expression levels with AF.. 177
Figure 4.6 SMR associations of antihypertensive target gene-expression levels with HF.. 180
Figure 4.7 SMR associations of antihypertensive target gene-expression levels with ischemic stroke. 182
Figure 4.8 SMR associations of antihypertensive target gene-expression levels with CKD. 185
Figure 4.9 SMR associations of antihypertensive target gene-expression levels with T2D.187

List of Tables

Table 1.1 Classification of hypertension stages. 17
Table 1.2 Comparison of observational study, Mendelian randomisation study, and randomised controlled trial. 44
Table 1.3 Comparison of one-sample MR and two-sample MR. 47
Table 3.1 Antihypertensive Drug Classes, Drug Names, and Corresponding Drug Target Genes. 105
Table 3.2 Genome-wide association studies used in Mendelian randomisation analyses. 109
Table 3.3 MR estimates of the effect of genetically predicted SBP on main outcomes. 112
Table 3.4 MR estimates of the effect of genetically lower SBP on eGFR, HbA1c, and fasting glucose outcomes. 113
Table 3.5 Genetic proxies identified for each antihypertensive drug class. 115
Table 3.6 MR analysis results for genetically proxied drug classes on the main outcomes. 125
Table 3.7 MR analyses on negative control outcomes, using inverse-variant weighted method. 130
Table 3.8 MR analysis results for genetically proxied drug classes on the main outcome, using blood pressure estimates from UKB GWAS. 132
Table 3.9 Difference between MR estimates for genetically proxied drug target perturbation and those for genetically proxied SBP lowering generally. 134
Table 3.10 Comparison of results from my MR study with those from RCT Meta-Analysis and previous MR studies. 137
Table 4.1 SMR associations of antihypertensive target gene-expression levels with SBP. 163
Table 4.2 Colocalisation analysis for SBP and gene expression levels 166
Table 4.3 SMR associations of antihypertensive target gene-expression levels with CAD. 171
Table 4.4 Colocalisation analysis of significant gene-disease associations. 173
Table 4.5 SMR results using summary statistics from eQTLGen Consortium. 173
Table 4.6 SMR associations of antihypertensive target gene-expression levels with MI. 175
Table 4.7 SMR results of drug target gene expression with the risk of AF. 178
Table 4.8 Colocalisation analysis of significant gene-disease associations. 178
Table 4.9 SMR associations of antihypertensive target gene-expression levels with HF. 181
Table 4.10 SMR associations of antihypertensive target gene-expression levels with ischemic stroke. 183
Table 4.11 SMR associations of antihypertensive target gene-expression levels with CKD. 186
Table 4.12 SMR associations of antihypertensive target gene-expression levels with T2D. 188
Table 4.13 Table illustrating the results from MR analysis (Chapter 3) and SMR analysis in this chapter. 191

Acknowledgements

I am deeply grateful to my supervisor, Professor Sandosh Padmanabhan, for his unwavering support and insightful guidance, along with the enriching opportunities and experiences he provided throughout my PhD journey.

I would like to express my profound thanks to Dr John McClure for being the co-supervisor for my PhD project.

I would like to thank Dr Dipender Gill for his expert advice on Mendelian randomisation.

I am sincerely grateful to all members of Professor Padmanabhan's lab for their support and friendship. The collaborative environment and the bonds of friendship we formed have been a source of motivation and inspiration throughout my research journey.

I dedicate this thesis to my parents and my brother. Their unwavering love and support have been my pillar of strength, and I am deeply grateful for their faith in me.

I owe a special word of thanks to my husband and colleague, Tran Quoc Bao Tran, for his love, enduring patience, kindness, understanding, and unwavering support throughout the rigorous journey of this PhD. This PhD journey has been an incredible chapter of my life, not only academically but also personally, as it led me to him and culminated in our marriage.

Author’s Declaration

I declare that this thesis is the result of my original work and has been completed independently by myself. This work was carried out under the supervision of Professor Sandosh Padmanabhan and Dr John McClure. The analysis was advised by Professor Sandosh Padmanabhan and Dr Dipender Gill.

I confirm that this thesis has not been submitted previously for any degree, neither at the University of Glasgow nor at any other institution. I affirm that all sources of information used in this thesis have been thoroughly acknowledged.

Nhu Ngoc Le

June 2024

List of Abbreviations

2SLS	Two-stage least-squares
Aao	Artery aorta
ABPM	Ambulatory blood pressure measurement
ACE	Angiotensin-converting enzyme
ACLY	ATP citrate lyase
ADRA2A	Adrenoceptor alpha 2A
ADRB1	Adrenoceptor beta 1
AdrG	Adrenal Gland
AF	Atrial fibrillation
AGT	Angiotensinogen inhibitor
AGTR1	Angiotensin II receptor type 1
ALLHAT	The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial
ARB	Angiotensin II receptor blockers
Asc	Adipose Subcutaneous
AT1	Angiotensin II receptor type 1
Atib	Artery tibial
AVO	Adipose Visceral Omentum
BC	Brain Cerebellum
Bcaud	Brain Caudate basal ganglia
BCH	Brain Cerebellar Hemisphere
BCor	Brain Cortex
BFCB9	Brain Frontal Cortex BA9
BMI	Body mass index
Bnacc	Brain Nucleus accumbens basal ganglia
BNF	British National Formulary
BP	Blood pressure
Bput	Brain Putamen basal ganglia
CACNA1B	Calcium voltage-gated channel subunit alpha1 B
CACNA1H	Calcium voltage-gated channel subunit alpha1 H
CACNB3	Calcium voltage-gated channel auxiliary subunit beta 3
CAD	Coronary artery disease
CCB	Calcium channel blockers
cGMP	3',5'-cyclic guanosine monophosphate
CI	Confidence interval
CKD	Chronic kidney disease
CRP	C-reactive protein
CVD	Cardiovascular disease
CYP11B1	Cytochrome P450 family 11 subfamily B member 1
CYP11B2	Cytochrome P450 family 11 subfamily B member 2
CYP19A1	Cytochrome P450 family 19 subfamily A member 1
DALY	Disability-adjusted life-year
DBP	Diastolic blood pressure
EDN1	Endothelin 1
EDNRA	Endothelin receptor type A
eGFR	Estimated glomerular filtration rate
eGFRcys	Glomerular filtration rate estimated by cystatin C
eQTL	Expression quantitative trait loci
ERA	Endothelin receptor antagonist
ESR1	Estrogen receptor 1
ET-1	Endothelin-1
ETA	Endothelin-A
ETB	Endothelin-B
FTO	Fat mass and obesity-associated
GLP1R	Glucagon-like peptide 1 analogue target
GTEx	Genotype-Tissue Expression
GTP	Guanosine-5′-triphosphate
GUCY1A3	Guanylate cyclase 1 soluble subunit alpha 3
GWAS	Genome-wide association studies

HAA	Heart Atrial Appendage
HbA1c	Glycated haemoglobin
HBPM	Home blood pressure monitoring
HDL	High-density lipoprotein
HEIDI	Heterogeneity in Dependent Instruments
HF	Heart failure
HFrEF	Heart failure with reduced ejection fraction
HLV	Heart Left Ventricle
HMGCR	3-hydroxy-3-methylglutaryl–coenzyme A reductase
ICPP	International Consortium of Blood Pressure
IL6	Interleukin 6
InSIDE	Instrument Strength Independent of Direct Effect
IV	Instrumental variable
IVW	inverse-variant weighted
KC	Kidney Cortex
KCNH2	Potassium voltage-gated channel subfamily H member 2
LD	Linkage disequilibrium
LDL	Low-density lipoprotein
LP-PAL2	Lipoprotein-associated phospholipase A2
MAF	Minor allele frequency
MI	Myocardial infarction
MR	Mendelian randomization
MSk	Muscle Skeletal
NICE	National institute for health and care excellence
NO	Nitric oxide
NPR3	Natriuretic peptide receptor 3
NR3C2	Nuclear receptor subfamily 3 group C member 2
Ntib	Nerve Tibial
NT-proBNP	N-terminal pro b-type natriuretic peptide
OPBM	Office based BP measurement
OR	Odds ratio
Panc	Pancreas
PCSK9	Proprotein convertase subtilisin/kexin type 9
PDE5	Phosphodiesterase type 5
PDE5A	Phosphodiesterase 5A
PHACTR1	Phosphatase and actin regulator 1
PheWAS	Phenome-wide association studies
PPARG	Peroxisome proliferator activated receptor gamma
PRECISION	Prospective Evaluation of Celecoxib Integrated Safety versus Ibuprofen or Naproxen
PRS	Polygenic risk scores
PSD	Potassium-sparing diuretic
PTGER1	Prostaglandin E receptor 1
RAAS	Renin-angiotensin-aldosterone system
RAS	Renin–angiotensin system
RCT	Randomised controlled trial
RR	Risk ratio
SBP	Systolic blood pressure
SCNN1D	Sodium channel epithelial 1 subunit delta
sGC	Soluble guanylate cyclase
siRNA	Small interfering RNA
SLC39A8	Solute carrier family 39 member 8
SLGT2	Solute carrier family 5 member 2
SMR	Summary-based Mendelian Randomization
SNP	Single nucleotide polymorphism
T2D	Type 2 diabetes
Thy	Thyroid
TZD	thiazide/thiazide-like diuretic
UKB	UK Biobank
UMOD	Uromodulin
WB	Whole Blood

Chapter 1 INTRODUCTION

1.1 Hypertension

1.1.1 Definition of blood pressure and hypertension

Blood pressure (BP) is the pressure exerted by circulating blood against the walls of blood vessels and plays a vital role in maintaining oxygen and nutrient delivery and cardiovascular homeostasis (1). BP is the product of cardiac output and peripheral vascular resistance (also known as systemic vascular resistance). Cardiac output is the amount of blood pumped by the heart per minute, depending on heart rate and stroke volume. Heart rate is the number of heartbeats per minute; stroke volume is the amount of blood pumped with each heartbeat. Peripheral vascular resistance refers to the resistance to the blood flow within the arteries and arterioles of the peripheral circulation, depending on the diameter of the blood vessels, the viscosity of the blood, and the length of the vessels.

BP is represented by systolic blood pressure (SBP) and diastolic blood pressure (DBP). SBP refers to the pressure exerted by blood against the artery walls during the contraction of the heart muscle, which represents the highest pressure within a cardiac cycle. DBP refers to the pressure exerted by blood against the artery walls while the heart muscle is resting between the contractions, which represents the lowest pressure in each cardiac cycle. On average, SBP increases linearly with age, whereas DBP only increases until the age of 50-60 years, after which it plateaus briefly before mildly declining. This pattern leads to a rise in pulse pressure (difference between SBP and DBP) with age (2). DBP has a stronger association with the onset of coronary heart disease (CHD) in younger individuals, while SBP is a more accurate predictor in individuals over the age of 60 (3). The linear, age-related increase in SBP levels in the general population suggests a harmful cycle related to high BP: rising BP can lead to damage or impairment to the vascular system, which subsequently triggers a detrimental cycle of accumulating vascular injury and rising BP (4).

BP follows a normal distribution in the general population, with a log-linear increase in the risk of cardiovascular disease (CVD) with each 1mmHg increase in BP (5). Hypertension is a chronic medical condition defined using an absolute BP value threshold at the upper end of the BP distribution where the benefits of intervention (lifestyle interventions or drug treatment) exceed the benefits of inaction (6). The threshold has been progressively decreased over the last 50 years (7). According to the National Institute for Clinical Excellence (NICE) guideline, a diagnosis of hypertension is confirmed in people with repeated clinical BP of 140/90 mmHg or above, and ambulatory BP measurement (ABPM)

daytime average or home BP monitoring (HBPM) average of 135/85 mmHg or above. Tests for target organ damage (fundoscopy, urinalysis, renal function, ECG) and a calculation of cardiovascular risk score are recommended to be conducted on patients who are newly diagnosed with hypertension (8). A minority of patients (5% to 10%) with hypertension exhibit underlying renal, adrenal, or monogenic disease as the cause for their increased blood pressure - their condition is labelled “secondary hypertension”. Whereas the remainder have essential or primary hypertension (prevalence of ~27% amongst adults worldwide) in which there is no clear single identifiable cause for their raised blood pressure (9).

The maintenance of normal BP involves a complex network of renal, neural, cardiac, vascular and endocrine mechanisms. Perturbation in any component of these mechanisms, which can derive from behavioural, environmental, or genetic factors or a combination of all, may contribute to the development of essential hypertension (10). This highlights the multifactorial nature of hypertension, which has been demonstrated in the Page model of hypertension (11). The age-related increase in SBP is observed only in industrialized societies, not in non-Westernized, tribal populations. These tribal populations do not have a routine intake of dietary salt and other environmental stressors and exposures that typically contribute to rising BP (12). This highlights the important role of lifestyle and environmental factors, such as salt intake and stress, in increasing BP levels. Other environmental and occupational factors that have a negative impact on BP include obesity, overconsumption of alcohol, smoking, insufficient physical activity, insufficient intake of potassium and calcium, and psychological stressors (13). Apart from behavioural and environmental factors, genetic factors also have a considerable influence on the risk of hypertension. Monogenic BP syndromes have substantial contributions to our understanding of BP regulation, the underlying mechanisms, and therapeutic strategies for hypertension (9). More than 30 genes identified are involved in the dysregulation of BP, with rare variants resulting in monogenic forms of hypertension or hypotension. In addition to the rare variants, over the past two decades, there has been a substantial increase in the number of validated single nucleotide polymorphisms (SNPs) in genome-wide association studies (GWAS) for BP phenotype, with more than 1500 common SNPs associated with BP phenotype. Although translation from GWAS findings into causal mechanisms and clinical applications is very challenging due to the pleiotropic nature of BP SNPs, SNPs from GWAS provide valuable insights into our understanding of the causal association between BP, risk factors and outcomes through Mendelian randomisation (section 1.3). Moreover, the findings from BP GWAS facilitate risk stratification using polygenic risk scores and offer opportunities for precision medicine and drug repositioning (10). Genetics of hypertension will be discussed in detail in section 1.2.

1.1.2 Burden of Hypertension

Hypertension is the leading modifiable cardiovascular risk factor in both developed and developing countries, with nearly 10 million deaths globally per year attributable to hypertension (14). 7.8 million deaths and 143 million disability-adjusted life-years (DALYs) in 2015 were estimated to be associated with an SBP of 140 mmHg or higher while 10.7 million deaths and 211 million DALYs were attributable to an SBP of 110 mmHg or higher (15). Evidence from a meta-analysis of individual data for 1 million adults in 61 observational studies highlights significant associations of increased SBP and DBP with elevated risk of cardiovascular diseases (5). The associations were reported over a wide age range from the age 30 years to beyond 80 years, and across diverse population groups (5, 16). For SBP levels ranging from < 115 mmHg to > 180 mmHg and DBP levels ranging from < 75 mmHg to > 105 mmHg, the risk of CVD is raised in a log-linear pattern. An increase of 20 mmHg in SBP or 10 mmHg in DBP was associated with a twofold increase in mortality from each of stroke, ischemic heart disease, or other vascular diseases (5). The association between hypertension and an increased risk of various diseases, such as myocardial infarction (MI), heart failure (HF), stroke, atrial fibrillation (AF), peripheral artery disease (PAD), abdominal aortic aneurysm, renal failure, and cognitive impairment, is also demonstrated in other observational studies (17, 18). Despite the high prevalence of hypertension and the significant health burden it exerts, nearly half of those affected remain unaware of their condition, emphasising the necessity for effective management strategies to alleviate the resultant health complications.

1.1.3 Classification of Hypertension

The classification of hypertension stages according to NICE guidelines (8) is summarised in **Table 1.1.**

Table 1.1 Classification of hypertension stages

NICE guidelines	Stage 1 hypertension	Clinic BP from 140/90 mmHg to 159/99 mmHg and ABPM or HBPM BP from 135/85 mmHg to 149/94 mmHg
	Stage 2 hypertension	Clinic BP from 160/100 mmHg to 180/120 mmHg and ABPM or HBPM ≥ 150/95 mmHg
	Stage 3 hypertension or severe hypertension	Clinic SBP ≥ 180 mmHg or Clinic DBP ≥ 120 mmHg

1.1.4 Blood pressure measurement

1.1.4.1 Office blood pressure measurement

The accurate measurement of BP is crucial in the diagnosis and management of hypertension. BP is typically measured in units of millimetres of mercury (mmHg). Despite the growing popularity of out-of-office BP measurements, conventional office BP measurement remains the most widely used method for diagnosing and managing hypertension. Office BP measurement (OBPM) is the most extensively researched method for assessing BP. It is the fundamental method for diagnosing hypertension, categorizing hypertension stages, understanding the role of BP as a cardiovascular risk factor, determining the benefits of antihypertensive treatment, and setting the BP thresholds and targets for therapeutic interventions (19). In the clinic or office setting, BP can be measured non-invasively using 2 techniques: the auscultatory or Korotkoff method and the oscillometric method. The accuracy of BP measurements is dependent on the use of standardized techniques and appropriate observer training irrespective of the methods employed.

The auscultatory method

The auscultatory method has been the traditional approach for the measurement of BP, using a device called a sphygmomanometer. The measurement process involves the placement of an inflatable cuff around the bare and outstretched upper arm of the patient, which is subsequently inflated to a pressure sufficient to halt the flow of blood in the brachial artery. Upon gradual deflation of the cuff, blood flow is restored and accompanied by sounds that can be detected using a stethoscope positioned over the brachial artery at the antecubital space. The rate of inflation and deflation can lead to errors in BP measurements (20). The sequence of sounds is characterised as:

- Phase 1: the first appearance of sharp tapping sounds, which is interpreted as SBP reading.
- Phase 2: swishing sounds.
- Phase 3: sounds become regular and intensify.
- Phase 4: sudden muffing of sounds.
- Phase 5: complete disappearance of the sounds, which is interpreted as DBP reading.

The oscillometric method

In the past two decades, there has been significant progress in the development and refinement of the oscillometric technique. Oscillometric devices are commonly used in various settings such as clinics, ambulatory, homes, and hospitals. This method involves the use of software embedded in the device to evaluate the oscillometric waveforms commonly

during the deflation of a BP cuff. BP readings are based on the amplitude of the oscillations detected in the lateral walls of the upper arm. The cuff pressure at the point of maximum oscillation amplitude is used to estimate mean arterial BP. Subsequently, SBP and DBP are estimated from the mean arterial, commonly using fixed ratios of the maximal oscillation amplitude (20, 21).

1.1.4.2 24-hour ambulatory blood pressure measurement

ABPM is performed using a non-invasive, fully automated technique to record BP over an extended period, typically 24 hours. The devices used to measure ABPM are mostly programmable and use the oscillometric technique. With ABPM measurements, multiple BP readings (average daytime, night-time, and 24-hour BP values) are recorded in conditions that reflect the usual environment including daily activities and sleep. During ABPM instructions are given to patients to keep a diary encompassing their daily activities, symptoms, times of meals and medication intake, sleep schedules, and any unusual issues (19). Thus, ABPM provides more comprehensive information on BP than office or home BP measurements. Previous studies showed that BP obtained through ABPM demonstrated a stronger association with and better prediction of hypertension-related target organ damage and clinical cardiovascular outcomes than BP measurements taken in the office (22). Importantly, ABPM is able to detect BP patterns such as sustained, white-coat, masked BP and the dipping status of BP, which cannot be detected by using office or home BP measurements (23).

The frequent use of ABPM is not practical due to its high cost, limited accessibility in primary care settings, requirement of additional clinic visits, and inducing discomfort to some patients, particularly during sleep. The main limitation of ABPM is the lack of outcome-based randomised controlled trials (RCTs) investigating the effects of ABPM-guided treatment in comparison to treatment guided by the OPBM. Currently, there are no directly established BP thresholds and treatment goals based on ABPM. It is recommended to use ABPM for confirming the diagnosis of hypertension subsequent to OBPM readings in some countries such as the United Kingdom (8) and the United States (24), however, currently there are no established BP thresholds and the treatment goals directly based on ABPM. As ABPM does not cause white-coat effect, ABPM readings are typically lower than OBPM readings (25). The threshold for ambulatory hypertension is established as 135 mmHg for daytime average SBP and 85 mmHg for daytime average DBP corresponding to an office BP of 140 mmHg SBP or 90 mmHg DBP, respectively (8). For patients with hypertension who have a white-coat effect or masked hypertension, ABPM or home BP monitoring is recommended in

addition to the OBPM to monitor the response to treatment, with the BP goals being inferred from their correspondence with the target office BP.

1.1.4.3 Home blood pressure monitoring

Although ABPM is suggested as the best out-of-office BP measurement method, home blood pressure monitoring (HBPM) is often a more practical approach in clinical practice as it is better tolerated, more widely available and has a relatively lower cost (26). HBPM refers to individuals having their BP measured at home. Similar to devices used for ABPM, the majority of HBPM is conducted using automated upper arm cuff devices with oscillometric technique applied. The conditions and posture during HBPM should be similar to those specified for OBPM. Some studies showed that HBPM is more reproducible than OBPM, and has better prediction for hypertension-mediated organ damage as well as cardiovascular diseases than OPBM (27, 28). As for ABPM, HBPM is also recommended for the confirmation of hypertension, and for the management of the treatment responses in patients with white-coat or mask hypertension (8). The integration of HBPM with telemonitoring and smartphone applications could provide benefits, particularly in the capacity to digitally store and transmit home BP data, thereby facilitating treatment assessment by healthcare professionals (29).

As most patients with HBPM do not experience a white-coat effect, HBPM values are typically lower than OBPM (25). Similar to ABPM, due to a lack of outcome-based RCTs evaluating the effectiveness of HBPM-guided treatment, the threshold for home hypertension is defined as 135 mmHg for SBP and 85 mmHg for DBP corresponding to an office BP of 140 mmHg SBP or 90 mmHg DBP, respectively. One of the limitations of HBPM is the requirement of patient training. Patients may conduct HBPM using inaccurate devices, which can cause anxiety and excessively frequent measurements. Additionally, HBPM also does not prove the night-time BP evaluation, although novel devices developed to measure BP automatically during sleep (19).

1.1.5 Management approaches

The effectiveness of BP-lowering treatments in preventing CVD is well established (30, 31). A systematic review and meta-analysis of 613,815 participants from 123 RCTs of BP-lowering treatment showed that every 10 mmHg reduction in SBP reduced the risk of major CVD events by 20%, CHD by 17%, stroke by 27%, and HF by 18%, which led to a significant 13% reduction in all-cause mortality in the population studies. Similar reductions in proportional risk were found in trials with either higher or lower mean baseline SBP (30). In a meta-analysis of individual participant-level data from 48 RCTs, a 5 mmHg reduction of SBP

was associated with an approximately 10% decrease in the risk of major cardiovascular events, regardless of previous diagnoses of CVD (32). Therefore, hypertension management is a key component of CVD prevention. Hypertension management has been improved considerably, attributable to the rise in understanding of the risk factors associated with hypertension, enhanced hypertension management from primary care practitioners, and the accessibility of efficacious therapeutic options. There are two main approaches in hypertension management: non-pharmacological and pharmacological approach.

1.1.5.1 Non-pharmacological approach

In all patients with suspected or diagnosed hypertension, lifestyle advice should be offered first and continue to be offered periodically (8). Lifestyle modifications are typically beneficial in mitigating a variety of CVD risk factors, including hypertension. They can be used as either a primary treatment before the start of medication therapy or an adjunct approach to medication therapy (33). Lifestyle interventions include a healthy diet (a diet that is high in fruits, vegetables, low-fat dairy, and low in saturated fats and cholesterol), low sodium intake, regular exercise, maintaining a healthy weight (weight loss for overweight or obese patients), moderate alcohol consumption, discouraging excessive consumption of coffee and other caffeine-rich products, and smoking cessation (8, 24). Each lifestyle modification can result in a reduction of a 4-5 mmHg decrease in SBP and a 2-4 mmHg decrease in DBP; a diet that is low in sodium, saturated fat, and total fat, and rich in fruits, vegetables and grains could reduce SBP by an estimated 11 mmHg (24).

1.1.5.2 Pharmacological approach: BP-lowering agents

According to the latest update of the NICE guidelines, treatment with antihypertensive drugs should be offered in addition to lifestyle advice for patients of any age with persistent stage 2 hypertension. The initiation of antihypertensive drug treatment should be discussed with those who are under 80 years of age and have persistent stage 1 hypertension, with either target organ damage, established CVD, renal disease, diabetes, or an estimated 10-year risk score of cardiovascular of equal to or greater than 10%. The treatment should be considered for adults who are under 60 years of age, have stage 1 hypertension and have an estimated 10-year risk score of less than 10%. Also, treatment should be considered for patients who are over 80 years of age, have stage 1 hypertension, and have clinic BP over 150/90 mmHg. For individuals under the age of 40 diagnosed with hypertension, it is important to consider evaluating secondary causes of hypertension. A more comprehensive evaluation of the long-term advantages and potential risks associated with treatment should be considered. The response to alterations in lifestyle or medication therapy in individuals with hypertension is monitored using clinical BP measurements. The BP targets vary

depending on the patient's age and health conditions. Generally, for patients aged under 80 with hypertension, the clinic BP target is <140/90 mmHg. However, the target is more stringent (<130/80 mmHg) for some specific circumstances, such as when patients have chronic kidney disease (CKD) with albumin to creatinine ratio (ACR) of 70 mg/mmol or more, or when patients have type 1 diabetes with ACR of 70 mg/mmol or more (8).

In the latest update of the NICE guidelines, angiotensin-converting enzyme inhibitors (ACEis), angiotensin receptor blockers (ARBs), calcium-channel blockers (CCBs), and thiazide/thiazide-like diuretics (TZDs) are recommended as first-line agents for the treatment of hypertension. The choice of medication depends on the patients' specific circumstances. Factors that can influence the choice of antihypertensive drugs for initial hypertension treatment include patients' age, ethnic background, presence of certain comorbidities (such as T2D, CKD, or HF), and patients' tolerability to the drugs (8). For example, for patients aged under 55, an ACEi or an ARB is usually the first choice. An ACEi or ARB is also a first choice for patients with T2D. For patients aged 55 or older, or if patients are of Black African or African-Caribbean family, CCBs are usually the first choice. If patients are unable to tolerate ACEi (for example due to dry cough), ARB should be offered as an alternative. Similarly, if patients demonstrate adverse reactions to CCBs, such as the onset of for example due to oedema, TZDs should be offered as an alternative (8).

If the goal of BP treatment cannot be achieved with a single therapeutic agent, a second drug (from the first-line agents) should be added to the initial agent in step 2 treatment. Combination use of medicines from two classes of renin-angiotensin hormone system (RAS) blocking agents (e.g. ACEi and ARB) is not recommended to treat hypertension due to an increased risk of hyperkalaemia, hypotension, and impaired renal function and lack of additional benefit (34, 35). Step 3 treatment combines an ACEi or ARB, a CCB and a TZD at the optimal tolerated doses. If the patient's hypertension remains uncontrolled following step 3 treatment, the patient is regarded as having resistant hypertension after checking adherence and confirming BP readings using ABPM or HBPM. If further treatment is indicated, a fourth antihypertensive drug should be added. Low-dose spironolactone may be most suitable for patients with a potassium level of 4.5 mmol/l or less, whereas an alpha- or beta-blocker may be considered for other patients (8).

1.1.6 Pharmacological Classes of Antihypertensive Drugs

1.1.6.1 Existing Antihypertensive Drug Classes

Angiotensin-converting Enzyme Inhibitors

Angiotensin-converting enzyme (ACE), a membrane-bound enzyme, is located on the surface of endothelial cells. This enzyme is especially prevalent in the lung which has a large surface area of vascular endothelium. The common isoform of ACE is not limited to endothelial cells, but is also located in other vascularized tissues, including heart, brain, striated muscle, and kidney. ACE plays an important role in the renin-angiotensin system (RAS). RAS, or renin-angiotensin-aldosterone system (RAAS), is a hormone system that works synergistically with the sympathetic nervous system. It provokes the release of aldosterone and plays a crucial role in regulating sodium excretion and fluid volume, and systemic vascular resistance. Renin, a proteolytic enzyme that is released by the kidneys, converts angiotensinogen (a plasma globulin made in the liver), to a decapeptide called angiotensin I. Angiotensin I is subsequently converted by ACE to an octapeptide called angiotensin II, which is a potent vasoconstrictor. The vasoconstriction action of angiotensin II is mediated via angiotensin II receptor type 1 (AT1) receptors, which belong to the family of G protein-coupled receptors. Angiotensin II also facilitates the release of noradrenaline, reinforcing sympathetic effects. It stimulates the release of aldosterone from the adrenal cortex, which acts on the kidneys to enhance the reabsorption of sodium and water, resulting in increased blood volume and arterial pressure (36). Angiotensin II can be formed locally in various vascular beds, which allows for local control independent of the formation of angiotensin II in the circulating plasma. ACE also inactivates bradykinin, which is a vasodilator molecule.

ACE inhibitors (ACEi) restrain the formation of angiotensin II and aldosterone, resulting in a reduction in vascular tone and extracellular fluid volume and, which consequently leads to a reduction in BP. Angiotensin-sensitive vascular beds, such as those of the heart, brain, and kidney, are preferential acting sites of ACEi. This specificity may be crucial in maintaining sufficient perfusion of the vital organs, especially when faced with decreased perfusion pressure. Moreover, ACEi lead to the accumulation of bradykinin, which is possibly responsible for dry cough - the commonest persistent adverse effect of ACEi (36). Some commonly used ACEi are captopril, benazepril, enalapril, lisinopril, perindopril, quinapril, ramipril, fosinopril, moexipril, and trandolapril. Our understanding of ACEi is based on extensive RCTs data (37-39), especially in individuals with HF, CAD, and at high cardiovascular risk. Systematic review and meta-analysis of RCTs demonstrated that ACEi reduced the risk of stroke, HF, CHD, and major cardiovascular events (which are a combination of stroke, CHD, and HF) (40). In head-to-head comparisons of various major antihypertensive drugs, ACEi were more effective than other drugs in reducing the risk of CHD and less in reducing the risk of stroke (41). A RCT on 5269 participants without CVD but with impaired fasting glucose levels or impaired glucose tolerance showed that 3-year

administration of ramipril does not lead to a significant reduction in the incidence of diabetes or death; however, it does significantly increase regression to normoglycemia (42).

Angiotensin II Receptor Blockers

Angiotensin II receptor blockers (ARBs), also known as angiotensin II receptor antagonists, are a group of pharmaceuticals that inhibit the AT1. ARBs have a comparable BP-lowering efficacy and protective effect as ACEi on major CV events and mortality outcomes, despite having different mechanism for inhibiting the RAAS and a smaller RCT database (43, 44). ARBs are among the first-line treatment of hypertension, with the lowest discontinuation rate in comparison with all other antihypertensive treatments (45). ARBs are not recommended to be used concurrently with other drugs that suppress the RAAS (e.g. ACEi, renin inhibitors) due to high risk of hypotension, hyperkalemia, and impaired renal function. Some commonly used ARBs are losartan, olmesartan, valsartan, azilsartan, telmisartan, irbesartan, candesartan, and eprosartan.

In clinical practice, ARBs tend to be offered if ACEis are not tolerated (such as because of dry cough) or not responded (36). Both ACEi and ARBs have been shown to reduce albuminuria and the progression of diabetic and non-diabetic CKD more effectively than other BP-lowering drug classes. RAS blockers are the only BP-lowering drugs that have available evidence of a reduction in end-stage renal disease risk (46). Both ACEIs and ARBs appear effective in reducing new AF in small or secondary studies of RCTs (47, 48); however, larger RCTs targeting AF did not result in a reduction in the recurrent AF (49, 50). ACEis and ARBs demonstrated efficacy in the prevention or regression of hypertension-mediated organ damage, including left ventricular hypertrophy and small artery remodelling (7). These drug classes are also recommended for patients with post-myocardial infarction and patients with chronic heart failure with reduced ejection fraction (HFrEF) (8).

Renin Inhibitors

Renin inhibitors prevent the first and rate-limiting step of the RAAS, which is the transformation of angiotensinogen into angiotensin I. This subsequently reduces the formation of angiotensin II, a key facilitator of blood pressure regulation. Aliskiren is the only renin inhibitor on the market currently, although it is not considered a first line of hypertension treatment due to its modest BP-lowering effects and lack of evidence regarding CV benefits (51).

Calcium Channel Blockers

Calcium channel blockers (CCBs) reduce intracellular calcium concentration by impeding the entry of calcium ions into vascular and cardiac cells. This results in relaxation and

vasodilation in arterial smooth muscle, leading to a decrease in arterial pressure. In the heart, CCBs diminish myocardial contractility, decrease sinoatrial nodal activity and atrioventricular conduction. The subsequent reduction in ventricular rate, contractility, and afterload reduces the oxygen requirement of the myocardium, thereby preventing angina. CCBs can broadly be divided into two subtypes: dihydropyridines and non-dihydropyridines. Dihydropyridines, such as amlodipine, nifedipine, and nimodipine, are relatively selective for the vasculature. While non-dihydropyridines, such as diltiazem and verapamil, are more selective for the heart.

CCBs are among first-line drug classes for the treatment of hypertension. CCBs are particularly effective in patients with African ancestry and the general older population (19). In comparison with other major antihypertensive drug classes, CCBs demonstrate similar efficacy in lowering BP, major cardiovascular events, and mortality outcomes. Although the majority of RCTs that demonstrate the effectiveness of CCBs on outcomes have used dihydropyridines, meta-analyses have shown that there is no substantial difference in effectiveness between two CCB subtypes (19). Head-to-head meta-analysis of RCTs has shown that CCBs are superior to other antihypertensive drug classes in reducing the risk of stroke, and to be inferior in lowering HF risk (30, 41). Moreover, CCBs are shown to be more effective than beta-blockers in decelerating the progression of carotid atherosclerosis, and in reducing left ventricular hypertrophy and proteinuria (7).

Diuretics

Diuretics indicated for BP management include loop diuretics, potassium-sparing diuretics, and thiazide diuretics. While these diuretics promote a natriuretic action and inhibit the reabsorption of sodium, they exert effects by blocking different transporters at various parts of the nephron. Loop diuretics (e.g. furosemide, torasemide) inhibit the sodium-potassium-chloride cotransporters (Na^+ - K^+ - $2Cl^-$) expressed in the thicks ascending limb of the loop of Henle. This leads to an increased excretion of water along with sodium, chloride, magnesium, calcium, hydrogen, and potassium ions. Although loop diuretics are not normally used in patients with uncomplicated hypertension, they are recommended for patients with CKD stage 4 and 5 or those with severe fluid retention (19). Potassium-sparing diuretics can be divided into two categories: epithelial sodium channel blockers and aldosterone receptor antagonists. Epithelial sodium channel blockers (e.g. amiloride, triamterene) inhibit the reabsorption of sodium via the epithelial sodium channels in the distal convoluted tubules of the kidneys. They are used as potassium-sparing diuretics in edematous states and facilitate potassium conservation when administered alongside thiazide or loop diuretics for the management of hypertension or HF (52). Aldosterone receptor antagonists (e.g. spironolactone, eplerenone) work by inhibiting the effects of

aldosterone at the mineralocorticoid receptor. While they are established treatments for HFrEF, their utilisation in hypertension management is confined to specific conditions such as hyperaldosteronism or resistant hypertension (53). Thiazides (e.g. hydrochlorothiazide, bendroflumethiazide) and thiazide-like diuretics (e.g. indapamide, chlortalidone) are used more commonly than other diuretics for the treatment of hypertension, and also are among the first-line antihypertensive treatments. Thiazides and thiazide-like diuretics have different chemical structures but exhibit similar therapeutic effects and clinical uses (19). Thiazides block the sodium-chloride co-transporter in the distal convoluted tubule of the nephron. This inhibits the reabsorption of sodium and its osmotically associated water, leading to a decrease in extracellular fluid volume. Over time, compensatory responses such as the activation of the renin-angiotensin system may partially counteract this effect. The long-term antihypertensive effect may be mediated by vasodilation, although the exact mechanism is not fully understood. RCTs and meta-analyses have demonstrated the effectiveness of thiazides in reducing the risk of stroke, CHD, HF and their composites, and cardiovascular mortality (40). In comparison with other major antihypertensives, thiazide or thiazide-like diuretics are superior in preventing HF (30, 41). Thiazide and thiazide-like diuretics can decrease serum potassium levels and have higher discontinuation rates in comparison with RAS blockers. They also can increase insulin resistance and the risk of new-onset diabetes. Previous studies highlighted the role of increasing potassium in the dysmetabolic effects of thiazides and thiazide-like diuretics (54, 55). The combination of thiazides or thiazide-like diuretics with either a potassium-sparing diuretic or an RAS blocker can mitigate the dysmetabolic impacts.

β-blockers

β-blockers are not a homogenous class, consisting of multiple subtypes. One of the main mechanisms of action of β-blockers is that they reduce the force of cardiac contraction and speed of conduction by blocking β_1 adrenoreceptors in the cardiac myocytes. Blocking β_1 adrenoreceptors leads to inhibition of the signal transduction pathway of G_s protein, which results in negative inotropic and chronotropic effects. The BP-lowering effect of β-blockers is complex, likely involving the β_1-mediated inhibition of renin release from the kidneys and subsequent decrease in AT II and aldosterone production.

The effectiveness of β-blockers in reducing the risk of stroke, HF, and major cardiovascular events in patients with hypertension was demonstrated in RCTs and meta-analyses (40). β-blockers are inferior to other antihypertensive drugs (ACEi, ARBs, CCBs, and thiazides) for the prevention of major CV events, stroke, and renal failure (30, 41). β-blockers have been demonstrated to be effective in managing hypertension under certain conditions such as symptomatic angina, for controlling heart rate, post-myocardial infarction, and HFrEF (8). β-

blockers have less favourable side effects than RAS blockers and have a higher rate of treatment discontinuation (45). In individuals predisposed to diabetes, particularly those with metabolic syndrome, the use of β-blockers and diuretics, especially in combination, has been linked to an increased risk of new-onset diabetes (56).

α-blockers

The pharmacological effects of α-blockers are primarily achieved by modifying the sympathetic nervous system. There are two subtypes of α-blockers, namely α_1 and α_2. The majority of the α_1 receptors are located on the vascular smooth muscle. α_1 receptors can be stimulated by catecholamines (epinephrine, norepinephrine), leading to vasoconstriction. α_2 receptors, which are situated on peripheral nerve endings, inhibit the release of norepinephrine when activated. α_1 is the primary receptor in managing hypertension. Non-selective α-blockers (e.g. phentolamine and phenoxybenzamine) induce vasodilation by inhibiting both α_1 and α_2 receptors. Selective α_1-blockers (e.g. doxazosin, prazosin, terazosin and tamsulosin) can cause vasodilation by inhibiting the activation of the α_1 receptors by norepinephrine, leading to a decrease in BP. Additionally, α_1-blockers promote the relaxation of smooth muscle in the prostate, making these drugs beneficial for the management of benign prostatic hyperplasia.

Although doxazosin was equally effective as chlorthalidone in preventing the primary endpoint (fatal CHD or nonfatal MI) in the ALLHAT trial, the doxazosin arm was stopped early as patients receiving doxazosin was associated with a substantial increase in the incidence of congestive heart failure (57). In the PATHWAY-2 trial, doxazosin demonstrated superior efficacy compared to a placebo in lowering BP in resistant hypertension, however, it was not as effective as spironolactone (58). While α-blockers are not typically first-line agents for the management of hypertension, they are established treatments for benign prostatic hyperplasia (58).

1.1.6.2 Emerging antihypertensive classes

Small Interfering RNA

RNA interference is a natural biological mechanism to defend against the invasion of exogenous genes. Small interfering RNA (siRNA) is an RNA interference modality that can knock down the expression of target genes in a sequence-specific manner by facilitating the degradation of a targeted mRNA. siRNA molecules are potential therapeutic agents across a wide range of diseases, including cancer, and viral and bacterial infections. The mechanism of action of siRNA involves complete base pairing with mRNA, while small molecule and monoclonal antibody drugs necessitate the identification of the complicated spatial structure

of particular proteins. The frequency of administration of siRNA treatment can be as infrequent as quarterly or even twice yearly, which is typically not feasible with small molecules and antibodies (59). Zilebesiran, an investigational siRNA therapeutic agent covalently linked to an N-acetylgalactosamine ligand, can bind with high affinity to the hepatic asialoglycoprotein receptor. The binding process leads to a decrease in angiotensinogen mRNA, which subsequently reduces the hepatic production of angiotensinogen. As a result, synthesis of both angiotensin I and II is inhibited, leading to a decrease in BP (60). As angiotensinogen is the only precursor of all angiotensin peptides, the RAAS suppression by using this siRNA therapy can potentially restrict the compensatory activation of angiotensin that is typically associated with ACE inhibition or angiotensin-receptor blockade. Zilebesiran specifically targets the angiotensinogen receptor in the liver, which plays a key role in most of the synthesis of angiotensinogen in the body (61). Although angiotensinogen is also produced by other organs such as brain, kidneys, heart, adrenal glands, and adipose tissue, previous studies on animal models showed that extrahepatic angiotensinogen does not have a significant impact on BP (62). However, further studies need to be conducted to understand whether these results are applicable to humans. Additionally, the preservation of angiotensinogen production in extrahepatic tissues may reduce the off-target effects in kidney and other tissues, which is supported by evidence from preclinical studies (63).

In phase 1 clinical trial of zilebsiran in patients with hypertension, no serious treatment-associated side effects were detected, while the most common treatment-related adverse effects were mild, transient injection-site reactions (60). No instances of hypotension, hyperkalemia, or deterioration in renal function were observed. Serum angiotensinogen levels were decreased rapidly and sustained up to 6 months after single subcutaneous doses of zilebesiran of 200 mg or more. The observed alterations were concomitant with dose-related 24-hour ABPM reductions, which were discernible by week 8 and sustained through the 24^{th} week. The BP changes were attenuated after exposure to high dietary salt intake and were augmented by coadministration with irbesartan (60). In phase 2 RCT of zilebesiran in patients with mild to moderate hypertension, treatment with zilebesiran in various dosages at 3-month or 6-month intervals has been found to significantly reduce 24-hour mean ambulatory SBP at month 3, as compared to placebo (64). The BP reductions associated with a single administration of zilebesiran were sustained until the sixth month, particularly with dosages of 300 mg and 600 mg. This aligns with the reduction in angiotensinogen demonstrated in pharmacodynamic data. These findings endorse further investigation of zilebesiran as a potential treatment approach for hypertensive patients (64).

Non-Steroidal Mineralocorticoid Receptor Antagonists

Aldosterone is the primary endogenous mineralocorticoid. The main action of aldosterone is to enhance the reabsorption of Na^+ by the distal tubules in the kidney and simultaneously increase the K^+ and H^+ excretion. The function of aldosterone is initiated through its interaction with the mineralocorticoid receptor (also known as corticosteroid receptor type I), an intracellular receptor of the nuclear receptor family. An overproduction of mineralocorticoids leads to a pronounced retention of Na^+ and water, resulting in an elevated extracellular fluid volume and sometimes hypokalemia, alkalosis and hypertension (36). Mineralocorticoid receptors demonstrate a comparable affinity for a range of other steroids, such as progesterone, cortisol, and cortisone, at a level that is comparable to that of aldosterone (65). Inappropriate activation of the mineralocorticoid receptor plays an important role in the onset of hypertension, CVD and CKD; inhibition of mineralocorticoid receptors is a clinically effective method to prevent these diseases (66). Spironolactone and the later-developed eplerenone are clinically approved mineralocorticoid receptor antagonists, which are used to treat primary or secondary hyperaldosteronism and used in combination with other drugs to treat resistant hypertension and HF. They are competitive antagonists of aldosterone, with the structure mimicking the molecular structure of progesterone. As they also partially inhibit the function of androgen and progesterone receptors, their uses are limited due to side effects such as inducing hyperkalaemia, gynecomastia, impotence, and amenorrhea. Novel non-steroidal selective mineralocorticoid receptor antagonists have been developed, with esaxerenone approved for the treatment of hypertension in Japan (67). The binding mode of esaxerenone with mineralocorticoid receptors is completely different to spironolactone and eplerenone, with high affinity to mineralocorticoid receptors and high selectivity over other steroid hormone receptors (68).

Esaxerenone was developed as a drug with high selectivity for mineralocorticoid receptors and a more prolonged half-life compared to previous mineralocorticoid receptor antagonists. In phase II clinical trial in patients with essential hypertension, esaxerenone demonstrated effectiveness in primary endpoint, which is the alteration from the baseline for sitting SBP and DBP, with no significant adverse events (67). Phase III clinical trial was conducted to compare the efficacy and safety of esaxerenone with eplerenone in patients with essential hypertension (69). Esaxerenone (2.5 mg/day) was found to decrease the sitting BP to levels that were equivalent to the eplerenone with a recommended dose of 50 mg/day. A decrease in the 24-hour SBP was observed in the esaxerenone arm irrespective of the dosage (2.5 or 5mg/day), whereas BP was significantly lowered by 5 mg/kg compared to eplerenone 50 mg/day. The superior and long-lasting antihypertensive effects of esaxerenone can be explained by its potent antagonism effect on mineralocorticoid receptors and its long half-life (18.6 ± 2.4 hours for esaxerenone, which is substantially longer than 5.0 ± 1.7 hours $t_{1/2}$ of

eplerenone) (70). Additionally, esaxerenone was found to reduce urinary albumin-to-creatinine-ratio, independent of the BP-lowering effect, supporting the potential of using esaxerenone in hypertensive patients with albuminuria (71). However, there is no evidence that esaxerenone associated with a lower risk of hyperkalemia onset compared to previous mineralocorticoid receptors, thus it is important to carefully monitor the levels of serum K^+ (72).

Aldosterone Synthase Inhibitors

Inhibition of aldosterone synthase is another potential approach to suppressing the activity of aldosterone. In addition to its classical genomic effects mediated via mineralocorticoid receptor as a transcription factor, some of the harmful effects of aldosterone may occur via non-genomic pathways which are independent of the mineralocorticoid receptor stimulation (73). Theoretically, the mineralocorticoid receptor antagonists do not inhibit the nongenomic effects of aldosterone. This may cause enhanced activation of sympathetic nervous system, adversely affect glucose homeostasis, and promote vascular contractility and stiffness (74). A previous study showed that a long-term administration of a mineralocorticoid receptor antagonist can lead to increased plasma aldosterone (75). Also, incidences of “aldosterone breakthrough”, aldosterone levels were initially decreased but then reverted to pretreatment levels or higher, were reported in patients using ACEi or ARB (76). The development of aldosterone synthase inhibitors is expected to address these issues, as an alternative approach to mineralocorticoid receptor antagonists (77).

The electrolyte concentrations in the plasma and the activity of the angiotensin II system are the primary factors regulating the synthesis and release of aldosterone. An increase in plasma K^+ concentrations or a decrease in plasma Na^+ concentrations directly stimulates the release of aldosterone from the zona glomerulosa cells in the adrenal cortex. The decrease of plasma Na^+ also triggers the activation of the RAAS, with one of the effects of angiotensin II being to increase the synthesis and release of aldosterone (36). Long-term production of aldosterone is controlled by the regulated transcription of the *CYP11B2* gene that encodes aldosterone synthase. Aldosterone synthase enzyme catalyses the final three rate-limiting hydroxylation steps from progesterone to aldosterone in the inner mitochondrial membrane of adrenal zona glomerulosa cells. The production of aldosterone is closely related to the production of cortisol. The final step of cortisol biosynthesis, which occurs in zona fasciculata of the adrenal cortex, is catalysed by the enzyme 11-β-hydroxylase encoded by the *CYP11B1* gene. Both *CYP11B1* and *CYP11B2* are situated in the short arm of chromosome 8. There is a substantial level of similarity (around 95%) in the amino acid sequences of the aldosterone synthase (*CYP11B2*) and cortisol synthase (*CYP11B1*), causing the major difficulty in the development of a potent inhibitor that selectively targets the aldosterone

synthase (78). The first oral aldosterone inhibitor, osilodrostat, was initially evaluated to treat primary aldosteronism and resistant hypertension (79). However, this drug is currently used in the treatment of hypercortisolism, also known as Cushing's syndrome.

A new-generation aldosterone synthase inhibitors, baxdrostat and lorundrostat, have shown improved potency and selectivity profiles. Baxdrostat provided > 100-fold selectivity for aldosterone synthase in vitro (80). In phase 1 trial, this drug demonstrated dose-depending effects on plasma aldosterone without effects on cortisol, and these effects were sustained over 10 days in healthy individuals. In a recent phase 2 trial in patients with resistant hypertension who were randomly assigned to receive baxdrostat for 12 weeks or a placebo, the treatment arm experienced a significant reduction in SBP compared to the placebo arm, with no serious treatment-related adverse effects or adrenocortical insufficiency (81). Lorundrostat, another selective inhibitor of aldosterone synthase, is also currently under investigation. Recently, a phase 2 clinical trial enrolled patients with uncontrolled BP despite taking 2 or more antihypertensive medications and randomised them to receive a placebo or lorundrostat (74). Following 8 weeks of treatment lorundrostat significantly reduced SBP in patients with suppressed plasma renin and elevated plasma aldosterone. This reduction was similar to BP reduction among patients without suppressed plasma renin. Lorundrostat was well tolerated, with small elevations in serum potassium and reductions in eGFR. There were no record cases of cortisol insufficiency (74). The promising findings for baxdrostat and lorundrostat support further studies to evaluate the potential of these drugs as an option for the treatment of resistant hypertension.

Soluble Guanylate Cyclase Stimulators

Soluble guanylate cyclase (sGC) is an intracellular enzyme expressed in various cell types including endothelial cells, vascular smooth muscle cells, platelets, and cardiomyocytes. sGC is the main target of nitric oxide (NO) in the cardiovascular system, eliciting many of the biological actions of NO via guanosine 3′,5′-cyclic monophosphate (cGMP) (36). The NO-sGC-cGMP signalling plays a pivotal role in the regulation of the cardiovascular, cardiopulmonary, and cardiorenal systems (82). NO is the endogenous activator of sGC. NO synthase (NOS) enzymes are central to the control of NO biosynthesis. NOS enzyme has three isoforms, namely endothelial nitric oxide synthase (eNOS or NOS3), neuronal nitric oxide synthase (nNOS), and an inducible form (iNOS or NOS2). In endothelial cells, NO is synthesized from L-arginine by eNOS, and then promptly diffuses downstream into smooth muscle cells where it combines with the heme-binding domain of sGC, which is the main target enzyme of NO (83). This binding results in the production of the intracellular cGMP, an important 'second messenger'. The consecutive interactions between cGMP and intracellular proteins such as cGMP-dependent protein kinases, cGMP-regulated ion channels, and

phosphodiesterase. One of the most potent effects of cGMP production is stimulating vasodilation. This is achieved through the activation of vasodilator-specific protein and the regulation of myosin light-chain phosphatase, leading to the reduction in the sensitivity of contractile filaments in smooth muscle cells (84). The NO-sGC-cGMP signalling pathway also leads to a variety of other physiological effects including leukocyte infiltration, antifibrotic and anti-inflammatory effects, angiogenic, antiproliferative effects, and platelet function (83). The development of some diseases especially in cardiovascular, cardiopulmonary, and cardiorenal systems has been associated with a deficiency in the bioavailability of NO. The deficiency in the bioavailability of NO results in reduced sGC activity, which in turn leads to a decrease in cGMP production (85).

sGC is increasingly being recognised as a potential treatment target in cardiopulmonary, cardiovascular and cardiorenal diseases. sGC has an additional regulatory site that is not influenced by NO (36). This site is the target for a therapeutic class known as sGC stimulators, with the first-generation including riociguat and vericiguat. sGC stimulators exhibit a dual mechanism, synergistically acting with endogenous NO while also directly stimulating sGC through the binding site independent of NO, leading to increased cGMP formation (86). Riociguat was approved for the treatment of pulmonary hypertension. In clinical trial phase 3 in patients with symptomatic pulmonary arterial hypertension, after 12 weeks riociguat significantly improved exercise capacity (the 6-minute walk distance), and secondary endpoints (change in pulmonary vascular resistance, N-terminal pro–brain natriuretic peptide (NT-proBNP) levels, World Health Organization (WHO) functional class, time to clinical worsening, score on the Borg dyspnoea scale, quality-of-life variables, and safety) (87). Vericiguat is a later developed sGC stimulator, which was approved for the treatment of HF. Vericiguat has a similar mechanism of action to riociguat – directly stimulating sGC through the regulatory site independent of NO, and also sensitizing sGC to endogenous NO (88). Clinical trial phase 3, which was conducted in patients with chronic HFrEF, demonstrated the efficacy of vericiguat in reducing the incidence of death from cardiovascular causes or hospitalization for HF compared to placebo (89). The encouraging outcomes from these two drugs led to further development of other sGC stimulators, such as praliciguat (in treatment of HF and diabetic nephropathy) (90), and olinciguat (in treatment of sickle cell disease and achalasia) (91). Recently, BAY 1165747 (BAY-747), a newly investigational sGC stimulator, have shown potential in the treatment of resistant hypertension (92). BAY-747 is a highly potent sGC stimulator in vitro, with long half-life and low peak-to-trough ratio. In vivo models of malignant and resistant hypertension, BAY-747 was found to induce a dose-related and long-lasting reduction in mean arterial blood pressure. Furthermore, when administered on top of losartan, amlodipine or spironolactone

and even on top of triple combinations of standard antihypertensive medications, BAY-747 demonstrated an enhanced therapeutic effect (92). Further clinical investigations in patients with hypertension are warranted to validate the safety and effectiveness of BAY-747.

Endothelin Receptor Antagonists

Endothelin, a 21-amino acid peptide, is a potent and long-acting vasoconstrictor factor released from endothelial cells. There are three isoforms: ET-1, ET-2, and ET-3. Among these isoforms, ET-1 is the most common and potent endothelin. It is the only endothelin expressed in endothelial cells, and its expression is also observed in numerous other tissues (36). ET-1 plays an important role in vascular tone regulation through the activation of two primary endothelin receptors, ET_A and ET_B, both of which are G protein-coupled receptors. ET_A receptors are predominantly activated by ET-1, while ET_B receptors are activated to a similar extent by each of the three isoforms of endothelin. The expression of mRNA for the ET_A receptor is found in many human tissues such as vascular smooth muscle, heart, lung, and kidney, but not found in endothelium (36). ET_A mediates vasoconstriction, bronchoconstriction, and the secretion of aldosterone. ET_A receptors are the principal mediators of vasoconstriction in the human vasculature, located on the underlying smooth muscle where a minor fraction of ET_B receptors may also exist. The vasoconstrictor responses triggered by ET-1 are generally sustained and long-lasting, with the magnitude in human-isolated coronary arteries being greater than those generated by angiotensin II or other vasoactive agents (93). ET-1-mediated vasoconstriction in isolated human vessels can be fully reversed by ET_A receptor antagonists. Previous studies demonstrated the beneficial effects of ET_A receptor antagonists such as vasodilation (94), BP maintenance (93), and mitigating adverse vascular remodelling (95). ET_B mediates the release of vasodilatory factors such as NO and prostacyclin (PGI_2) in the endothelium, whereas ET_B receptor mediates vasoconstriction in vascular smooth muscle cells (36). Additionally, ET_B receptors undergo endocytosis to beneficially eliminate ET-1 from the circulation, playing an important role in the deactivation of ET-1 in addition to metabolism by neutral endopeptidase (96). ET-1 is also present in the nephrons of the kidney and is thought to regulate sodium and water excretion under physiological conditions (93). This is likely a contributing factor to fluid retention, a primary adverse effect of ET receptor antagonists (97). However, the main mechanism through which ET receptor antagonists induce fluid retention is intricate, given the multitude of biological functions that ET-1 executes in the kidney (93). Emerging strategies combining ET receptor antagonists with other drugs, such as phosphodiesterase-5 inhibitors and sodium-glucose co-transporter-2 antagonists, have the potential to reduce side effects (98).

The first ET-1 receptor antagonist, bosentan, was approved more than twenty years ago (99). Since then, several ET-1 receptor antagonists, such as macitentan (100) and ambrisentan (101) have been approved, although their application has been primarily confined to the treatment of pulmonary arterial hypertension. In the past 5 years, studies on the endothelin system have demonstrated the promising potential of endothelin receptor antagonists in treating a range of CVDs, including resistant hypertension, microvascular angina, and post coronavirus disease (COVID-19) conditions (93). Aprocitentant, a dual antagonist for ET_A and ET_B receptors, is an orally administered drug taken once daily with a half-life of 44 hours. The recent clinical trial, PRECISION (Parallel-group, phase 3 study with aprocitentan in subjects with resistant hypertension), has shown positive results in the safety and efficacy of aprocitentan on patients with resistant hypertension (102). Patients were eligible for randomisation if within 1-year screening phase before randomisation their sitting BP remained uncontrolled despite being established on a maximally tolerated treatment comprising amlodipine, valsartan, and hydrochlorothiazide. 730 eligible patients were initially randomised to placebo, aprocitentan 12.5 mg or aprocitentan 25 mg for 4 weeks. Following this, all patients received aprocitentan 25 mg for the next 32 weeks. They were subsequently re-randomised to aprocitentan 25 mg or placebo for an additional 4 weeks. SBP was significantly reduced in the treatment arm after a 4-week double-blind phase. This BP lowering effect was maintained throughout a 32-week non-placebo-controlled phase with 25 mg aprocitentan. After 4 weeks of withdrawal phase, SBP significantly increased in the placebo arm compared to the aprocitentan arm. The most frequent adverse event was mild-to-moderate oedema or fluid retention, with a rate of 9-18% reported during the 4-week double-blind part (102). This highlights the importance of adequate diuretic therapy, along with the minimisation of nonadherence to combination therapy, to maintain the long-term success of treatment with aprocitentan for patients with resistant hypertension.

1.2 Genetics and pharmacogenomics of hypertension

1.2.1 Genome-wide association studies

Genome-wide association study (GWAS) is an experimental design aiming to detect associations between genetic variants and complex diseases or traits in samples from a population. The process of GWAS involves testing a vast number of genetic variants, ranging from hundreds of thousands to millions, across the genomes of numerous individuals to identify associations between genetic variants and phenotypes. The initial phase of a GWAS involves the determination of the disease or trait under investigation and the selection of a suitable study population. This could be, for instance, cases and controls for a disease, or an unselected sample from the population for a trait. The majority of genetic

variants examined through GWAS are common in the population, with a minor allele frequency (MAF) generally exceeding 1% (103). Genotyping can be conducted using single-nucleotide polymorphism (SNP) arrays in conjunction with imputation or whole-genome sequencing (WGS). Association tests are performed to identify genomic regions associated with the phenotype of interest at a genome-wide significance level (5×10^{-8}). Meta-analysis is frequently utilised to enhance the statistical power for detecting associations (104). It's important to note that causal variants are typically not directly genotyped but are in linkage disequilibrium (LD) with the genotyped SNPs (104). These genotyped SNPs, referred to as tag SNPs, are highly correlated with adjacent SNPs. Tag SNPs act as proxies for extensive genomic regions encompassing unmeasured SNPs. Consequently, functional analyses of genetic variants are often required to transition from statistical association to causal variants and genes.

The widespread application of GWAS has sparked a transformative shift in the research of human genetics and the genetic elements that are fundamental to complex diseases. More than 50,000 genome-wide significant ($p < 5 \times 10^{-8}$) associations have been reported between genetic variants and common diseases and traits (105). The findings from GWAS have facilitated the discovery of the biological foundation of complex diseases, such as the involvement of autophagy in the pathogenesis of Crohn's disease (106) and the role of complement component 4 in the development of schizophrenia (107). Recently, polygenic risk scores (PRS) – quantitative measures of risk summed across various SNPs – hold promise in guiding early intervention or preventive strategies for common diseases (108). Selecting therapeutic targets supported by genetic evidence from GWAS could potentially lead to a two-fold increase in the success rate of clinical development for new medicinal treatments (109).

1.2.2 Genetics architecture of hypertension

The significant role of genetics in BP regulation has been highlighted in observational studies. A significant heritable component to BP was demonstrated in previous research involving families and twins, with a heritability range of 15-40% for clinic SBP, 15-30% for clinic DBP, and 69% for ambulatory SBP and 51% for ambulatory DBP (110, 111). Individuals with a parental or grandparental history of hypertension face an elevated risk of developing hypertension, even after accounting for secular trends and intergenerational differences in lifestyle, behaviour, physical activity, and dietary sodium intake (112).

The genetic architecture of a trait encompasses the entirety of genetic elements and their properties that contribute to the manifestation of a specific phenotype. This includes the number of genetic variants that have impacts on the phenotype, the extent of their effects,

the prevalence of these variants within the population, and their intricate interactions with both other genetic factors and the external environment (113). The genetic architecture of hypertension currently involves more than 30 genes harbouring rare variants associated with monogenic forms of hypertension or hypotension, and over 1,500 single nucleotide variants associated with BP phenotypes in GWAS (114). These variants encompass both low-frequency and common variants, indicating the continuous spectrum of risk alleles affecting BP and consequently hypertension.

Monogenic forms of hypertension or hypotension primarily involve the RAAS and the adrenal glucocorticoid pathway. A smaller proportion of monogenic BP syndromes is attributed to neuroendocrine tumours of the sympathetic and parasympathetic nervous systems. The genetic variants linked to these monogenic syndromes exert significant effects on BP, resulting in deviations of approximately 20–50 mmHg from normal values. Although these variants do not substantially contribute to the overall public health burden of hypertension, their identification has been instrumental in elucidating the key regulatory pathways governing BP (10).

The SNPs identified in GWAS for BP phenotype account for about 27% of the estimated 30–50% heritability of BP phenotype and approximately 5.7% of the phenotypic variance of SBP (115). Unlike monogenic variants, each SNP exerts a small effect on the BP phenotype - approximately 1 mmHg for SBP and 0.5 mmHg for DBP (115). Most BP loci identified in GWAS are common variants (with allele frequencies greater than 1% in the general population) located in non-coding regions (103). Only a small number of these SNPs are in genes known to contain variants for monogenic BP syndromes (10).

1.2.3 Pharmacogenomics

The growing list of genomic variants linked to BP and hypertension offers new opportunities to broaden our understanding of the pathophysiology of hypertension and to revise and update its molecular classification, which will contribute to precision medicine and drug repurposing for hypertension. Genetic studies have pinpointed numerous genes that encode proteins targeted by currently used drugs. For instance, *HMGCR* is the target for statins, *ESR1* is the target for tamoxifen, *CYP19A1* is the target for aromatase inhibitors, *NR3C2* is the target for finerenone, and more recently *PCSK9* is a novel drug target for the management of hypercholesterolemia (116). Gene-drug interactions for the plausible genes mapped to the regions of BP SNPs were extracted in a previous study (115), which demonstrated that all the major antihypertensive drug classes are captured by pharmacogenetic interaction with the BP GWAS loci (10). Furthermore, gene-drug interactions have uncovered a considerable number of drugs that influence BP. Many of

them (e.g. riociguat (117), nesiritide (118), sodium-glucose cotransporter 2 inhibitors (119)) commonly used for other conditions have either decreased or increased BP being known as their secondary or adverse effect (10). These findings highlight the potential of using evidence from genetic studies to drive drug discovery and drug repurposing.

Genetic variants linked to the targets of common antihypertensive drugs can be utilised as a proxy for the effect of the drugs in Mendelian randomization studies (section 1.3). This approach facilitates the identification of potential causal associations of the drug targets with adverse events. For example, beta-blockers and thiazide diuretics have been associated with an increased risk of T2D. ACEi may elevate the risk of schizophrenia (120) and colorectal cancer (121). Furthermore, MR investigation of the drug targets provides opportunities for drug repositioning (122). CCBs have been associated with a reduced risk of stroke, AF, and CAD (122), while ACEi may reduce the risk of T2D (42) (Section 1.3).

Another area of interest is RNA interference, a natural mechanism involving short strands of RNA that lead to targeted gene suppression (123). Zilebesiran, an RNA interference therapeutic that targets hepatic angiotensinogen synthesis, has shown promising results in phase I and phase II trials, sustaining reductions in serum angiotensinogen and BP over six months (60, 64) (Section 1.6.2).

1.2.4 Limitations of post-GWAS

GWAS have yielded significant achievements, evidenced by the statistical association of many thousands of genetic loci with risk for diseases and traits. GWAS have significantly advanced our understanding of complex traits, however, the translation of GWAS findings into causal mechanisms and clinical applications remains limited (124). Indeed, very few SNPs identified through GWAS have been translated into causal mechanisms and clinical applications, leading to potential pathophysiological pathways involved in blood pressure regulation. The identification of SNP rs13333226 (125) in the *UMOD* gene through a GWAS on BP plays an important role in discovering a novel pathway involving the interaction of uromodulin with renal sodium-potassium-chloride cotransporters in the thick ascending limb. Based on the findings from GWAS and animal studies, the repositioning of a loop diuretic within the hypertension management pathway is being explored by an ongoing clinical trial (126). Three SNPs identified from GWAS on BP, namely rs9349379 (located in *PHACTR1*), rs1630736 (in *EDN1*), and rs10305838 (in *EDNRA*), have been linked to the endothelin pathway (115). The minor allele of rs9349379 in *PHACTR1* is associated with a lower risk of hypertension and this SNP has subsequently been shown as a distal regulator of *EDN1* expression (127). Phase II and III studies of aprocitentan (102, 128), a dual endothelin receptor subtype A and B antagonist, demonstrated favourable safety profiles and beneficial

effects of this drug in treating resistant hypertension, which led to the FDA's approval of aprocitentan in the treatment of resistant hypertension. Other potential targets for BP management, which were identified through GWAS for BP, include the *NPR3* locus and *SLC39A8* locus. The former encodes the natriuretic peptide receptor C and contributes to blood pressure elevation via enhanced vascular smooth muscle cell proliferation, angiotensin II-induced calcium flux, and vessel contraction (129). The latter is linked to an Ala391Thr variation, where the BP-raising Ala391 variant is associated with elevated ERK2 phosphorylation, nuclear factor-κB activation, cadmium accumulation, and decreased vascular endothelial cell viability (130).

Despite the identification of a large number of SNPs associated with complex traits or diseases, these variants often explain only a small fraction of heritability. This discrepancy is known as missing heritability, which indicates that the sum of genetic contributions identified by GWAS accounts for only a small proportion of the estimated heritability of a complex trait. Missing heritability can be explained by several factors. Firstly, some SNPs, which do not reach the genome-wide significance level, have small effects contributing to the complex trait. With the expansion of sample sizes and the advancement of methodologies and research designs, it is anticipated that GWAS may be able to elucidate a considerable proportion of heritability in the future (104). Missing heritability can be due to the existence of rare genetic variants (MAF < 0.5%), which are often not sufficiently frequent to be captured by GWAS. Rare variants may exert considerable effects on the diseases, without demonstrating the distinct patterns of Mendelian inheritance. Although these variants are rare, they could collectively contribute to missing heritability. Identifying the contribution of rare variants in heritability is challenging. It often necessitates large sample sizes and specialised study designs to ensure adequate statistical power for the detection of their effects (131). Interactions between genes, as well as between genes and environments are anticipated to account for a proportion of missing heritability. This could particularly apply to diseases where environmental factors significantly contribute to the likelihood of susceptibility (132). Moreover, overestimation of heritability can contribute to the problem of missing heritability. Family or twin studies are often used to estimate the heritability. In twin studies, one of the main assumptions is that the environments experienced by identical twins are no more than the environments experienced by non-identical twins. The violation of this assumption can result in an overestimation of heritability (131).

While thousands of GWAS have been conducted leading to the discovery of thousands of loci for a wide range of traits and diseases, the number of functional follow-up studies is substantially lower (124). A major challenge in post-GWAS lies in linking SNPs to causal genes and their functions, primarily owing to the predominance of GWAS signals in non-

coding regions, the haplotype structure of the genome, and the influence of pervasive pleiotropy.

More than 90% of the genetic variants identified through GWAS are located in non-coding regions, with many of them being situated at a substantial distance from the nearest known gene (133). The non-coding regions contain a variety of regulatory elements that control gene expression, such as promoters and enhancers, each of which can interact with multiple genes in intricate manners, thereby increasing the complexity of interpreting these non-coding variants (134). Also, non-coding variants can influence the traits via a distal regulation of gene expression, in which the gene nearest to the top GWAS signal is not the causal gene (135). The enrichment of causal variants as well as tag SNPs in predicted transcriptional regulatory regions has been demonstrated in previous studies, suggesting that the significant GWAS loci can affect the phenotype by altering the genetic regulation of one or more target genes (136). Understanding these regulatory mechanisms is essential yet challenging, necessitating advanced analytical approaches.

A significant variant-phenotype association detected from GWAS does not specify the causal variant of the phenotype. There are often many co-inherited genetic variants in strong LD with the significant phenotype-associated variant, also known as tag SNPs (137). GWAS SNP arrays consist of tag SNPs that serve as markers for common variants in the genome via LD. These tag SNPs can capture information about the common patterns of variation in a particular region of the genome, comprising a haplotype. Tag SNPs and haplotypes have been exploited in genetic studies, allowing researchers to avoid genotyping SNPs with redundant information. Local correlation of multiple genetic variants due to LD is helpful in the initial detection of a locus that is linked to the phenotype, however, the LD structure itself poses a major challenge in discovering the causal variant. Within a haplotype, the statistical association of the causal variant with the phenotype is nearly indistinguishable when compared to those of tag SNPs (124). As a result, further experimental studies are usually required to ascertain which of the linked variants are functional.

Pleiotropy refers to a phenomenon in genetics where a single genetic locus affects two or more traits. There are two main types of pleiotropy: vertical and horizontal. Vertical pleiotropy, also known as spurious pleiotropy, refers to a phenomenon where a genetic variant affects one trait that subsequently causes perturbations in other traits (138). For example, a genetic variant in the *FTO* locus that affects body mass index (BMI) probably through its impact on caloric intake might also have an additional corresponding effect on other phenotypes such as hypertension, CAD, and T2D due to the causal relationship between higher BMI and the risk of these diseases. Thus, these associations can be considered as vertical pleiotropy. Advanced genetic methods such as Mendelian

randomization have been proposed to conceptualize and quantify vertical pleiotropy (section 1.3). Horizontal pleiotropy, also known as biological pleiotropy, refers to a phenomenon where a single genetic variant directly influences multiple phenotypes (139). Horizontal pleiotropy is pervasive across the human genome, driven by highly polygenic phenotypes. Phenome-wide association studies (PheWAS), in which each genetic variant is analysed for associations with a wide range of phenotypes, have unveiled a widespread pleiotropy in the majority of blood pressure SNPs. The most abundant pleiotropic signals are found in anthropometric traits, comprising height, BMI, and measures of adiposity and visceral fat. Following these are haematological traits, including red blood cell, white blood cell, and platelet measurements (10). Further investigation is required to ascertain whether the associations identified from GWAS indicate causal relationships with BP regulation or are indicative of confounding or reverse causation.

Given the aforementioned challenges, additional analyses are warranted for post-GWAS to translate the discoveries from GWAS into biological insights. Some analyses that have been used in the post-GWAS analysis include re-sequencing and fine-mapping of multi-ethnic or mixed-race populations, functional analyses, evolutionary genetic analyses, and methodological developments (124). Mendelian randomisation is one of the powerful tools in post-GWAS analysis. It is an advanced statistical method that uses genetic variants as proxies to infer causality, enabling the discovery and validation of causal associations suggested by GWAS findings.

1.3 Mendelian randomisation studies

1.3.1 Limitations of Current Research Methods

Understanding the causal effects of a biomarker or behaviour on adverse health conditions is crucial to evidence-based medicine, drug development, and enhanced decision-making in medical practice. The evidence of causal effects of modifiable exposures on disease ideally comes from well-conducted RCTs, which are considered the “gold standard” in epidemiology (140). A double-blinded randomisation is usually used in well-conducted RCTs, which ensures an equitable distribution of confounding factors between the treatment arm (such as receiving an active drug or active intervention) and the control arm (such as receiving a placebo or control study).

However, RCTs are resource-intensive, often limited in scope, and may not always align with ethical guidelines or be feasible in certain contexts (140). Some exposures, such as alcohol consumption, smoking, physical activity, and complex aspects of nutrition, may not be feasible to be investigated using RCTs. As RCTs are expensive and to reduce the cost,

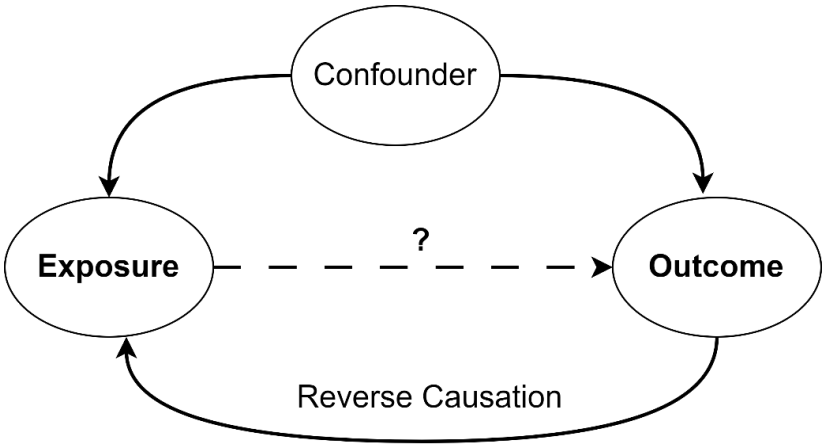
RCTs often focus on specific patient populations such as individuals with a high risk of the disease to enhance the probability of demonstrating the effectiveness of the intervention. This leads to the constraint in extrapolating the findings to the broader and more diverse hypertensive population. Additionally, RCTs typically have limited follow-up periods, restricting their capacity to evaluate the long-term safety and efficacy of drugs. Also, RCTs are designed to test a particular hypothesis about drug efficacy and safety, thus, they may not have sufficient power to uncover rare but clinically adverse effects due to insufficient statistical power. Consequently, treatment guidelines, which frequently rely on extrapolated evidence from RCTs, may not fully capture the potential benefits or adverse effects of drugs. Some antihypertensive drugs or candidates, which have been withdrawn from the market or have their indication restricted due to safety concerns, include aliskiren (141), sitaxentan (142), debrisoquine (143), guanethidine (144), mibefradil (145), reserpine (146), methyldopa (147), doxazosin (148), fenoldopam (149), and omapatrilat (150). On the other hand, some drugs have demonstrated potential for repurposing beyond their original therapeutic indications. For example, sodium-glucose cotransporter 2 (SGLT2) inhibitors are a class of medications developed for the treatment of T2D. SGLT2 inhibitors target the SGLT2 protein in the kidneys, which leads to blocking the reabsorption of glucose in the kidneys and lowering blood glucose levels. In addition to blood glucose control, some SGLT2 inhibitors (such as empagliflozin (151), canagliflozin (152), and dapagliflozin (153)) have shown their effectiveness in lowering BP as well as reducing major cardiovascular events in patients with T2D. These findings suggest that SGLT2 inhibitors can be a novel therapeutic approach for the treatment of hypertension, especially in T2D patients.

Observational epidemiological studies hold several advantages over RCTs, including lower costs, extended study periods, and the inclusion of a more diverse patient population. Observational studies are primarily utilised to discover risk factors and prognostic indicators. They are particularly useful when RCTs would be unethical or not feasible to conduct. However, observational studies are prone to confounding, reverse causation, and measurement error, limiting their reliability. There are some cases that a probably strong causal association was found in observational epidemiological studies but later was rendered implausible by large-scale RCTs. Some important examples include the association between vitamin β-carotene (154) and smoking-related cancers, vitamin E and CHD (155), and between plasma vitamin C levels and CHD mortality (156).

Confounding is often the primary cause of misleading results in observational studies. It refers to a situation where an exposure, which does not have a direct causal effect on disease, is associated with other factors (i.e. confounders) that modify the risk of the disease (**Figure 1.1**). Confounders influence both the exposure and the diseases, introducing biases

and leading to spurious associations between the exposure and the outcome. Some common confounders in observational studies include socioeconomic status, behavioural patterns, environmental factors, or the use of health services (157). Reverse causality, another form of confounding, occurs when the outcome or early signs of the disease that result in the outcome influence the risk factor (**Figure 1.1**). For example, observational studies suggested that lower SBP levels may be associated with higher risks of mortality in the elderly (>75-80 years of age), whereas evidence from a RCT (Systolic Blood Pressure Intervention Trial) showed that lowering SBP to a target of < 120 mmHg was associated with 34% reduction in cardiovascular events and 33% reduction in all-cause mortality in people with greater than 75 years of age (158). This contrast between observational data and RCT data suggests the presence of reverse causality where the outcome (higher mortality) may be affecting the exposure (lower SBP) (159). To improve control for confounding factors, a range of modifications in the design of the study and the analysis are required. Numerous epidemiological methods, such as regression (160), g-computation (161), and propensity score-based methods (162), have been proposed to adjust for confounding factors. However, these methods can only account for known and accurately measured confounders; they are unable to adjust for confounders that are unknown or not measured, which remains a constraint of observational studies. Moreover, measurement errors can occur during the measurement of confounding factors, which leads to residual confounding even after statistical adjustment.

Figure 1.1 Factors introducing bias to the association between exposure and outcome in observational study.



1.3.2 Mendelian randomisation

As discussed in the previous section, RCTs, while the gold standard for assessing drug efficacy, are resource-intensive and often limited in scope, focusing on specific patient populations and short-term outcomes. Observational studies, despite their broader

applicability, are prone to confounding, reverse causation, and measurement errors, limiting their reliability. MR offers a robust alternative to traditional observational studies by using genetic variants as instruments to estimate the causal effect of risk factors or exposure on health outcomes. Under specific assumptions, MR can help to overcome some major limitations of observational studies including bias that could arise from unobserved confounding, measurement error and reverse causality (157) (**Figure 1.2 and Table 1.2).**

The principles of MR are based on Mendel's laws of inheritance and instrumental variable (IV) analysis (157). The first law, the law of segregation, implies that for each location in the autosomal genome, the offspring randomly inherit one allele from the maternal parent and one allele from the paternal parent. The second law, the law of independent assortment, indicates that the inheritance of one trait is independent of the inheritance of other traits. MR exploits the random allocation of genetic variants from parents to offspring at conception as a means of mitigating confounding factors when investigating the association between exposure and disease.

MR is a form of instrumental variable analysis which was developed to control for unmeasured confounders, with the instrumental variable being central to the study design. The concept of using instrumental variables was first proposed in econometrics, by Phillip and Sewall Wright in 1928. It was subsequently referred to as "instrumental variables" by Reiersol in 1941 (163). An instrumental variable is a measurable variable which is associated with the exposure of interest, but not with any other factors or confounders. In MR analysis, genetic variants, which are randomly allocated from parents to offspring, are used as such instrumental variables as they can provide a source of exogenous and unconfounded variation (164). There are important assumptions that a genetic variant should satisfy to be considered as an instrumental variable for a given exposure: 1) the genetic variant is associated with the exposure; 2) the genetic variant is not related to confounders of the exposure-outcome association; 3) the genetic variant does not affect the outcome other than through the exposure. Single nucleotide polymorphisms (SNPs), the most common form of DNA variation among people, are commonly used as IVs in MR analysis. With millions of SNPs in the human genome and the significant advancements of GWAS over the last decades, SNPs have become a rich source of potential instruments for various exposures in MR analysis.

Figure 1.2 Comparison of study designs for observational cohort study, Mendelian randomisation study, and randomised controlled trial.
The exposure is systolic blood pressure, and the outcome is the risk of cardiovascular disease.

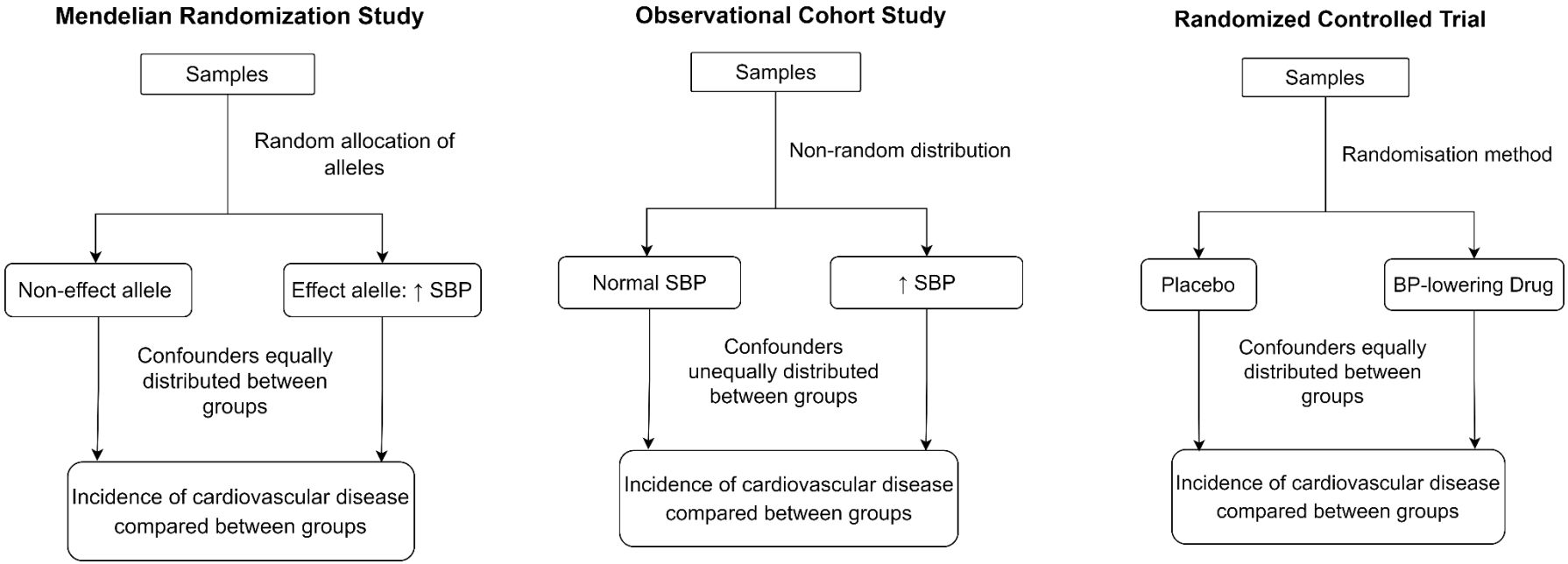


Table 1.2 Comparison of observational study, Mendelian randomisation study, and randomised controlled trial.

	Observational study	**Mendelian randomisation study**	**Randomised controlled trial**
Aim	Single point association	Causal inference (Life-long effect)	Causal inference (2-10 years effect)
Strength	Low cost Extended follow-up period Diverse patient population	No confounding No reverse causation No measurement error	No confounding No reverse causation No measurement error
Weakness	Confounding Reverse causation Measurement error	Instrumental variable assumptions	High cost Patients with a high risk of disease Short follow-up duration Ethical issues in some cases

MR was first applied using parent-offspring data where the alleles are randomly allocated from the parents to their offspring with an equal probability. However, due to data limitations and low statistical power, most MR is conducted within whole populations rather than within families. By conducting at a population level, MR studies can leverage large-scale genetic data from diverse populations, thereby increasing statistical power. Individuals within populations share a substantial common ancestry, and through the random assignments of alleles across a series of mating events the genetic constituents of each individual can be retraced. The application of MR in population studies was based on the premise that at a population level, genetic variants can distinguish groups that exhibit average differences regarding modifiable exposure. In such studies, the classification based on genetic variations should be independent of confounding factors that introduce bias in traditional observational studies, such as behavioural and environmental factors that take place post-conception (165). Thus, the association between genetic variants and traits should be free from confounding factors. Any differences in outcomes between groups that are defined by genetic variants can be due to the exposure, assuming no selection bias relating to that genetic variation (165).

Mendelian Randomisation vs. Randomised Controlled Trials

In RCTs, participants are randomly assigned to either a treatment or a control group. As this random allocation is almost certain to ensure the elimination of potential confounding associated with the treatment and the outcome, the causal effect of the treatment on the disease can be estimated in a manner that is likely to be free from bias. The rationale of MR is closely analogous to that of RCTs (157) (**Figure 1.2)**. MR exploits the random allocation of genetic variants to derive estimates of the causal effect of an exposure on an outcome, which can overcome some types of confounding that undermine observational research (166). In MR, randomisation is essentially the random allocation of genetic variants at conception. We will take a hypothetical example where MR is applied to investigate the causal effect of elevated SBP (exposure) on the risk of CVD (outcome) by leveraging genetic variants associated with the exposure. In a given population, individuals can be segregated into subgroups according to their genetic variants. We have a group of people with the genetic allele A associated with normal SBP, and another group of those with genetic allele B for the same genetic variant associated with high SBP. Due to the random allocation of genetic variants at conception, which is independent of other traits, environmental influences, and other risk factors in the population, the risk of disease can be directly compared between the two groups (group with allele A vs. group with allele B). A significant difference in the disease risk between the two groups will indicate the causal effect of elevated SBP on the disease.

RCTs provide estimates of the effects of an intervention or a treatment during a period of the study, whereas MR estimates the effects of an exposure on an outcome over a lifetime, which is based on the assumption that the association between the genetic variants and the exposure remains unchanged throughout time (167). The differences in the duration for which the effects are estimated can result in significant differences in the effect estimates derived from MR studies and RCTs. Moreover, in the scenario of population-based MR involving unrelated individuals, the expected similarity between allele groups is not necessarily assured, as it typically would be in a well-conducted RCT. Population stratification or assortative mating can lead to associations between allele distribution and certain traits at the population level (168). Therefore, in genetic association studies, MR provides an estimated inference rather than an absolute determination.

A MR study typically involves the following key steps:

1. Identify the primary hypothesis, exposure of interest and outcome of interest.
2. Identify appropriate data sources.
    a. Type of MR investigation: one-sample or two-sample.
    b. Type of data: individual-level data or summary statistics from genome-wide association studies.
3. Select genetic instruments and check the validity of MR assumptions.
4. Conduct the primary MR method.
5. Conduct sensitivity analysis to check the robustness of the MR findings.
6. Interpretation of the results in the context of current knowledge and other studies.

1.3.3 Data sources for MR studies

MR studies can be performed using data from a single sample (known as one-sample MR) or from two samples (known as two-sample MR). In one-sample MR genetic variant-exposure and genetic variant-outcome associations are measured in the same individual, whereas in two-sample MR the associations of the genetic variants with the exposure and outcome are obtained from two independent samples. Two-sample MR has facilitated the application of the MR approach in research, especially when it is challenging to obtain a dataset that includes data on genetic variants, exposure, and outcome due to the infrequency of the trait or the high cost of measuring. Both one-sample and two-sample MR have their benefits and limitations, as illustrated in **Table 1.3.**

Table 1.3 Comparison of one-sample MR and two-sample MR.

IVs, instrumental variables

One-sample MR	Two-sample MR
Exposure and outcome measured in the same individual	Exposure and outcome measured in two independent samples
Individual-level data	Individual-level or summary data
Able to explore associations of IVs with confounders	May not be possible to explore associations of IVs with confounders
Winners' curse bias	Avoid Winners' curse bias when using non-overlapping samples
Weak instruments bias towards confounded exposure-outcome association	Weak instrument bias is expected towards the null
Possible overfitting	No overfitting when using non-overlapping samples
Can do sub-groups and non-linear effects with sufficient sample size	Difficult to do sub-group or non-linear effects when using summary data. Possible to conduct if using individual data.
Not subject to bias from adjustments made in GWAS	Potential bias from adjustments made in GWAS if using summary data. Not relevant if using individual data.

Two-sample MR offers some advantages over one-sample MR. Weak instrument bias occurs when the associations between genetic variants and exposure are weak. In one-sample MR, weak instruments tend to bias towards the confounded exposure-outcome association, leading to inflated false positive rates. In two-sample MR without overlapping samples, the weak instrument bias is expected to towards the null, and does not result in false positive findings (169). Furthermore, one-sample MR can be affected by the Winners' curse, a phenomenon where the discovery results for an association tend to be exaggerated while replication studies tend to be more modest. In one-sample MR, if the genetic variant-exposure association is based on the discovery GWAS that used the same sample as the sample for testing the genetic variant-outcome association, the genetic variant-exposure effects tend to be exaggerated whereas the genetic variant-outcome effects will be potentially underestimated. This results in the underestimation of true causal effects in one-sample MR, especially when multiple genetic variants are used as instrumental variables. The Winners' curse is unlikely to happen in two-sample MR with non-overlapping samples (170). Using summary data from GWAS consortia in two-sample MR can increase the statistical power, especially when examining the effects on binary disease outcomes (171). When using summary data for MR estimation, bias due to overfitting can occur if there are

overlapping participants in the datasets utilised for estimating genetic associations with the exposure and the outcome. However, recent studies have indicated that this overlap may not distort the estimates as much as previously assumed, unless the genetic instruments have a weak association with the exposure. Methods have been developed to measure and adjust for this bias (172).

A key assumption for two-sample MR estimation is that two samples come from the same underlying population. More specifically, the association of genetic variants with exposure should not differ between the two samples, although the exposure may not be measured in the outcome dataset. A particular issue emerges when the two samples are from different ethnic groups are the potentially different patterns of linkage disequilibrium between population ancestry groups. Consequently, the association of genetic variants with the exposure may not be consistent (a weaker association or no association) in the outcome dataset (168). Thus, to meet the assumption of the underlying population, the two-sample MR approach typically uses data from two samples that have the most similar underlying population as possible in terms of genetic ancestry and contextual factors (165). One-sample MR analyses are not affected by these concerns, and they also do not necessitate the harmonisation of the genetic variants throughout the dataset. Moreover, in a one-sample MR setting, the results of MR analysis can be compared with those from traditional observational analysis in the same individuals (173).

Although the utilisation of summarised data is usually equated with the two-sample approach, the advantages and disadvantages of the two options (one-sample vs. two-sample, and individual-level vs summarised data) are distinct. One-sample MR can only be conducted using individual-level data, whereas two-sample MR can be conducted using either individual-level or summarised data. Summarised data are often extracted from GWAS with large sample sizes, which can increase the power to detect a causal effect. Also, conducting MR with summarised data allows for the replication of the analysis by a third party, provided they have access to the same data. However, having access to solely to summarised data imposes limitations on the range of analyses that can be performed. For example, it is unlikely to be able to conduct subgroup analyses and non-linear effects with summary data from large GWAS consortia. Whereas, with sufficient sample sizes and data on the relevant stratifying risk factors as well as their genetic instruments, these analyses can be conducted with individual-level data (174). Moreover, if the GWAS has made adjustments for a variable causally downstream of the exposure or outcome, a two-sample MR using summarised data from this GWAS is subjected to collider bias and should provide commentary on the probable implications of this. (Collider bias occurs when exposure and outcome independently affect a common third variable, which is controlled for in the study

design or analysis). By contrast, MR using individual-level data is not affected by collider bias as it can decide which variables to adjust for (175).

The statistical power of a MR analysis that uses individual-level data increases in correspondence with the sample size and the proportion of the variability in the exposure that is explained by the genetic variants. Summary data are often extracted from GWAS, which is a research approach used to estimate the association between genomic variants and the risk for a disease or a particular trait. The statistical power of a MR study that uses summary-level data depends on the precision of the estimated genetic variants-outcome associations. Given genetic variants usually account for a small proportion of the variance in the phenotype, conducting a robust MR study normally necessitates a large dataset (165).

1.3.4 Selection of genetic instruments

A crucial step in designing a MR study is the selection of genetic instruments, which directly influences the validity of MR analysis. There are primarily two strategies for selecting genetic instruments: a biological approach and a statistical approach. These two approaches are not mutually exclusive and can be used together in a complementary manner. In the biological approach, the selected genetic variants are from genetic regions that are biologically connected to the exposure. As the biological understanding is often not comprehensive, the biological approach should be applied with caution. In the statistical approach, the selection of genetic variants commonly includes the SNPs identified from reliable GWAS, which can detect a substantial number of SNPs that have a strong association with a specific phenotype, as genome-wide significant proxies of the exposure for MR estimation. When the selection of genetic variants is only based on their association with the exposure without reference to their functionality, it is important to be mindful of the potential pleiotropy of the variants.

The MR analysis can be performed using genetic variants from a single gene region or using variants from multiple regions of the genome (also known as polygenic analysis). If a gene region demonstrates a distinct biological association with the exposure, MR analysis using variants from this single gene region appears to be a more credible approach for evaluating the causal association of that exposure of interest in comparison to the polygenic analysis that uses all genome-wide significant proxies of the exposure. MR analyses for investigating potential drug targets often use genetic variants from a single gene region which encodes the protein target of the drug (176). For example, in the MR analysis for investigating ACEi, genetic variants in the region of gene *ACE* were considered as proxies for the effect of ACE inhibition (122). However, for a polygenic or complex trait, which is influenced by multiple genes and can be influenced by environmental factors, polygenic analysis is more

appropriate as there is no single gene linked to the trait. The polygenic analysis approach enables researchers to examine the consistency of the causal finding across multiple variants that affect the exposure through various biological mechanisms (168).

1.3.5 Estimation of the causal effect

In one-sample MR with individual-level data, estimation of causal effects is commonly obtained using the two-stage least-squares (2SLS) method. The first stage involves regressing the exposure of interest on the genetic variants and any relevant covariates. In the second stage, the outcome is regressed upon the predicted value of the exposure from the first stage and the same covariates (177). The covariates incorporated into both stages of the estimation should be considered carefully as these covariates can be affected by the exposure or the outcome; in such cases, adjustment for the covariates may introduce bias to the MR estimates (178). It is recommended to only include as covariates age, sex, genomic principal components of ancestry, and technical covariates (168).

The 2SLS method can also be applied to two-sample MR, provided that individual-level data is available for both samples. Summarised data are commonly used in two-sample MR setting. In two-sample MR with summarised data, if only a single genetic variant is used as an instrument, the causal effect of the exposure on the outcome is simply estimated as the ratio of the variant-outcome association divided by the variant-exposure association. If multiple genetic variants are used as instruments, the most commonly used method to obtain the MR causal estimate is the inverse-variant weighted (IVW) method (179). With uncorrelated genetic instruments, the IVW effect estimate can be obtained by averaging the ratio estimates from the individual genetic instruments using a formula that applies inverse-variance weighting derived from meta-analysis studies (180). The IVW estimate is equivalent to the 2SLS estimate obtained from MR with individual-level data when outcomes are continuous.

One important assumption for the IVW method is that the ratio estimates from each variant provide independent evidence on the causal effect, which occurs when the genetic variants are independent of each other. This assumption is usually satisfied by removing one of each pair of genetic variants that are in linkage disequilibrium, which is an important step in conducting MR analysis (181). Furthermore, genetic variant-exposure and genetic variant-outcome associations are normally reported per additional copy of a particular allele. Given the fact that not all publicly available data resources are consistent in reporting strand information accurately, it is important to harmonise data to ensure that genetic associations refer to the same effect allele when combining summarised data on genetic associations (182).

The IVW method can be performed using a fixed-effects or random-effects meta-analysis model. For summarised data, the IVW method with multiplicative random-effects is recommended as the primary analysis method due to its efficiency with valid instrumental variables and its ability to account for heterogeneity in variant-specific causal estimates. This recommendation holds unless the number of variants is minimal, or all variants originate from the same gene region. The random-effects analysis aligns with the fixed-effect analysis when the heterogeneity among ratio estimates for the individual variants is as expected by chance. However, in the presence of excess heterogeneity, the fixed-effect analysis may provide misleadingly narrow confidence intervals, making it unsuitable. The multiplicative random-effects IVW method yields valid causal estimates under the assumption of balanced pleiotropy, where pleiotropic effects on the outcome have an equal probability of being positive or negative (183).

1.3.6 Assumptions underlying MR

As MR is a form of instrumental variable (IV) analysis, there are three critical IV assumptions for valid MR analysis (**Figure 1.3**):

- Assumption 1 (the relevance assumption): The genetic variant is associated with the exposure.
- Assumption 2 (the independence assumption): The genetic variant is not associated with the outcome due to confounding pathways.
- Assumption 3 (the exclusion restriction assumption): The genetic variant does not affect the outcome other than through the exposure.

The first assumption ensures that, based on the genetic variants, genetically distinct subgroups will exhibit different average levels of the exposure. This ensures the existence of a systematic difference between the subgroups. The first assumption would be violated if the genetic variants used in the analysis are not strongly associated with the exposure, these variants are commonly known as weak instruments.

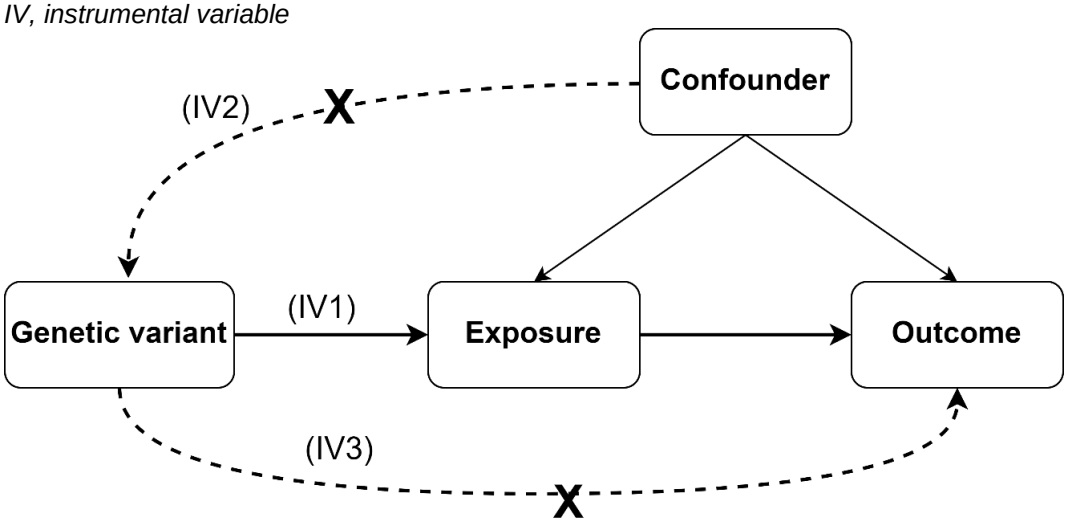
The second assumption ensures that all variables, excluding the exposure and its downstream effects, are evenly distributed between the subgroups defined by the genetic variants. The violation of the second assumption can be due to the association of the genetic variant with confounding factors of the exposure-outcome association. This assumption would also be breached if the genetic variant and the outcome share a mutual cause, like ancestry, unless this is accounted for in the analysis (173).

The third assumption ensures that the sole pathway from the genetic variant to the outcome is through the exposure. This implies that the genetic variant does not have a direct impact

on the outcome, and there is no alternative pathway for the variant to affect the outcome except via the exposure (173).

Figure 1.3 Instrumental variable assumptions.

IV, instrumental variable



Assessing the plausibility of these assumptions is vital for MR findings. While the first assumption can be formally evaluated, the second and third assumptions can be disproved but cannot be demonstrated to be true (165).

The weak instrument is the main factor that causes the violation of the first assumption. A genetic variant, which is not strongly associated with the exposure of interest in statistical terms, is commonly known as a weak instrument. The strength of the instrument can be evaluated by using first-stage F-statistics; a strong instrument is often defined as one with an F-statistic greater than 10 as a rule of thumb (184). Using weak instruments for MR analysis can introduce weak instrument bias into the estimation. Specifically, MR causal estimates are biased towards the observational confounded association in one-sample MR using the 2SLS method. By contrast, in the two-sample MR setting, the bias is towards the null. A weak instrument, unlike an invalid instrument, has the potential to be strengthened through the accumulation of more data. Aggregating genetic variants into an allele score is a way to increase instrument strength (184).

The validity of the second and third IV assumptions can be disproven and otherwise examined using a range of sensitivity analyses (section 1.3.7); however, they cannot be proven to be true (165).

As genetic variants are fixed at conception, they cannot be affected by conventional confounders (e.g. age, sex, environmental factors, etc). Nevertheless, population stratification or assortative mating can introduce bias to the MR estimates. Population stratification occurs when genetic variants are associated with phenotypes due to the underlying structure within the population under study. Different ethnic origins or non-

random mating are factors that can lead to population stratification. Population stratification can result in associations between the genetic variants and the outcome at the population level (168), violating the second IV assumption. This confounding is not easily corrected with current MR methods. MR using family data can account for this confounding at the population level; however, within-family MR is limited by the availability of data and lower power to detect a causal association (165). In GWAS studies, population stratification is typically addressed using principal components analysis of the genetic variants or using linear mixed models. However, it is important to note that this approach may not always fully account for the stratification for some phenotypes (185).

Violations of the third IV assumptions can be caused by pleiotropy. Pleiotropy refers to a phenomenon where a genetic variant is associated with multiple traits. Pleiotropy is a frequent phenomenon in the human genome and, if not adequately accounted for, can violate the IV assumptions and introduce bias to the MR estimates (138). In circumstances where a genetic variant, which is used as a proxy for the exposure, is also associated with another risk factor for the outcome, it leads to the violation of either the second or third IV assumption. Specifically, if pleiotropy results in the association of the genetic variant with the outcome through a confounding variable, it would lead to a breach of the second IV assumption. In another scenario where pleiotropy leads to an alternative causal pathway from the genetic variant to the outcome not through the exposure, it would result in the violation of the third IV assumption.

Pleiotropy can be categorised into two main types: horizontal pleiotropy and vertical pleiotropy. Horizontal pleiotropy, also known as direct pleiotropy, occurs where a genetic variant is associated with multiple variables, each leading to the outcome through an independent causal pathway (186) **(Figure 1.4 A)**. In other words, the genetic variant not only affects the exposure of interest but also has distinct, direct effects on the outcome. This can violate the third IV assumption and introduce bias into the MR estimates, either exaggerating or diminishing it. Vertical pleiotropy, also known as mediation, refers to a phenomenon where a single genetic variant is linked to multiple variables that are on the same causal pathway to the outcome (138). This could take place either before or after the exposure of interest **(Figure 1.4 C, 1.4 D)**. Vertical pleiotropy is the phenomenon that MR seeks to identify - a sequence of causal effects from the genetic variant, through the exposure, to the outcome. Vertical pleiotropy, unlike horizontal pleiotropy, does not breach the core assumptions of MR and is typically not problematic. Differentiation between horizontal pleiotropy and vertical pleiotropy requires knowledge of the associations between the variables in MR analysis. For example, in a MR study that investigates the causal association between high-density lipoprotein (HDL) cholesterol and heart disease, genetic

variants associated with HDL were used as genetic instruments for the analysis. As genetic variants associated with HDL are likely also associated with low-density lipoprotein (LDL) and triglycerides, horizontal pleiotropy bias can be introduced to the MR estimates as the association of LDL or triglycerides with heart disease is not completely mediated through HDL levels (187). Another example is the genetic variation in the *FTO* gene. Variants in the *FTO* gene are associated with satiety, which is the sensation of being full of food. Mediation analysis demonstrated that the association of variants in the *FTO* with increased adiposity was partially explained by their influence on satiety (188). Variants in the *FTO* gene can be used as instruments for BMI if both satiety and BMI follow the same causal pathway to the outcome, and there is no causal pathway from the genetic variant to the outcome not through BMI. Thus, this illustrates an example of vertical pleiotropy.

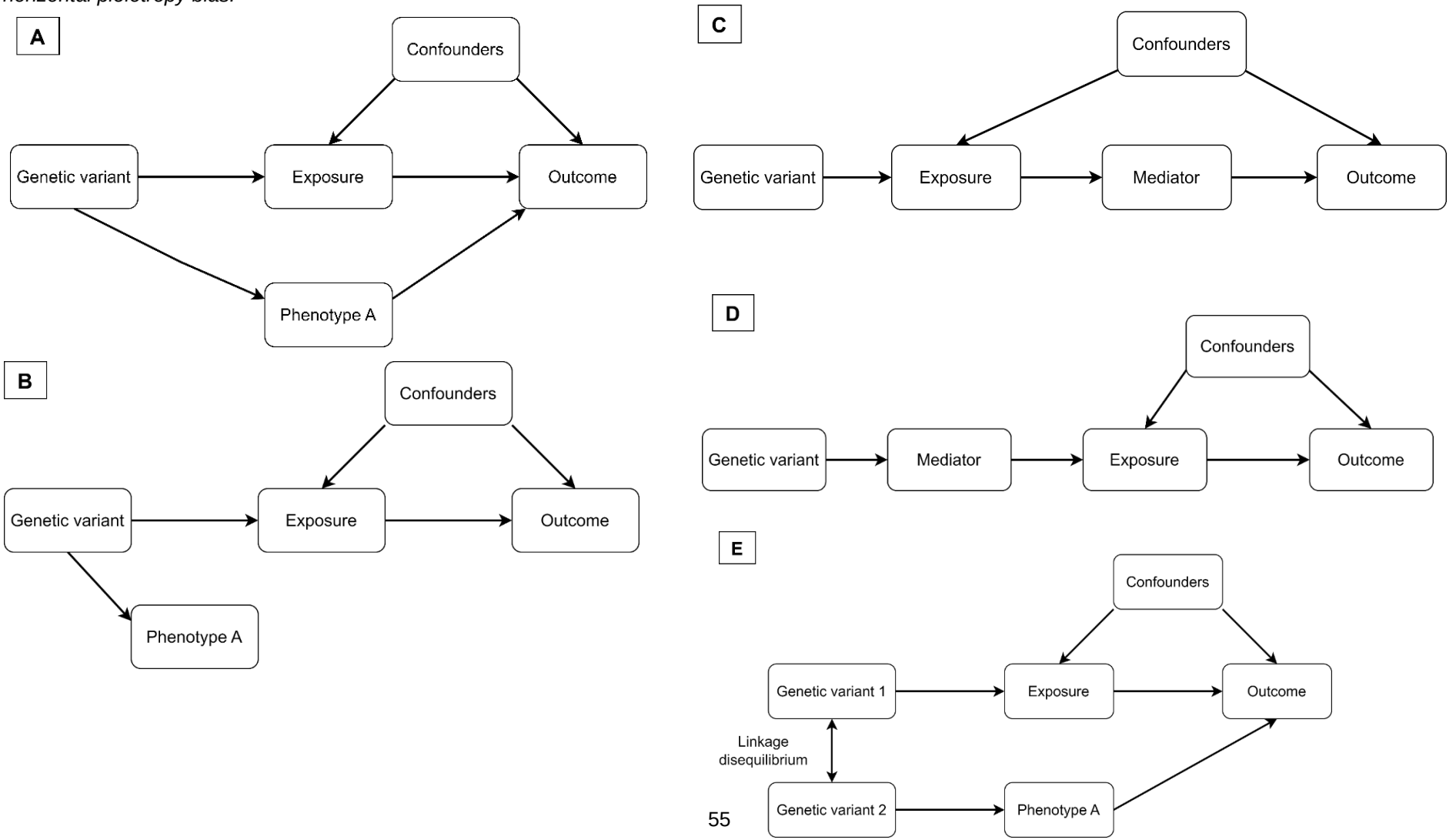
Linkage disequilibrium (LD) refers to the non-random allocation between alleles at different loci on the chromosome in a given population (189). Genetic variants located in close physical proximity to each other on the same chromosome may be in LD and thus be inherited together. LD can introduce bias to the MR estimates when the genetic instrument in LD with other genetic variants that are associated with competing risk factors of the outcome (**Figure 1.4E**). In this case, the bias introduced by LD is similar to the bias created by horizontal pleiotropy (173).

Although excluding horizontal pleiotropic variants can produce more reliable results, this should be done with caution as removing vertical pleiotropic variants can introduce bias to the MR estimates (168). To mitigate the risk of horizontal pleiotropy, genetic instruments can be selected by using genetic variants located in genes whose biological functions are comprehensively understood. For example, variants in the *LDLR* can be considered instruments for high LDL cholesterol, or variants in the *CRP* can be used to proxy the effect of C-reactive protein (173).

Various MR methods have been developed to account for horizontal pleiotropy, such as MR Egger, weighted median, simple median, and Mendelian Randomization Pleiotropy RESidual Sum and Outlier (MR PRESSO) (Section 1.3.7). These methods are robust to different forms of pleiotropy and can evaluate the robustness of MR estimates obtained from the IVW method. For example, a MR study investigates the causal association between HDL cholesterol and heart disease, using genetic variants associated with HDL as genetic proxies. The IVW detected an association between HDL and CVD. However, genetic variants associated with HDL are likely also associated with LDL and triglycerides, which can introduce bias to the HDL-heart disease association. MR analysis using MR Egger, weighted median, and weighted mode resulted in a null association between HDL and CVD. These findings suggest that the association detected in IVW is driven by genetic pleiotropy (187).

Figure 1.4 Figures illustrating different types of pleiotropy.

(A), Horizontal pleiotropy introducing bias: *The genetic variant is associated with another risk factor (phenotype A) of the outcome.* ***(B), Horizontal pleiotropy not introducing bias:*** *The genetic variant is associated with another trait unrelated to the outcome.* ***(C)&(D), Vertical pleiotropy:*** *Mediator is on the pathway of causal effects from genetic variant, through the exposure, to the outcome.* ***(E), Confounding by linkage disequilibrium:*** *Another genetic variant, correlated with the variant used in MR analysis, affects the outcome through a pathway that is not via the exposure, creating a bias analogous to the horizontal pleiotropy bias.*



1.3.7 Sensitivity analysis

Many robust methods have been developed to detect and correct for horizontal pleiotropy, which leads to the violation of the third IV assumption, in MR analysis with multiple genetic variants. As any cause of instrument invalidity can be mathematically represented as bias resulting from pleiotropy, these methods may be able to assess the sensitivity of the MR findings to instrument invalidity more generally rather than only from pleiotropy (190). These methods can result in valid causal estimates under weaker assumptions than the standard IVW methods. Robust MR methods can primarily be divided into three main categories: consensus method, outlier-robust method, and modelling method (191). Consensus methods, such as weighted median (192) and mode-based estimation (193), take a summary measure of the distribution of the ratio estimates as the causal estimate, rather than taking the weighted mean of the ratio estimates as in the standard IVW method. For example, the weighted median method takes the median of a distribution of the ratio estimates in which genetic variants with more precise ratio estimates receive more weight as the causal estimate (192). This method can produce an unbiased estimate if up to 50% of the weight comes from valid genetic variants. The consensus methods are robust to outliers as the median and mode of a distribution are not influenced by the magnitude of outliers (191). Outlier-robust methods, such as MR PRESSO (186), MR Robust (194), and MR-Lasso (194), identify and then either down weight or remove genetic variants from the MR analysis that have outlying causal effect estimates. The standard IVW method is subsequently performed on the set of genetic variants after adjusting for or removing outlying variants. Modelling methods, such as MR-Egger (195), MR-RAPS (196), MR-Mix (197) and contamination mixture method (198), construct a model of the distribution of causal estimates arising from invalid IVs or establish a specific assumption regarding how the IV assumptions are violated. The MR-Egger method is performed similarly to the IVW method, except for the inclusion of an intercept term in the regression model. The MR-Egger method requires the Instrument Strength Independent of Direct Effect (InSIDE) assumption to be satisfied to provide consistent estimates of the causal effect. The InSIDE assumption states that the pleiotropic effects of genetic variants must be not correlated with the association between genetic variants and exposure. Furthermore, the MR-Egger intercept test, which is a part of the MR-Egger method, also provides a test of IV assumptions. If the intercept term differs from zero, it suggests that the average pleiotropic effect is not zero, indicating evidence of directional pleiotropy (195). The number of new methods for MR analysis is growing regularly, but it is unlikely that one approach is robust to all types of pleiotropy. These methods can give valid causal inferences under less stringent assumptions than the standard IVW method, although many of these have less statistical

power than the traditional MR, resulting in very broad confidence intervals. Therefore, rather than performing all available MR methods, a sensible range of methods should be chosen for sensitivity analysis to be performed in conjunction with the standard IVW MR (168). Weighted median and weighted mode methods are usually included in summary data MR analysis, although other appropriate methods can be selected for the specific application of the MR study (165).

1.3.8 Interpretation

Well-conducted MR studies can provide more reliable evidence than conventional observation studies as MR can overcome major limitations such as confounding, measurement errors, and reverse causation. However, it is recommended that the causal effect demonstrated by a MR study should be interpreted with caution (199). The reliability levels of the MR findings depend on several factors, including the validity of the MR assumptions, the consistency of estimates obtained from different methods, and the sensitivity analysis results (168). Suppose a MR study demonstrates robust evidence of association that remains consistent across sensitivity analyses, with no major violations of MR assumptions. In that case, the exposure-outcome association can be interpreted as compatible with causality (200). Triangulating the MR findings with findings from other study designs, such as observational studies and RCTs, is a powerful approach to form a comprehensive understanding of the causal relationship under investigation. For example, several large MR studies were conducted to investigate the causal effect of C-reactive protein (CRP) concentration on the risk of CVD (201, 202). These studies did not detect any causal association, whereas evidence from previous traditional observational studies suggested a causal association of high CRP concentration with an increased risk of CVD (202). This suggests that evidence from observational studies might be biased by confounding and/or reverse causality, and that targeting CRP is unlikely to be a promising therapeutic approach for preventing CVD.

MR studies have shown their capacity to detect causality which is qualitatively aligned with findings from RCTs (200). However, it is likely to see the discrepancies between estimates derived from MR studies and those obtained from RCTs. This is due to the differences in how genetic variants and clinical or pharmaceutical interventions respectively affect the exposure. MR estimates specifically relate to changes in the exposure caused by the genetic variants that are utilised as instrumental variables. The genetic variants, which are fixed at conception, typically affect usual levels of exposure over a lifelong period. Therefore, MR analysis can be interpreted as evaluating the lifelong impact of an exposure on an outcome. On the other hand, clinical trials assess the estimated impact of intervening on the exposure

(by medicines or interventions) at a specific point in time. Hence, the estimates from MR analyses are usually larger than those obtained from observational studies and clinical trials. For these reasons, the MR estimate is more appropriately viewed as a statistical test for a causal hypothesis and an indicator of the effect direction. It should not be interpreted as the estimated effect of a medical intervention at a particular time point in life (168).

Clinical guidelines should not be formulated solely based on the results from MR studies (187). It is important to interpret MR findings within the broader context of existing evidence and biological knowledge. Davies et al. illustrated an example of how the understanding of biological context plays a crucial role in interpreting findings from MR studies (187). Genetic variants in the *CHRNA5* gene are associated with both increased frequency of tobacco smoking and increased risk of lung cancer. Individuals carrying these variants are also expected to face greater challenges in reducing smoking following a diagnosis of lung cancer. As the genetic variant carriers would have engaged in more intense smoking before the diagnosis, it would be inappropriate to use these genetic variants as instruments to investigate the causal association of reduced smoking on mortality in people after they are diagnosed with lung cancer. The association between the genetic variants with mortality reflects the causal effect of life-long smoking, rather than the effect of current smoking status, on the mortality of lung cancer patients (187).

1.3.9 Future of MR

Large-scale GWAS, molecular epigenetics, advanced -omics technologies, data availability of biobanks, and advanced methodological MR methods have enabled the exploitation of MR approaches in a wider scope.

The data availability of biobanks has facilitated the development of one-sample MR studies. The availability of biobanks that provide genetic data at the individual level, such as UK Biobank, and BioBank Japan, enables the non-randomisation MR analysis and subgroup analysis, which typically necessitate individual-level data to be conducted (173).

GWAS studies have driven the emergence of MR, facilitating the identification of genetic instruments for MR analysis as well as providing datasets for two-sample MR. The rapid development of GWAS over the last decade has facilitated more comprehensive and larger MR studies for known risk factors as well as novel exposures. Including more genetic variants as instruments can increase the statistical power of MR to detect causality. Given that some robust methods necessitate many variants for giving reasonable inferences, the availability of many variants enables more robust findings via the comparison of results between different methods (i.e. triangulation of evidence). Moreover, a large number of variants also enables the categorization of variants based on their associations with different

mechanisms, which is particularly useful when the genetic variants affect the exposure via different mechanisms. By doing this, we can obtain separate MR estimates for each category of variants, enabling the understanding of underlying disease mechanisms. Given the advantages of a large number of variants, researchers need to be cautious about the increased amount of potential noise to analysis and the increased heterogeneity in the variant-specific estimates.

The advancements in high-throughput analytical techniques enable the availability of a wide range of omics data, including gene expression and methylation, proteins, lipids, transcription factors and metabolites (203). This provides a large number of risk factors for MR in addition to traditional epidemiological risk factors. Genetic variants are expected to explain more variance in the omics risk factors because the variants are more proximal to these risk factors than the traditional epidemiological risk factors. One major challenge of using omics in MR studies is the shared genetic predictors between related omics risk factors (e.g. multiple genes whose expressions share the same genetic predictors), with novel methods being developed to address this challenge (204).

Although the developments in data availability and methodological practice facilitate the performance of MR analyses, data analysts play a pivotal role in conducting the analysis, from formulating the study designs, and choosing the appropriate dataset, to critically interpreting the findings. Although some parts of a MR study can be performed automatically using available MR platforms such as MR-Base (205), conducting the MR study always requires critical thought to obtain the most reliable findings (206).

1.3.10 Mendelian randomisation analysis in drug development

1.3.10.1 Need for MR in Drug Development

The pharmaceutical industry faces significant challenges in drug development, with high failure rates, particularly for targets associated with new mechanisms. This is largely due to the poor predictability of preclinical models and the translational gap between animal models and human biology (207). These failures lead to increased costs for drug development and a decrease in the availability of novel agents (208, 209). Drug targets based on human genetic evidence have a more than twofold likelihood of progressing into clinical practice (109). Two-thirds of the drugs approved in the US in 2021 were supported by evidence from human genetics (210). About 70 out of the 670 known effects of licensed drugs have been validated through associations at the loci of the genes that code for their respective target proteins (211), using GWAS. These advantages have prompted the broad acceptance of genetic data as a crucial source of evidence in drug discovery. The accessibility of extensive genetic

association data and genotyped biobanks provides an opportunity for investigations that are both time and cost-efficient (205).

1.3.10.2 How to Conduct MR Analysis in Drug Development

MR can be utilised to investigate drug effects, where the exposure under study is a perturbation of a drug target (212). This MR approach can be applied in drug development to study the efficacy and potential adverse drug effects of newly developing drug targets as well as explore opportunities for repurposing existing drugs (212). Due to the random allocation of genetic variants at gametogenesis, MR can reduce confounding and reverse causality issues inherent in traditional observational studies, providing a robust framework for drug target identification and validation. Similar to the original MR approach, MR analysis in identifying and testing drug targets also relies on key assumptions for causal inference including the relevance, independence, and exclusion restriction assumptions, as mentioned in section 1.3.2**.**

In MR investigations of drug effects, genetic instruments are selected to mimic the perturbation of the drug protein targets. Genetic variants can be selected as IVs based on two steps. The first step is the identification of a gene or group of genes that encodes the drug target proteins. The second step is the selection of genetic variants, typically located in the vicinity of the genes (cis-variants), to proxy perturbation of the drug protein targets. IVs can be cis-variants associated with known downstream functions of the drug target (such as effects on metabolites levels or biomarkers), gene expression quantitative trait loci (cis-variants associated with expression of the gene encoding the drug target protein), or protein expression quantitative loci (cis-variants associated with levels of the target protein) (168). The selection of genetic instruments for the MR study on drug effects typically focuses on the cis-variants rather than the variants that are located outside the encoding gene, often on a different chromosome (trans-variants). One proposed explanation for this is that cis-variants are more likely to have effects specific to the protein of interest than trans-variants, which was also illustrated in a paper using CRP as an example (213). Currently, there is no established consensus on how proximal cis-variants should be to the corresponding gene and on whether we should consider enhancer or promoter regions for genes as cis-acting. The essential requirement is that the variant should be associated with the function or expression of the drug target (168). Indeed, one approach used by previous MR studies for the validation of the selected instruments is to examine the association of those instruments with a well-known effect of the drug of interest (214). For drugs with protein targets that are encoded by multiple genes such as calcium channel blockers, previous MR studies have included genetic variants related to the individual genes as IVs and combined their data to

examine the effect of perturbing the drug target. LD clumping was performed to obtain uncorrelated variants before running the MR analysis (122).

Despite the advances MR brings to drug development, the approach is not without its challenges. Methods normally used for MR (mentioned in sections 1.3.5 and 1.3.6) can be utilised for MR investigation of drug target perturbation. The number of genetic instruments selected to proxy drug target perturbation may be minimal as these are usually from a specific gene region rather than from the entire genome. Given the limited number of variants, many of the methods used for assessing the sensitivity of the MR findings to horizontal pleiotropy may not be applicable. Altering the dataset and choice of genetic variants, utilising positive and/or negative control outcomes, genetic colocalisation, subgroup analysis and investigating the associations with potentially pleiotropic variables can help to assess the robustness of the MR findings (168). MR provides a powerful tool for drug development by offering a genetic approach to validate drug targets and predict drug effects. By integrating MR into the drug discovery process, pharmaceutical research can become more efficient and cost-effective, ultimately leading to the development of safer and more effective therapies, with putative examples described next.

1.3.10.3 Applications in Drug Development

The development of proprotein convertase subtilisin/kexin type 9 (PCSK9) inhibitors showcases how human genetics can inform drug development and subsequently lead to successful drug approval. PCSK9 inhibitors are a new drug class for lowering cholesterol, being approved for the prevention of coronary artery disease. The identification of the drug target started with the discovery of genetic mutations in the *PCSK9* gene that are causally associated with LDL levels (215). Loss-of-function alleles in *PCSK9* were associated with decreased LDL levels and a reduced risk of coronary heart disease (216). There were no obvious adverse events observed in carriers of loss of function alleles in *PCSK9* (217). In terms of mechanism, PCSK9 protein is thought to act through endogenous and exogenous mechanisms. Exogenously, circulating PCSK9 protein binds to the LDL receptor on the surface of hepatocytes. Once internalised into the liver cell, the PCSK9 protein directs the LDL receptor to the lysosome for degradation. Endogenously, PCSK9 is synthesized and secreted from the Golgi apparatus. It then binds the LDL receptor prior to the receptor reaching the cell surface, leading it to lysosomal degradation. When PCSK9 is inhibited, this process is disrupted, leading to an increase in LDL receptor expression. As a result, more LDL receptors are available to remove LDL cholesterol from the bloodstream, decreasing serum LDL levels. Evidence from genetic studies and knowledge of mechanisms lead to the development of PCSK9 monoclonal antibodies that target the PCSK9 protein. PCSK9

inhibitors have demonstrated efficacy in lowering LDL levels and reducing the risk of CVD in clinical trials, leading to the approval of evolocumab (218) and alirocumab (219).

The successful story of *PCSK9* is an example that illustrates key principles of how MR can be applied to identify drug targets that can lead to drug approval. Specifically, loss-of-function variants in *PCSK9* act as genetic proxies for the effect of LDL lowering through the inhibition of *PCSK9*. These variants are strongly associated with LDL levels, satisfying the first IV assumption. As these variants are located within the protein-coding sequence of *PCSK9*, they are anticipated to manifest their effects by modulating the function of *PCSK9*. Moreover, these variants are not associated with confounding variables, including age, sex, smoking, and T2D. This evidence supports the second and the third IV assumptions (220). The loss-of-function variants in *PCSK9* were directly examined on their associations with CVD, illustrating the simplest form of a MR analysis (221). Inverse-variant weighted or 2-stage least square methods can be applied to weigh the effect on CVD by a unit increase/decrease in the exposure of interest (e.g. LDL cholesterol level in case of *PCSK9*).

Over the past few years, MR has been extensively utilised in the field of drug development (221). MR can provide insights at all stages of drug development, from target discovery to clinical trial design and repurposing existing drugs for new indications.

1.3.10.4 Drug Target Discovery

The process of identifying drug targets is crucial to drug discovery. Drug targets can be first identified from GWAS, hypothesis-driven studies, or high-throughput screens (220). MR can be subsequently used to examine the potential causal effects of the drug targets on the phenotype of interest, using genetic variants as instruments for the analysis. Small molecule and biologic drugs primarily function by interacting with protein targets which are coded by genes (222). Thus, genetic variants can serve as potential instruments in MR analysis for investigating the impact of pharmacologically disturbing these protein drug targets**.** One example highlights the potential of MR for predicting novel drug target opportunities. ATP citrate lyase (ACLY), an emerging target for pharmacotherapy, is an enzyme in the cholesterol–biosynthesis pathway situated upstream of 3-hydroxy-3-methylglutaryl–coenzyme A reductase (HMGCR). In a MR study using individual-level data from 470,478 participants, the effects of genetic proxied ACLY and HMGCR inhibitors were found to be associated with similar patterns of changes in plasma LDL levels and with similar effects on the risk of CVDs per each unit decrease in LDL. This was determined using genetic scores composed of independently inherited variants in the genes encoding ACLY and HMGCR as instruments. Based on the findings from the MR analysis, the paper highlighted the

mechanism by which ATP citrate lyase inhibition decreases plasma LDL levels and suggested the validation for ATP citrate lyase inhibition as a potential genetic target (223).

1.3.10.5 Phase 1 and Phase 2 Applications

MR can identify the potential adverse drug effects of drug target perturbation. An example focuses on the potential of MR for predicting adverse drug events. Statins, one of the most frequently prescribed medications to decrease LDL levels for primary and secondary prevention of CVDs, have been proven to be effective in large-scale RCTs (224). Although statins are generally safe, evidence from RCTs suggests that statins potentially increase the risk of new-onset T2D. A MR study was conducted to investigate whether the inhibition of HMGCR results in the effect of increasing the risk of T2D. By using a single SNP located on the *HMGCR* gene that was associated with decreased LDL levels in a GWAS as an instrument in MR analysis, this study demonstrated the association of this genetic variant with lower LDL levels, higher plasma insulin and higher plasma glucose concentration. Based on this evidence the paper highlighted the at least partial involvement of HMGCR inhibition in the increased risk of T2D (225).

PheWAS is a method that performs hypothesis-free scans of the entire human phenome to identify potential associations between a set of specific genetic variants and a wide range of phenotypes. PheWAS can be combined with MR to perform hypothesis-free scans for adverse effects of drug targets across the human phenome (226). For example, a PheWAS was conducted to identify potential adverse effects of genetically proxied IL-6 inhibitors. The PheWAS suggested that genetically predicted IL-6 inhibition is associated with an increase in the risk of atopic dermatitis, cellulitis, urinary tract infections, and cholecystitis (227). Theoretically, MR can be able to predict the adverse effects of an investigational drug prior to the drug being licensed, thereby contributing evidence for the prediction of the balance between the benefits and risks of the new drug candidate. Furthermore, MR analysis using the outcome of lifespan can be applied to evaluate whether the benefits of a drug target outweigh its adverse effects. By using this approach a MR analysis demonstrated that genetically proxied inhibition of IL-6 signalling was associated with a possible lifespan benefit in a general population, supporting the hypothesis that the benefits of the suppression of IL-6 signalling outweigh its' adverse effects (228). It is important to underscore that these analyses, which are performed using genetic associations with lifespan in the general population, may not accurately reflect the effects expected in specific subgroups in clinical trials (220).

MR method can be used to predict the efficacy of potential therapeutic targets in the early stage of the drug development process. Darapladib is a drug that was developed as an

inhibitor of lipoprotein-associated phospholipase A2 (Lp-PAL2). LP-PAL2, also known as plasma platelet-activating factor acetylhydrolase, is a Ca^{2+}-independent enzyme that circulates bound mainly to the LDL. Lp-PAL2 is found to accumulate in unstable and unruptured plaques. Evidence from observational studies showed that higher Lp-PAL2 levels in blood were associated with an increased risk of major cardiovascular outcomes (229). A loss of function variant (rs76863441) in the *PLA2G7* gene, which encodes Lp-PLA2, was utilised to evaluate the causal effect of Lp-PLA2 on the incidence of major vascular events. The study showed that the loss of function variant was not associated with major vascular events as well as other vascular outcomes (230). These results were consistent with findings from phase III trials of the Lp-PLA2 inhibitor darapladib, which turned out to be failed clinical trials (231). This example underscores the potential role of genetic studies in predicting the effectiveness of drug targets before initiating RCTs, which can help to mitigate trial failures and reduce the costs associated with drug development.

1.3.10.6 Pre-phase 3 and Phase 3 Considerations

MR can be used to inform clinical trial design by helping to construct primary outcomes, investigate drug effects in population subgroups, and test for drug-drug interactions (220).

MR can be used to inform clinical trial design by helping to construct primary outcomes, investigate drug effects in population subgroups, and test for drug-drug interactions (222). For example, MR studies consistently highlighted the protective effects of genetically proxied inhibition of IL6 signalling against CAD and stroke (228, 232, 233), whereas no association was detected for pulmonary arterial hypertension (234). These findings suggest that CAD and stroke are suitable primary outcomes for clinical trials, while pulmonary arterial hypertension may not be appropriate.

MR can help investigate drug effects in population subgroups. For example, an MR study investigated the association of genetically proxied IL-6 inhibitors with CVD risk stratifying by baseline high-sensitivity C-reactive protein (CRP) and CV risk factors. The population subgroup with the highest baseline high-sensitivity CRP levels had a larger absolute benefit, suggesting that this population should be prioritised for IL-6 inhibition therapies in clinical trials (235).

A 2x2 factorial MR stratified across levels of genetic proxies for the drug targets can be applied to study drug-drug interactions. A previous MR study employing this approach suggested additive protective effects against CVD risk of genetically proxied IL-6 inhibition and HMGCR inhibition, indicating the need for further research on this combination for CVD protection (236).

1.3.10.7 Post-phase 3 and Drug Repurposing

The low success rate of the traditional drug development approaches can be attributed to several challenges including substantial expenses, low success rate in preclinical trials, poor concordance between preclinical and clinical outcomes, and difficulties in predicting human drug toxicity (237). Drug repurposing, also known as drug repositioning, is a process of researching new indications for established drugs or unapproved drugs that were studied advancingly. Drug repurposing can address several challenges associated with traditional drug development approaches. Some reviews report that repurposing efforts tend to have a higher success rate than new drug applications. Specifically, nearly 30% of repurposing efforts are successful and receive approval, whereas the figure for new drug applications is about 10% (238). Drug repurposing has received renewed interest during the COVID-19 pandemic. Several drugs, such as remdesivir, dexamethasone, and tocilizumab, were granted by Food and Drug Administration (FDA) to treat COVID-19.

MR can help identify potential new indications for established drugs, offering a pathway for drug repurposing based on genetic evidence. One example is the potential repurposing of sulfonylurea and glucagon-like peptide 1 receptor (GLP1R) agonists as candidates for the treatment of Alzheimer's disease. A MR study was conducted to examine the association between genetic variation in the target of four antidiabetic drug classes and Alzheimer's disease risk. Both genetic variation in the sulfonylurea target (*KCNJ11*) and the glucagon-like peptide 1 (GLP-1) analogue target (*GLP1R*) was associated with lower blood glucose (which is their known clinical effect) and associated with a lower risk of Alzheimer's disease (239). Notably, these same genetic variants were not associated with body weight reduction, which is also a known clinical effect of GLP1R agonism. This raised concerns over the validity of genetic variants used to proxy the effects of GLP1R agonists. However, genetic proxies for glucose-lowering effects of GLP1R agonists may not necessarily be associated with body weight due to distinct signalling mechanisms underlying effects on glycaemic control and body weight reduction (240). Phase III trial in Alzheimer's disease with semaglutide, an oral GLP1R agonist, is currently in progress (241).

1.3.10.8 MR studies on effects of systolic blood pressure and anti-hypertensive drugs

MR studies have explored the causal links between elevated BP and outcomes including both CVD and non-CVD. Genetically proxied SBP has been causally linked with hypertension-related CVDs such as CAD, stroke, HF, AF, and a variety of other CVDs including aortic aneurysm, aortic stenosis, dilated cardiomyopathy, endocarditis, peripheral vascular disease, and rheumatic heart disease. It was also found to be inversely associated

with venous thromboembolism (242). The long-term impact of BP lowering at a population level was evaluated using the UK Biobank, with an overall 16.9%, 30.8%, and 56.2% decrease in morbidity for a 5, 10, and 22.7 mmHg decrease in SBP respectively (242). While MR studies found no causal effect of genetically determined BP on estimated glomerular filtration rate (eGFR), a lower genetically proxied eGFR is linked with higher BP in other MR studies using eGFR as an exposure.

While SBP and DBP are correlated, only SBP is incorporated into CVD risk prediction models. Epidemiologically, DBP is more strongly associated with the onset of coronary heart disease in younger populations, whereas SBP is a better predictor in individuals over 60 years of age. Although most SNPs identified in BP GWAS are associated with both SBP and DBP, a small subset of SNPs are exclusively associated with one or the other. A MR study utilising three distinct sets of BP GWAS SNPs - 242 SNPs associated with both SBP and DBP, 120 SNPs exclusive to SBP, and 80 SNPs exclusive to DBP - was conducted to elucidate the separate effects of SBP and DBP on hypertension outcomes (243). The study found that SBP is the causal driver for CAD, stroke, and ischemic stroke, while DBP is the causal driver for small vessel stroke. Additionally, SBP is exclusively associated with HF, AF, and T2D (243).

MR analysis can help to identify potential causal associations of genetically determined antihypertensive drug classes with unintended effects of the drugs, providing valuable perspectives for drug repositioning. A MR study conducted by Gill et al (122) evaluated the effects of ACEi, beta-blockers, and CCBs on the risk of CAD and stroke. They utilised the genetic variants located in the genes encoding protein targets for each drug class. Their study found that CCBs and beta-blockers had protective effects against CAD, while ACEi and CCBs were protective against stroke. These findings were comparable to results from RCTs, highlighting the potential of the MR approach to complement evidence from RCTs and observational studies (122). Using a similar MR approach for first-line antihypertensive drug targets, other studies revealed that CCBs can mitigate the risk of HF and AF (244); while ACEi have been shown to decrease the risk of T2D (245). Both ACEi and CCBs have been found to have renal-protective effects (246). On the other hand, beta-blockers and thiazide diuretics have been found to elevate the risk of T2D (245); ACEi could potentially increase the risk of schizophrenia and colorectal cancer (120, 121). Furthermore, research also finds the positive impact of antihypertensive medications on Alzheimer’s disease risk, which is perhaps attributed to their influence on SBP (247).

1.3.11 Limitations of MR

Despite its potential in a wide range of medical research, MR has limitations.

Low power

Typically, only a small proportion of the total variance in a trait is explained by genetic variants. Therefore, a very large sample size is usually required to have sufficient statistical power in MR studies. A single SNP normally explains < 1% of the total variance in traits, thus researchers usually combine multiple genetic instruments to maximize the statistical power of the instrument. Despite doing that, the total variance explained is usually still small. In our previous MR study, 242 SNPs selected as genetic instruments for SBP account for only 2.82% of the total variance in SBP (243). Calculation tools have been developed to estimate statistical power in MR studies (248).

Lack of genetic instruments

With the advancements in genetic technologies, there has been a significant surge in the number of GWAS, making it more feasible to find genetic instruments in MR studies. However, the majority of large-scale GWAS have focused on traits that are relatively straightforward or cost-effective to measure, with the most studied population in GWAS being European (200). Thus, there is still a lack of availability of GWAS for certain traits, especially with other non-European ancestries. In the context of MR analysis used for drug development, there is also a lack of GWAS that uses molecular data from the appropriate tissue or cell type, which is required to study the effect of certain drug target (220).

Complexity of association

Another limitation is the poor understanding of the underlying biological pathways of the detected associations, which sometimes can lead to misinterpretation. For example, a MR study found that a SNP (rs2228145, also known as p.Asp358Ala) located in the *IL6R* gene was associated with increased circulating levels of IL-6R and IL-6, and decreased risk of CAD. By contrast, evidence from observational studies indicated that higher circulating levels of IL-6R and IL-6 were associated with an increased risk of CAD. One potential explanation for the observed contradiction is that an increase in circulating levels of IL-6 in carriers of rs2228145 is not a result of enhanced protein synthesis. Instead, it might be attributed to a reduction in the cellular binding of IL-6, which leads to increased circulating levels of IL-6 but reduced IL-6 cell signalling (232). This example highlights the importance of integrating in-depth biological understanding into detected associations in MR studies.

Canalization

MR estimates may be biased by canalisation, which refers to a process by which lifelong genetic effects of the genetic variant on a phenotype are buffered by compensatory developmental processes. If canalization occurs, the actual effect of the exposure on the

outcome might be less than expected. The impact of canalization is expected to be less problematic for exposures that occur after fetal development (e.g. smoking, drinking, etc). However, it can pose a greater challenge for severe protein-coding mutations expressed during fetal development (200). Currently, there is no method established to address canalization bias in MR.

Pleiotropy

One of the IV assumptions in MR is the exclusion restriction, which states that there should be no bias due to horizontal pleiotropy. Given the wide spread of pleiotropy in the genome, the exclusion restriction assumption may be violated in some instances. Phenome-wide association studies (PheWAS) tool, such as Phenoscanner, examines associations of IVs with many traits. This approach facilitates the identification of pleiotropic signals, however, the interpretation may be complicated by multiple testing issues and normally requires a profound understanding of the associations (168). Many MR methods have been developed to account for horizontal pleiotropy, such as MR Egger, weighted median, simple median, and MR PRESSO (Section 1.3.6). For example, a MR study investigates the causal association between HDL cholesterol and heart disease, using genetic variants associated with HDL as genetic proxies. The IVW detected an association between HDL and CVD. However, genetic variants associated with HDL are likely also associated with LDL and triglycerides, which can introduce bias to the HDL-heart disease association. MR analysis using MR Egger, weighted median, and weighted mode resulted in a null association between HDL and CVD. These findings suggest that the association detected in IVW is driven by genetic pleiotropy (187).

Weak instrument bias

Weak instruments provide a poor prediction of the exposure. Studies using weak instruments will have low statistical power to reject the null hypothesis. Also, violation of the IV assumptions such as the ‘exclusion restriction’ assumption can be amplified by using weak instruments. First-stage F-statistics can be used to detect weak instrument bias. Strong instruments are often defined as F-statistic > 10 as a rule of thumb (187). Weak instrument bias can be mitigated by using two-sample MR, and increasing sample size. For two-sample MR, the estimate using weak instruments is biased towards the null, however, it has limitations such as bias induced by overlapping samples (section 1.3.3). A large sample size enhances the statistical power, but it can be resource-intensive and not always feasible (184).

1.4 Summary-based Mendelian Randomization Analysis

1.4.1 About Summary-based Mendelian Randomization

Given the complex nature of linkage disequilibrium structure in the genome, the association between a genetic variant and a disease or trait in GWAS does not automatically imply that the genetic variant is causal or functional in the disease progress. There is still limited understanding of the underlying mechanisms of the associations identified from GWAS, particularly at the genetic and molecular levels. Summary-based Mendelian Randomization (SMR) (249), an extension of the Mendelian Randomization analysis, has been proposed as a tool to prioritise genes underlying GWAS hits for follow-up functional studies (249). SMR uses genetic variants, associated with gene expression levels (also known as eQTL), as proxies for the effect of the gene expression levels on outcomes. By integrating summary-level data from independent GWAS and data from expression quantitative trait loci (eQTL) studies, SMR tests whether genes whose expression levels are associated with complex traits. This method can avoid major limitations linked to observational studies and RCTs, such as confounding bias, reverse causation, and feasibility. By employing cis-eQTLs as instruments for mRNA or protein expression, we can significantly reduce the likelihood of horizontal pleiotropy (213), thereby helping to ensure the validity of the MR assumptions.

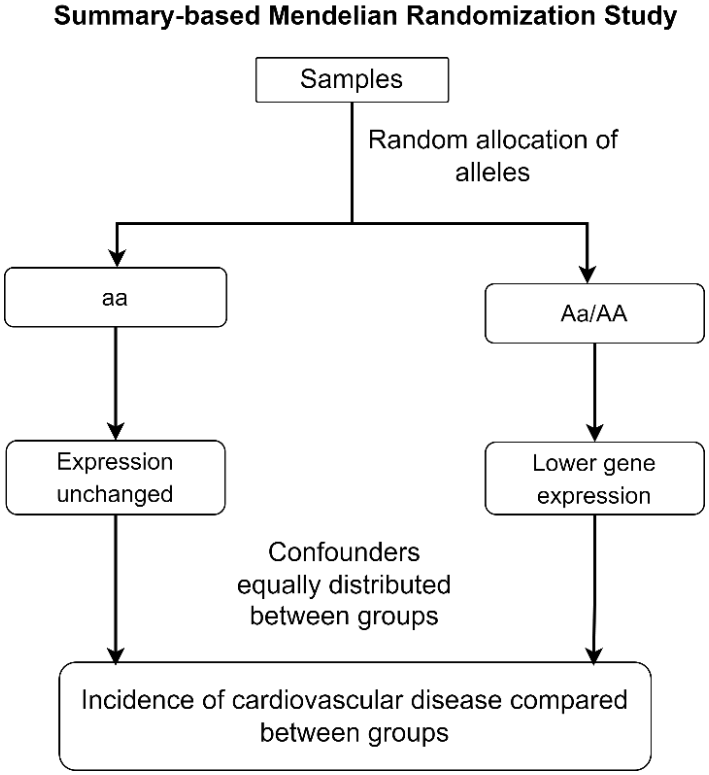
The rationale behind this method is akin to that of MR analysis. If a genetic variant has an influence on a particular gene expression level, this will result in different gene expression levels among individuals with different genotypes of the genetic variant (e.g. AA, Aa, and aa). Then if the expression level of the gene influences a trait, phenotypic differences will be observed among the different genotype groups. In other words, the genetic variant will also have an effect on the trait (249). As eQTL are stable across the life course and generally not influenced by confounding from non-genetic factors, eQTL can be used as proxies for investigating associations between gene expression and outcomes (**Figure 1.5**).

The SMR approach selects a single most significantly associated cis-eQTL (i.e. eQTL located near the target gene) in a particular eQTL study as an instrument. The method then tests the association between a trait and the expression level of a gene using the effect size of the cis-eQTL from the eQTL study and the effect size of the same cis-eQTL from a GWAS. The SMR estimates are obtained using a two-stage least-squares approach (249).

One issue that arises in the SMR approach is that the association observed in the SMR analysis does not necessarily mean that the selected cis-eQTL is a causal variant for both gene expression and the trait of interest. The selected cis-eQTL can be in LD with two distinct causal variants - one influences gene expression while the other affects trait

variation. In order to discover the linkage scenario, the authors of the SMR method proposed a method called heterogeneity in dependent instruments (HEIDI) (249).

Figure 1.5 Figure illustrating the rationale behind the summary-based Mendelian randomisation analysis.



1.4.2 Heterogeneity in Dependent Instruments Test

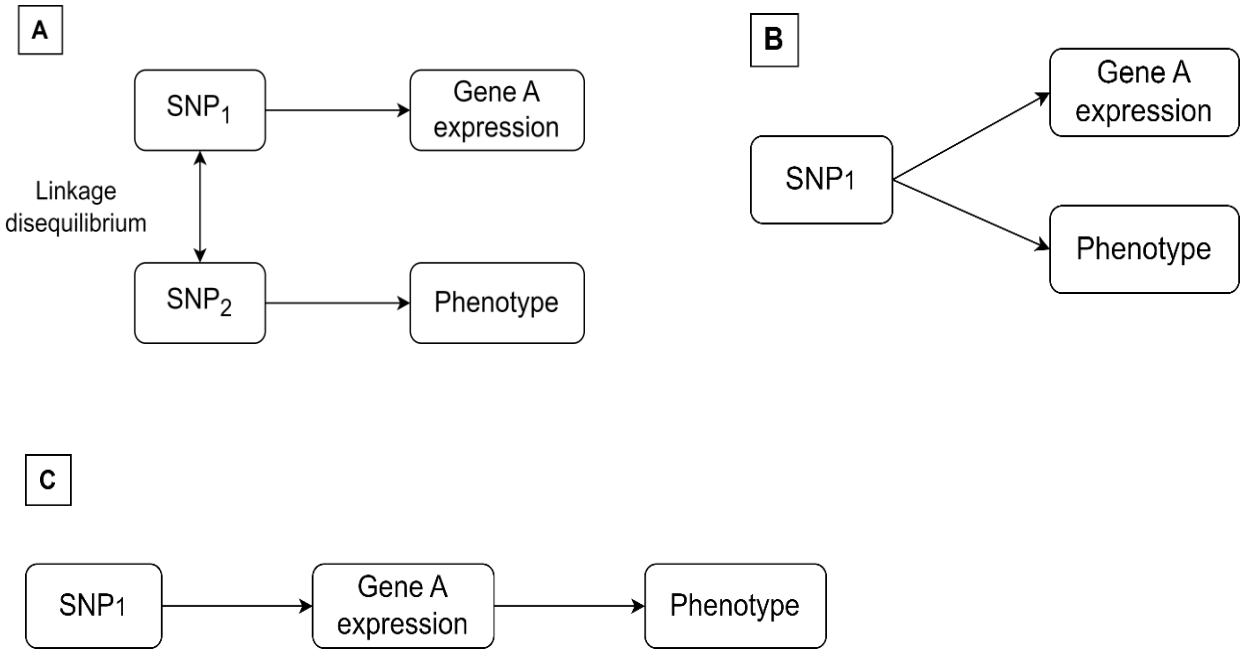
The SMR approach incorporates the HEIDI test (249) to evaluate whether the observed association between gene expression and the outcome is a result of a linkage scenario. In this scenario, the SNP impacts the outcome not through the regulation of gene expression, but because the SNP that affects expression is in LD with another SNP that independently affects the outcome. This would dispute one of the MR assumptions. **Figure 1.6** illustrates three possible explanations for an association observed in an SMR test – the association can be attributed to a linkage scenario, pleiotropy, or causality.

If gene expression and phenotype share the same causal variant (**Figure 1.6B, 1.6C**), the SMR-estimated effect values of the gene expression on the phenotype, estimated for any SNPs in LD with the causal variant, remain identical. The HEIDI method has been developed to examine whether there is heterogeneity in the estimated effect values for multiple SNPs in the cis-eQTL region. The null hypothesis is that there is no heterogeneity, which is equivalent to that there is a single causal variant for both gene expression and the phenotype (**Figure 1.6B, 1.6C**). Contrary to most conventional analyses that disregard

results with high statistical p-values (such as p > 0.05), HEIDI analysis disregards results with low heterogeneity test p-values – the lower the p-value in the HEIDI test, the higher the probability of rejecting the null hypothesis, indicating that the observations align with a linkage model (**Figure 1.6A)**. The SMR authors recommended that a HEIDI test p-value less than 0.01 is indicative of an association due to a linkage scenario (120, 249).

Moreover, as HEIDI leverages local LD to differentiate models, it is not able to provide a perfect distinction between pleiotropy and linkage (249). The statistical power to distinguish between pleiotropy and linkage decreased with the LD between two causal variants, which was shown in simulations under a linkage model with two distinct causal variants (249). In a situation where the two causal variants are in perfect LD, it becomes impossible for any statistical test to differentiate between the two models of pleiotropy and linkage (249). Therefore, SNPs in very high LD with the top associated cis-eQTL (at $r^2 > 0.9$) or in very low LD or not in LD ($r^2 < 0.05$) with the top associated cis-eQTL will be omitted. Additionally, SNPs in the cis-eQTL region with $P_{eQTL} > 1.6 \times 10^{-3}$ (equivalent to $\chi_1^2 < 10$) were also removed from the HEDI test to avoid weak instrumental variables (249).

Figure 1.6 Possible explanations for an observed association between the gene expression level of a gene and a phenotype.
(A) Linkage disequilibrium: *The top associated eQTL (SNP_1) is in linkage disequilibrium with another SNP (SNP_2), one SNP affects the expression levels of the gene while another affects the phenotype;* ***(B) Pleiotropy:*** *SNP_1 is the causal variant for both gene expression and phenotype, with their effects being independent of each other.* ***(C) Causality:*** *SNP_1 affects the phenotype via the expression level of the gene.*



Moreover, as HEIDI leverages local LD to differentiate models, it is not able to provide a perfect distinction between pleiotropy and linkage (249). The statistical power to distinguish

between pleiotropy and linkage decreased with the LD between two causal variants, which was shown in simulations under a linkage model with two distinct causal variants (249). In a situation where the two causal variants are in perfect LD, it becomes impossible for any statistical test to differentiate between the two models of pleiotropy and linkage (249). Therefore, SNPs in very high LD with the top associated cis-eQTL (at $r^2 > 0.9$) or in very low LD or not in LD ($r^2 < 0.05$) with the top associated cis-eQTL will be omitted. Additionally, SNPs in the cis-eQTL region with $P_{eQTL} > 1.6 \times 10^{-3}$ (equivalent to $\chi_1^2 < 10$) were also removed from the HEDI test to avoid weak instrumental variables (249).

1.4.3 Applications of SMR

The SMR method has been applied to three traits (BMI, height, and waist-to-hip ratio adjusted by BMI) and two diseases (rheumatoid arthritis and schizophrenia) in its original paper, conducted by Zhu et al. (250). The study aimed to prioritise genes for the five diseases or traits, leveraging gene expression data in peripheral blood and GWAS data from the large-scale GWAS meta-analyses. The study retained 5,967 probes that had at least one cis-eQTL at the genome-wide significant level ($p_{eQTL} < 5\times10^{-08}$). Top associated cis-eQTL were used as proxies for the effect of gene expression levels. Results from SMR analyses identified 289 genes whose gene expression levels in peripheral blood were associated with the outcomes of interest. Subsequently, a HEIDI test was conducted to test whether the observed association was due to a linkage scenario. The study found 104 genes that passed the HEIDI test, including *TRAF1* and *ANKRD55* for rheumatoid arthritis, and *SNX19* and *NMRAL1* for schizophrenia. About 20% of the identified genes are novel candidates, which passed the SMR and HEIDI test and are situated more than 1 Mb away from the nearest GWAS hit. About 60% of the prioritised genes are not the closest annotated gene to the top associated GWAS SNP. These findings demonstrate the potential of the SMR approach in identifying the functionally relevant genes underlying GWAS hits that can be prioritised for follow-up functional studies. SMR and HEIDI tests have been applied for other diseases and traits to identify novel genes, including CAD (251), multiple sclerosis (252), cancers (253), and lipid traits (251).

SMR can be applied to evaluate the effects of drugs on complex diseases. In this approach, cis-eQTL of a gene that encodes the target of a drug will be used as instrumental variables to proxy the effects of a drug target on a disease outcome. The rationale behind this is that the alleles for an eQTL that increase/decrease the expression level of a gene, leading to greater/lesser gene product, can be used as a proxy for the effect of a drug that increases/decreases the level of the same gene product. Thus, the expression level of a gene can be used as an exposure to study the causal effects on the outcomes of interest.

For example, cis-eQTL for gene *ACE* can be used as an instrumental variable in SMR analysis to examine the effect of ACEi on outcomes of interest (120). The SMR method has shown its potential in predicting the association of several drug classes with diseases, such as antihypertensive drug classes and mental health diseases (120), and lipid-lowering drugs and non-alcoholic fatty liver disease (254).

1.4.4 Limitations of SMR

The statistical power to detect causal effects in tissue-specific SMR analysis is restricted by the sample size of available eQTL datasets and the strength of the instrumental variable. SMR analysis typically requires a large sample size for detecting small effect sizes, which is not always available with eQTL datasets. Moreover, SMR analysis normally uses summary statistics from GWAS and eQTL studies. Cell types and environment tissues may differ between GWAS and eQTL studies, potentially resulting in a variation in the regulatory environments of the tissue types under investigation. Another limitation is that when using SMR to investigate the effect of perturbating the target of a drug, variants affecting gene expression or protein levels may not necessarily affect the function of the protein product and, thus the effects observed downstream may not be comparable to those attainted by the pharmacological alteration of the drug target (221).

1.5 Colocalisation analysis

1.5.1 Colocalisation analysis

Although hundreds of genomic loci that influence complex diseases have been identified and replicated by GWAS, deciphering these results, especially identifying the underlying molecular mechanisms, has been a significant challenge. One approach to enhance biological understanding following GWAS is to incorporate functional data with GWAS results (255). A common approach is to look for overlaps between complex trait-associated variants and eQTL variants, which helped to identify the causal gene and the tissue where the effect was exerted in previous studies (255). If a single genetic variant at a gene locus is associated with both traits (the two traits are also called “colocalised traits”), the probability of a shared causal mechanism between two traits is significantly increased (255). Alternatively, the observed associations could be the result of distinct causal variants which are correlated due to linkage disequilibrium. Given the multitude of significant associations identified by GWAS, the likelihood of random overlap between association signals for different traits is high (255). As a result, overlap alone does not enable the identification of causal variants (255). Statistical colocalisation methods (256-259) were developed to

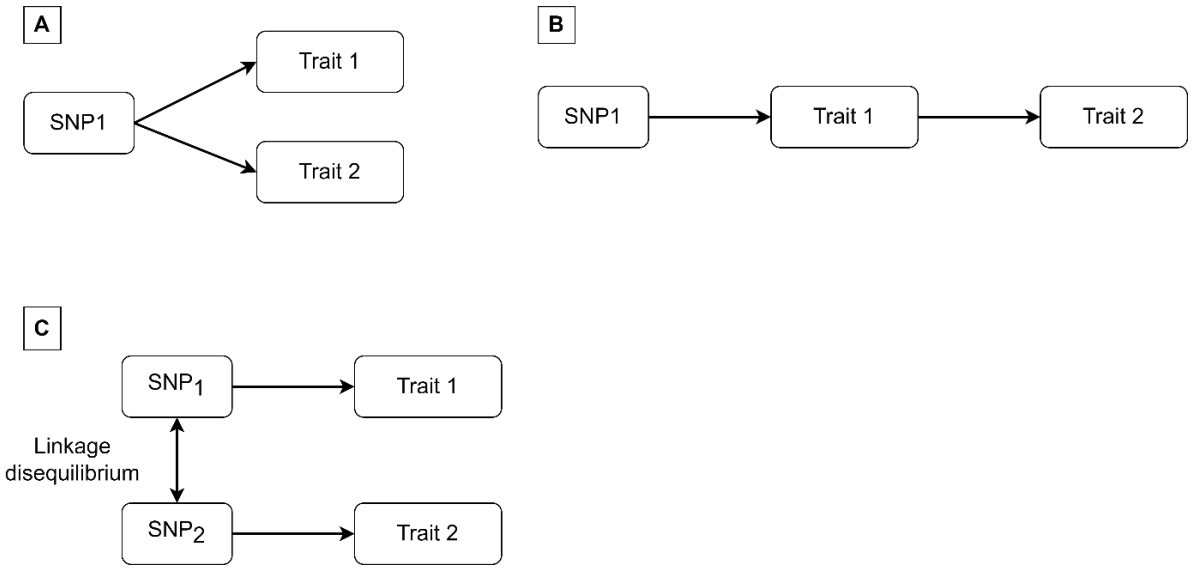
address this issue and have been effective in determining whether a molecular trait (e.g. gene expression) and a disease trait share a causal variant in a genomic region, and potentially prioritise a candidate causal gene.

The overlap between association signals for two different traits at a genetic region can be explained by two possible scenarios (**Figure 1.7**). In the first scenario, there are two distinct causal variants, possibly in linkage disequilibrium, each associated with a different trait. Conversely, in the second scenario, there is one shared causal variant that is associated with both traits. Colocalisation has been proposed to differentiate between these two potential situations.

There are two families of methodological approaches that have been developed for colocalisation: proportional colocalisation and enumeration colocalisation (260). Proportional colocalisation tests a null hypothesis of proportionality under which, if two traits share the same causal variants in a specific genetic region, the associations of these traits with the causal variants are proportional to each other (261). If the statistical model indicating proportional genetic associations is contradicted by the evidence, there is no colocalisation evidence. Proportional colocalisation depends on the assumption that there is either a single causal variant in the genetic region or the traits are on the same causal pathway with all variants primarily exerting influence on the same upstream trait. The HEIDI test is categorised under the proportional colocalisation family (260).

Figure 1.7 Diagrams illustrating possible situations underlying an overlap between association signals for two different traits at a genetic locus.

(A) *Two unrelated traits share a single causal variant;* ***(B)*** *Two traits have a single causal variant and share the same causal mechanism;* ***(C)*** *Two traits have distinct causal variants that are in linkage disequilibrium. Scenarios A and B exemplify the concept of colocalisation.*



Enumeration colocalisation uses a Bayesian framework to assess whether two association signals are colocalised, leveraging summary statistics for both traits across multiple SNPs in

a particular genetic region. Evidence for different hypotheses in the Bayesian framework is compared, based on the pre-specified prior probabilities and the summarised association data. Several methods (255) have been developed based on this framework, each with varying assumptions. The coloc method, proposed by Giambartolomei et al. (255), works by systematically comparing evidence for five different hypotheses in the Bayesian framework. The hypotheses are:

H_0: No causal variants for either trait

H_1: Causal variant for trait 1, not for trait 2

H_2: Causal variant for trait 2, not for trait 1

H_3: Distinct causal variants for trait 1 and trait 2

H_4: One shared causal variant for trait 1 and trait 2

In this framework, the test for colocalisation can be re-formulated as evaluating the evidence support for the hypothesis H_4. A large posterior probability for hypothesis 4 (PP_4) supports the model with a shared causal variant affecting both traits. Whereas a large posterior probability for hypothesis 3 (PP_3) supports the model with two independent causal SNPs associated with each trait, indicating confounding by linkage disequilibrium.

The coloc method works under the assumption that there is at most one causal variant for each trait. It also requires the definition of prior probabilities, which are the prior probability for a genetic variant being associated with trait 1 (p_1), the prior probability for a genetic variant being associated with trait 2 (p_2), and the prior probability for a variant being associated with both traits (p_{12}).

1.5.2 Applications of colocalisation

Previous studies utilising colocalisation methods have identified multiple regions associated with various autoimmune or autoinflammatory diseases, such as type 1 diabetes and celiac disease (262, 263). In recent years, colocalisation has been increasingly used as a sensitivity analysis for MR studies. Zheng et al. conducted a phenome-wide association study to systematically investigate the causal role of plasma proteins in the human phenome (264). To select the genetic instruments for the plasma proteins, protein quantitative trait loci (pQTL) from five published GWAS were pooled and cross-validated. Two sample MR analysis was conducted to systematically investigate the causal effects of 1,002 plasma proteins on 153 diseases and 72 disease-related risk factors. 413 associations between proteins and traits were observed in the MR analysis. Colocalisation analysis was conducted as a sensitivity analysis to investigate whether the observed association between protein and trait are due to a shared causal variant. To mitigate the risk of inaccurate interpretations from colocalisation analysis due to the potential existence of multiple conditionally distinct

predictors in the genetic region, the authors conducted pairwise conditional and colocalisation analysis. Evidence from colocalisation analysis supported 283 of 413 protein-phenotype associations, with the posterior probability for H_4 > 80% being considered as colocalisation. Furthermore, by evaluating the data from historic drug development programs, the authors reported that target-indication pairs, supported by both MR and colocalisation, had a higher likelihood of receiving approval (264). This study emphasizes the value of the combination of MR and colocalisation in detecting and prioritizing potential therapeutic targets.

1.5.3 Limitations of colocalisation

Colocalisation is a useful tool in MR studies, particularly in providing evidence that supports or challenges the MR assumptions. This can be seen in the context of gene expression and protein levels (265), which are often instrumented with a single genetic variant. However, colocalisation analysis has several limitations (168). One of the limitations is the challenge of identifying the true causal exposure. This is particularly pertinent when the exposure is gene expression because the colocalisation results are often varied depending on the tissue selected. Another limitation is the statistical power of the analysis. In proportional colocalisation, failure to reject the null hypothesis does not only indicate colocalisation. It could also be due to either disease not being associated or insufficient statistical power, which could be a result of a limited number of samples or an incomplete genetic map (266). Colocalisation methods that are based on the Bayesian approach (e.g. coloc (255), coloc-SuSiE (267), HyPrColoc (259)) typically need genetic variants in a specific genetic region under study to have robust associations (p less than 10^{-4} or 10^{-5}) with both traits in order to affirm the colocalisation. For example, in coloc method (255), if genetic variants are strongly associated with only one of the traits, the method may support hypothesis 1 (one causal variant for trait 1 but not trait 2) or hypothesis 2 (one causal variant for trait 2 but not trait 1), rather than hypothesis 3 (distinct causal variants for traits 1 and 2) or hypothesis 4 (a shared causal variant for both traits). Thus, in these cases, colocalisation methods may not provide robust evidence either to support or against the colocalisation. Another limitation relates to the assumption that there is at most one causal variant for each trait, which might restrict the accuracy of the analysis in the scenario of multiple causal variants in the given genetic region. Although there are some methods proposed for the scenario of multiple causal variants, these methods require LD information and are highly sensitive to LD misspecifications (268).

1.6 Gaps in Knowledge in Hypertension Research that Can Potentially Be Answered by Mendelian Randomisation

Despite significant advances in understanding the genetic determinants of blood pressure and hypertension through genomic and MR studies, several critical gaps remain that MR can potentially address.

Limited Scope of Cardiovascular Outcomes:

Most MR studies have primarily focused on traditional cardiovascular outcomes such as myocardial infarction, stroke, and coronary artery disease. There is limited understanding of the broader spectrum of cardiovascular effects, including peripheral artery disease, HF, and arrhythmias.

Insufficient Exploration of Renal and Diabetic Outcomes:

While some research has explored the relationship between BP and renal outcomes, there is a lack of comprehensive studies that integrate other causal determinants of renal and diabetes outcomes in hypertension. Understanding the interplay between BP, renal function, and diabetes is crucial given the interrelated nature of these conditions.

Underrepresentation of Newer and Non-First-Line Antihypertensive Drugs:

Most existing MR studies have concentrated on first-line antihypertensive drugs (e.g., ACEi, beta-blockers). There is a paucity of research on newer and non-first-line drug classes, which could offer insights into alternative therapeutic strategies.

Mechanistic Insights into Drug Action:

While associations between BP and various health outcomes have been established, there is limited understanding of the underlying mechanisms, particularly at the genetic and molecular levels. The role of gene transcription influenced by antihypertensive drugs remains underexplored.

Transcriptional Mediation:

Few studies have investigated whether the effects of antihypertensive drugs are mediated through the transcription of genes encoding targeted proteins. This gap is crucial for understanding the direct and indirect pathways through which these drugs exert their effects.

Diverse Populations and Genetic Variability:

Existing MR studies often lack diversity in their study populations, leading to potential biases and limiting the generalisability of findings. There is a need for studies that encompass diverse genetic backgrounds to ensure broader applicability of results.

1.7 Aim of this thesis

The overall aim of this thesis is to fill some of those critical gaps by providing a more detailed understanding of the full spectrum of antihypertensive drug effects on a wide range of cardiovascular, diabetic, and renal outcomes. This will be achieved through two main objectives:

Systematic Evaluation Using Mendelian Randomisation:

This thesis will use the MR method to systematically evaluate the causal effects of genetically predicted systolic blood pressure and various BP-lowering drug classes on a comprehensive range of outcomes. By including both established and newly developed drugs, as well as first-line and non-first-line therapies, this research will offer a more complete picture of antihypertensive drug effects.

Transcriptional Mediation Analysis:

The second objective is to investigate if any detected associations are mediated through the transcription of genes encoding proteins targeted by the corresponding antihypertensive drug classes. This will involve using the SMR method to identify and elucidate the genetic and molecular mechanisms underlying these drug effects.

By addressing these gaps, this research will contribute to a deeper understanding of how antihypertensive drugs impact cardiovascular, renal, and diabetic outcomes, potentially informing better therapeutic strategies and improving patient care. The insights gained could also lead to the identification of novel drug targets and the optimization of current treatment regimens based on genetic and transcriptional profiles.

Chapter 2 DATA AND METHODS

In this chapter, I will detail the methodologies employed to conduct Mendelian randomization (MR), summary-based MR, and colocalisation analyses, as well as describe the source datasets utilised throughout this thesis. The principles of MR have been outlined in the introductory chapter. Here, I will expand on the specific analytical steps, the selection and validation of genetic instruments, and the statistical frameworks applied to investigate the causal relationships between genetic variants, gene expression, and disease outcomes. This includes a description of the two-sample MR approach, the utilisation of summary data from large-scale genome-wide association studies (GWAS) and expression quantitative trait loci (eQTL) studies, and the implementation of robust statistical tests such as the HEIDI test and colocalisation analysis to distinguish true causal associations from those confounded by linkage disequilibrium.

All data used in this thesis are publicly available summary statistics from anonymised individuals, with ethical approval and informed consent from all participants being obtained from their original studies.

2.1 Two-sample Mendelian randomisation

Two-sample Mendelian randomization (MR) was performed to examine the causal effects of genetically proxied SBP reduction and BP-lowering drug classes on cardiovascular, diabetic, and renal outcomes.

2.1.1 Genetic instrument selection

To proxy the perturbation of the drug-protein targets, genetic instruments were selected based on two steps: (i) identifying a gene or group of genes that encode the proteins targeted by BP-lowering drugs. The antihypertensive drug classes and their active ingredients were obtained by reviewing the British National Formulary (BNF) (269). The pharmacologically active protein targets of each drug and the corresponding genes were identified using the Drugbank database. (ii) Selection of genetic variants that proxy the drug target perturbation: genetic variants located within the 100 kb on either side of the gene and strongly associated with SBP, a known downstream function of the BP-lowering drugs. Currently, there is no consensus on how proximal the genetic variants should be to the gene centre, and the essential requirement is that the variant should be associated with the function or expression of the drug target (168). Previous MR studies have used the 100 kb window to select the genetic instruments for the cardiovascular drug effects (270, 271). Although antihypertensive drugs have shown effectiveness in reducing both SBP and DBP,

SBP was used as the exposure biomarker instead of DBP due to the driven role of SBP on the risk of CVDs, as demonstrated in previous observational and MR studies (243). Genetic variants associated with SBP in a GWAS meta-analysis of BP traits (115) at a genome-wide significance level of 5×10^{-08} were considered strongly associated with SBP. F-statistics were calculated to evaluate the strength of the variants, with F-statistics greater than 10 indicating a low risk of weak instrument bias (184). Clumping was performed to ensure the genetic instruments for the drug target were independent. SNPs were clumped to a linkage disequilibrium (LD) threshold of $r^2 < 0.1$ using the 1000 Genomes European reference panel (272). Data clumping was performed using PLINK (205).

2.1.2 Exposure data

Genetic association estimates for SBP were extracted from a GWAS meta-analysis of 757,601 European-ancestry participants drawn from the UK Biobank (UKB) and the International Consortium of Blood Pressure Genome-Wide Association Studies (ICBP) (115). ICBP GWAS dataset consists of 77 independent cohorts of 299,024 European-descent participants, with all of them being genotyped and imputed to either the 1000 Genomes Reference panel or the HRC panel. UKB is a prospective cohort study of 458,577 European-ancestry participants, with 408,951 participants being genotyped by customized arrays and imputed to the Haplotype Reference Consortium (HRC) panel. For the UKB GWAS, the mean SBP and DBP values were calculated from two automated or two manual BP measurements. For participants taking antihypertensive drugs, adjustments were made to the medication use by an increment of 15 mmHg in SBP and 10 mmHg in DBP. Sex, age, age^2, and BMI were adjusted within the association analysis. Meta-analysis of summary results from the UKB and ICBP GWAS was performed using a fixed-effects inverse-variance-weight method, comprising up to 757,601 participants and approximately 7.1 million SNPs with minor allele frequency (MAF) ≥1% (115). GWAS summary data for SBP were extracted from the GWAS Catalog (105) using the study accession code GCST006624.

2.1.3 Outcome data

Genetic association estimates for CAD were obtained from a GWAS meta-analysis for CAD comprising 181,522 cases among 1,165,690 individuals of predominantly (> 95%) European descent (273). Summary data for MI were extracted from a GWAS meta-analysis of 171,875 individuals of predominantly European ancestry, out of whom 43,676 were cases (274). Genetic association data for AF was extracted from a GWAS meta-analysis of more than 1,000,000 European-descent individuals, including 60,620 atrial fibrillation cases and 970,216 controls (275). Summary data for HF was obtained from a GWAS meta-analysis of

HF on European-ancestry individuals, comprising 47,309 cases and 930,014 controls (276). Summary data for ischemic stroke was extracted from a GWAS meta-analysis of ischemic stroke on European ancestry participants, including 62,100 cases and 1,234,808 controls (277). Genetic association data for CKD was obtained from a European-ancestry specific GWAS meta-analysis of CKD comprising 41,395 cases and 439,303 controls (278). Summary data for T2D was extracted from a European ancestry specific GWAS meta-analysis of T2D comprising 80,154 cases and 853,816 controls (279).

GWAS data for the outcomes of interest were extracted from GWAS Catalog using the following study accession codes (CAD: GCST90132314, MI: GCST003117, HF: GCST009541, ischemic stroke: GCST90104540). Summary data for AF were downloaded at https://csg.sph.umich.edu/willer/public/afib2018/; summary data for CKD were downloaded at https://ckdgen.imbi.uni-freiburg.de/datasets; summary data for T2D were downloaded at https://diagram-consortium.org/downloads.html.

2.1.4 Instrumental variable assumptions

MR analysis relies on three main assumptions: (i) the genetic variant is strongly associated with the exposure; (ii) the genetic variant is not associated with the outcome due to confounding pathways; (iii) the genetic variant does not affect the outcome other than through the exposure. The validity of the first assumption can be assessed by using the F-statistics (280). The second and third assumptions can be examined using a range of sensitivity analyses, such as MR-Egger, weighted median, simple median, and negative control outcomes. These analyses can evaluate the robustness of MR findings against possible pleiotropy and other biases, as discussed in detail in section 1.3.7.

2.1.5 Data harmonisation

In two-sample MR the associations of the genetic variants with the exposure and outcome are obtained from two independent samples. Data harmonisation was conducted to ensure that genetic associations with the exposure and outcome refer to the same effect allele when combining summarised data on genetic associations. Data harmonisation includes the following steps: (i) Check if the GWAS studies for the exposure and the outcome using the same reference strand: this can be provided by the original GWAS studies, otherwise, can be checked using the reference alleles in the two datasets. For example, if the reference/other alleles in one dataset are C/A, respectively, and the reference/other alleles in the second dataset are C/A, respectively, regardless of the order – i.e., this could be A/C – then this indicates that this SNP is coded on the same reference strand. By contrast, if the SNP in the first dataset had a C/A coding and in the second dataset it had a G/T coding

(regardless of the order) then this indicates that these datasets have used different reference strands. For palindromic SNPs that have alleles on the forward strand that are the same as on the reverse strand (A/T or C/G), the reference strand can be checked using the allele frequency of the effect allele unless the allele frequency is close to 0.5. In this current MR study, the GWAS dataset for the exposure and the outcomes are coded on the same reference strand, therefore, palindromic SNPs were not dropped from the MR analysis. (ii) After confirming that both datasets used the same reference strand, check if the effect allele is the same for variant-exposure and variant-outcome associations. If the effect alleles are different, the SNP needs to be flipped and the effect estimate in the outcome dataset needs to be reversed (multiplied by -1). Data harmonisation was performed using the TwoSampleMR package (version 0.5.10) (205).

2.1.6 Perform MR

Inverse-variance weighted method was used as a primary MR analysis to estimate the effect of genetically proxied antihypertensive drug classes on the outcomes. The Wald ratio was used instead if only one single genetic variant was used as an instrumental variable for the analysis. The MR estimate corresponds to the odds ratio for each binary outcome or beta for each continuous outcome per 10 mmHg decrease in SBP, orientating to the BP-lowering effect of antihypertensive drug classes.

If a single genetic variant (G_j) is used as an instrument, the causal effect of the exposure on the outcome can be estimated as the ratio of the variant-outcome association divided by the variant-exposure association (Wald ratios):

$$\hat{\beta}_{IV_j} = \frac{\hat{\beta}_{Y_j}}{\hat{\beta}_{X_j}}$$

where $\hat{\beta}_{Y_j}$ is the estimated coefficient from univariable regression of the outcome on the genetic variant *j* with variance $se(\hat{\beta}_{Y_j})^2$, and $\hat{\beta}_{X_j}$ is the estimated coefficient from univariable regression of the exposure on the genetic variant *j* with variance $se(\hat{\beta}_{X_j})^2$.

When multiple genetic variants are used as instruments for two-sample MR, the MR causal estimate is estimated using the inverse-variance weighted (IVW) method (179). The ratio estimates from each genetic instrument are averaged using a formula that applies inverse-variance weighting derived from the meta-analysis studies. The variance is the first term from a delta method expansion for the ratio estimate: $\frac{se(\hat{\beta}_{Y_j})^2}{\hat{\beta}_{X_j}^2}$; the pooled IVW estimate (assuming a fixed-effect model) can be calculated as:

$$\hat{\beta}_{IVW} = \frac{\sum_{j=1}^{N} \hat{\beta}_{Y_j} \hat{\beta}_{X_j} se(\hat{\beta}_{Y_j})^{-2}}{\sum_{j=1}^{N} \hat{\beta}_{X_j}^2 se(\hat{\beta}_{Y_j})^{-2}}$$

Where N is the total number of genetic variants selected as instruments. This IVW estimate is equivalent to the two-stage least squares estimate obtained from individual data when the genetic variants are uncorrelated.

One important assumption for the IVW method is that the ratio estimates from each variant provide independent evidence on the causal effect, which occurs when the genetic variants are independent of each other. This assumption is usually satisfied by removing one of each pair of genetic variants that are in linkage disequilibrium via the LD clumping step (181). Furthermore, as genetic variant-exposure and genetic variant-outcome associations are normally reported per additional copy of a particular allele, it is important to harmonise data to ensure that genetic associations ($\hat{\beta}_{X_j}$ and $\hat{\beta}_{Y_j}$) refer to the same effect allele when combining summarised data on genetic associations (182).

IVW analyses were conducted using the MendelianRandomization package (version 0.6.0) in R (281).

The propagation of error method (282) was used to investigate whether the MR estimates for the drug target perturbation significantly differ from those obtained for genetically proxied SBP-lowering effect more generally.

$$\beta_{diff} = \beta_{SBP} - \beta_{target}$$
$$SE_{diff} = \sqrt{SE_{SBP}^2 + SE_{target}^2}$$
$$z = \frac{\beta_{diff}}{SE_{diff}}$$
$$p = 2 \times (1 - pnorm(abs(z)))$$

Where β_{SBP} indicates the MR effect size of genetically proxied SBP reduction on the outcome; β_{target} indicates the MR effect size of genetically proxied drug target perturbation on the outcome; SE_{SBP} and SE_{target} target indicate the standard error of the effect estimate for genetically proxied SBP and the drug target, respectively; z indicates the z-score.

2.1.7 Sensitivity analysis

The IVW method gives valid causal estimates under the assumption that all instruments are valid (173). This does not always hold given the complexity and pervasiveness of pleiotropy in human genetics. Horizontal pleiotropy is a major factor that can lead to the violation of IV assumptions. Robust methods have been proposed to detect and correct possible horizontal pleiotropy. These methods can give valid causal estimates under weaker assumptions than the IVW method (190).

In this MR study, to evaluate the robustness of the MR results against possible horizontal pleiotropy, various sensitivity analyses were conducted, including simple median, weighted median and MR Egger. In the median method, instead of taking a weighted mean of the ratio estimates as in the IVW method, the overall MR estimate is the median of the variant-specific ratio estimates. The median of the ratio estimates will tend towards the true causal effect, under the assumption that over 50% of the genetic variants are valid instruments. In the simple median method, each genetic variant is given an equal weight in the analysis. Whereas in the weighted method, each genetic variant is given a weight corresponding to the inverse of the variance of the ratio estimate, identical to the weighting approach used in the IVW method (192). This means that genetic variants with more precise ratio estimates receive more weight as the causal estimate. Subsequently, the median of a distribution of the variant-specific ratio estimates is taken as the weighted median estimate (192).

The principle behind the MR Egger method is to regress the variant-outcome associations in the variant-exposure associations, allowing for a non-zero intercept in the regression (195). MR Egger can give consistent estimates of the causal effect, under the Instrument Strength Independent of Direct Effect (InSIDE) assumption. The InSIDE assumption states that any direct impacts of the genetic variants on the outcome are uncorrelated with the impacts of these variants on the exposure. The MR Egger intercept represents an estimate of the average pleiotropic effect, which also provides a test of the IV assumptions. A non-zero intercept indicates directional pleiotropy, which leads to the violation of MR assumptions (195).

Apart from the aforementioned statistical test, a negative control outcome test was also performed as a sensitivity analysis. A negative control outcome refers to an outcome that is believed not to be causally affected by the exposure. If a MR investigation detects a causal association between a negative control outcome and the exposure, then this is an indication of violating the instrumental variable assumptions, which could be through pleiotropy or population stratification. Negative control outcome is particularly useful when there is a limited number of genetic instruments, such as MR analysis on a single gene region, this is because the aforementioned statistical methods typically need a large number of SNPs to increase the reliability of the tests (173).

Simple median, weighted median, MR Egger, and negative control outcome analyses were conducted using the MendelianRandomization package (version 0.6.0) in R (281).

2.1.8 Example 1: C reactive protein and coronary heart disease

Background

C reactive protein (CRP) is an acute-phase protein that the liver produces in response to inflammation. CRP binds to low-density lipoproteins and is found in atherosclerotic plaques. High CRP concentration is associated with an increased risk of coronary heart disease in observational studies. However, whether the association is causal remains unclear as observational studies are subjected to confounders and reverse causation. The observed association was attenuated after adjusting conventional risk factors, such as age, sex, BMI, and diabetes, and was attenuated to near null after further adjustment for fibrinogen (283). Given inconclusive evidence for the causal role of CRP on CHD, Wensley et al conducted a study to examine the causal effect of genetically proxied CRP concentrations and CHD, using MR analysis based on individual-level data (202).

Methods

The study performed one-sample MR based on individual-level data. Genetic variants in the *CRP* gene were used as genetic proxies for serum CRP levels (the exposure). The outcome of interest is coronary heart disease. Individual-level data are obtained from 47 studies in the CRP CHD Genetic Collaboration, encompassing a total of 194,418 participants, out of which 46,557 were cases. 89% of the total participants were of European descent.

The authors used 4 SNPs (rs3093077, rs1205, rs1130864, and rs1800947) in the *CRP* gene as genetic instruments, which capture 98% of the common haplotype of the *CRP* gene in the European ancestry population.

Checking MR assumptions:

(i) The relevance assumption: 4 SNPs are in the gene, which fully covers the common variations of the gene in populations of European descent. Each of these variants was associated with base-line CRP. The authors did not perform the F-statistic test or show the total variance in the exposure explained by these SNPs.

(ii) The independence assumption: the authors evaluate the association of these SNPs with a range of other risk factors, including conventional risk factors and other inflammatory markers. They detect no associations between each SNPs and the risk factors tested.

(iii) The exclusion restriction assumption: the authors test for pleiotropic effects by measuring heterogeneity, using I^2 statistic. There was modest heterogeneity in the MR estimates for each of the 4 SNPs, and little heterogeneity in the MR estimates for haplotype analyses.

Regarding the MR method, the study used a two-stage least squared approach, which is appropriate for one-sample MR, to calculate risk ratios for CHD per addition of a prespecified risk allele of each SNP. Natural log risk ratios were calculated separately within each study

and were subsequently pooled across studies by random effects meta-analysis. To calculate risk ratios for CHD per 1 SD higher genetically increased CRP as determined by all 4 SNPs the authors used a hierarchical Bayesian meta-analysis framework. I^2 statistics were used to assess heterogeneity.

To check for the robustness of the MR finding, the authors conducted a range of sensitivity analyses, including using different genetic models, excluding variants or studies that deviated from Hardy-Weinberg equilibrium or were evaluated as insufficiently strong genetic instruments, using fixed effect meta-analysis models; excluding participants who reported using cardiovascular drugs (including statins) at baseline; excluding participants in clinical trials; conducting the analysis on only people of European ancestry; comparing prospective and retrospective studies; and comparing larger studies versus smaller studies.

Results

The MR findings are nicely presented using forest plots. The plots illustrated clearly the associations of each genetic variant with a wide range of risk factors, with natural log concentration of CRP and risk of CHD, the comparison between estimates of the association of circulating CRP levels and genetically increased CRP levels with the risk of CHD.

The authors found that genetically increased CRP concentrations were not associated with conventional risk factors and risk of CHD, suggesting that CRP concentration is unlikely to have a causal effect on the risk of CHD.

Strengths and limitations

Strengths:

- Selected genetic variants to cover the full range of common variability in the *CRP* gene.
- Associations of genetic variants with potential confounders were formally tested.
- Various sensitivity analyses were applied. Results from primary analysis were compared with those from sensitivity analyses.
- Interpret their findings in a broader context: compare with in vivo studies, and in vitro studies.
- With individual-level data, the study was able to conduct analyses (such as haplotype analysis) that could not be performed with summary data.

Limitations:

- Should perform formal tests to evaluate weak instrument bias, such as using F-statistics.
- Should mention the total variance in the exposure explained by the genetic instruments.

- Proxies were used for studies in which the CRP SNPs were not measured. This potentially leads to residual confounding due to unidentified effects of genotypes on other risk factors, as well as compensatory adaptation during early life to counterbalance the genetically increased CRP.
- Further research is warranted in populations of non-European ancestry, especially considering the ethnic-specific variations in the genetic architecture of CRP.

This study demonstrates the value of MR in clarifying the causal relationships between biomarkers and disease outcomes, by refuting a causal role for CRP in CHD. The study highlights the need for further research in non-European populations to account for ethnic-specific genetic variations.

2.1.9 Example 2: C reactive protein and 32 complex somatic and psychiatric outcomes

The study investigated the causal relationship between genetically proxied CRP levels and 32 complex somatic and psychiatric outcomes, showcasing the versatility of MR in exploring diverse health outcomes.

Background

Observational studies have reported the association of CRP with a wide range of common complex outcomes diseases; however, it remains controversial whether these associations are causal. To address this knowledge gap, Prins et al conducted a MR study to investigate the causal association of genetically proxied CRP with 32 complex somatic and psychiatric outcomes (284).

Methods

The study performed two-sample MR based on summary data from large-scale GWAS consortia. The authors used two genetic risk scores (GRSs) as genetic proxies for serum CRP concentrations (the exposure). The first GRS consists of 4 SNPs in the *CRP* gene, and the second GRS consists of 18 SNPs that were associated with CRP levels at the genome-wide significance level ($P < 5\times10^{-8}$) in the largest GWAS at the time of publication. The outcome of interest is CHD. Individual-level data are obtained from 47 studies in the CRP CHD Genetic Collaboration, encompassing a total of 194,418 participants, out of which 46,557 were cases. 89% of the total participants were of European descent. There are 32 outcomes of interest, categorised as either immune, cardiometabolic, or psychiatric traits and diseases. Summary statistics for the outcomes were obtained from GWAS consortia (sample size per outcome, 1,566 – 184,305, participants of European descent).

Checking MR assumptions:

(i) The relevance assumption: 4 SNPs are from the gene *CRP*, explaining about 2% of the variance in the exposure. 18 SNPs were associated with CRP levels in the largest GWAS, explaining about 5% of the variance in the exposure. F-statistics were calculated to assess the instrument strength for the GRSs.

(ii) The independence assumption: the authors have not examined the association of 18 SNPs from GWAS with potential confounders, which is one of the limitations of a two-sample MR setting.

(iii) The exclusion restriction assumption: the authors test for pleiotropic effects by measuring heterogeneity. Goodness-of-fit tests were conducted to correct both GRSs for heterogeneity. The authors did not perform advanced MR methods such as MR-Egger to test and correct for the potential pleiotropy.

Regarding the MR method, the study used the inverse variant weighted (IVW) method, which is appropriate for two-sample MR, to estimate the causal effect of CRP level on the outcomes. Proxies were used if the CRP SNPs did not have the summary statistics in the outcome GWAS dataset. For multiple testing corrections, a Bonferroni-corrected p-value was applied to test for a significant association; a p-value of 0.05 was considered a nominal association.

To check for the robustness of the MR finding, the authors conducted goodness-of-fit tests to correct for the heterogeneity. Sensitivity analysis was performed after stepwise removal of SNPs exhibiting potential pleiotropic effects until significant heterogeneity was no longer detected. To validate the detected significant association of CRP GRS_{GWAS} with schizophrenia risk, an individual-level dataset from the Psychiatric Genomics Consortium (PGC) was used to perform a follow-up analysis.

Results

The MR findings are nicely presented by type of instrumental variables (genetic risk score using either SNPs in CRP gene or SNPs from GWAS). The authors found that genetically increased CRP levels were associated with the risk of schizophrenia, after adjusting for multiple testing. This association was also significant when performing the analysis using individual-level data from schizophrenia PGC. Additionally, the study found a nominal association of CRP levels with SBP, DBP, knee osteoarthritis, and bipolar disorder.

Strengths and limitations

Strengths:

- With 2-sample MR, the study can perform the analysis on a wide range of disease outcomes, using data from large-scale GWAS.

- Large sample size can increase statistical power to detect causal associations.
- Interpret their findings in a broader context: the authors compared their findings with previous observational studies. They also explored the possible underlying pathways for the CRP-schizophrenia causal association by conducting a follow-up in silico functional pathway analysis. One possible explanation for the association may be through the CRP-interferon pathways.
- F-statistics were conducted to examine instrument strength. The total variance in the exposure that can be explained by IVs was calculated.

Limitations:

- Associations of genetic variants with potential confounders were not formally tested.
- The study did not perform advanced MR methods such as MR-Egger and weighted median to evaluate and correct for the potential pleiotropy.
- Data harmonisation was not mentioned in the study.
- Further research is warranted in populations of non-European ancestry, especially considering the ethnic-specific variations in the genetic architecture of CRP.

In this example, the broad range of outcomes analysed, and the exploration of underlying biological pathways underscores the value of MR as a powerful tool in genetic epidemiology.

2.2 Summary-based Mendelian randomisation

SMR was performed to examine the association between the expression of genes encoding targets of the BP-lowering drugs and a range of cardiovascular, diabetic and renal outcomes.

2.2.1 Summary-based Mendelian randomisation

The SMR approach selects a single most significantly associated cis-eQTL (i.e. eQTL located near the target gene) for a gene in a particular tissue as an instrument for the expression level of a gene of interest. The method then tests the association between a trait and the expression level of a gene using the effect size of the cis-eQTL from the eQTL study and the effect size of the same cis-eQTL from a GWAS. The SMR estimates are obtained using a two-stage least-squares approach (249).

If we denote z as a genetic variant (for example, eQTL), x as the expression level of a gene (exposure) and y as the outcome of interest, $\hat{b}_{zx}$ as the estimate of the genetic variant effect on gene expression (estimated by the eQTL study), $\hat{b}_{zy}$ as the estimate of the genetic variant effect on the outcome (estimated by the GWAS). The estimate of the effect of x on y is determined as:

$$\hat{b}_{xy} = \frac{\hat{b}_{zy}}{\hat{b}_{zx}} \text{ (Wald ratio)}$$

An approximate calculation of the sampling variance of $\hat{b}_{xy}$ can be obtained using the Delta method:

$$var(\hat{b}_{xy}) = \frac{b_{zy}^2}{b_{zx}^2}\left[\frac{var(\hat{b}_{zx})}{b_{zx}^2} + \frac{var(\hat{b}_{zy})}{b_{zy}^2} - \frac{2cov(\hat{b}_{zx},\hat{b}_{zy})}{b_{zx}b_{zy}}\right]$$

Where $cov(\hat{b}_{zx},\hat{b}_{zy})$ is 0 if b_{zx} and b_{zy} are estimated from independent samples. Even if b_{zx} and b_{zy} are estimated from the same sample, it remains zero if x and y are independent. In practical setting, the expected values of b_{zx} and b_{zy} can be replaced by their estimates, resulting in an approximate χ^2 test statistics of

$$T_{SMR} = \frac{\hat{b}_{xy}^2}{var(\hat{b}_{xy})} = \frac{z_{zy}^2\, z_{zx}^2}{z_{zy}^2 + z_{zx}^2}$$

Where z_{zy} and z_{zx} are the z statistics derived from the GWAS and eQTL study, respectively (249).

One issue that arises in the SMR approach is that the association observed in the SMR analysis does not necessarily mean that the selected cis-eQTL is a causal variant for both gene expression and the trait of interest. The selected cis-eQTL can be in LD with two distinct causal variants - one influences gene expression while the other affects trait variation. In order to discover the linkage scenario, the authors of the SMR method proposed a method called HEIDI (249).

2.2.2 Heterogeneity in Dependent Instruments Test

Heterogeneity in Dependent Instruments (HEIDI) test (249), which is incorporated with the SMR approach, was conducted to evaluate whether the observed association between gene expression and the outcome is due to a linkage scenario or due to a shared causal variant.

If a single causal variant (the top cis-eQTL) influences both a trait and gene expression, b_{xy} derived using any significant eQTL in LD with the causal variant remains identical. Under Hardy-Weinberg equilibrium, for any SNP_i in LD with the causal variant SNP_c, the effect of the gene expression level on the outcome using that SNP as an instrument is demonstrated as:

$$b_{xy_i} = \frac{b_{zy_i}}{b_{zx_i}} = \frac{b_{zy_c} r_{ci}\sqrt{h_c/h_i}}{b_{zx_c} r_{ci}\sqrt{h_c/h_i}} = \frac{b_{zy_c}}{b_{zx_c}} = b_{xy_c}$$

Where a subscript c represents the causal variant; a subscript i represents a SNP in LD with the causal variant; b_{xy_c} is the effect of the gene expression level on the outcome using the

causal variant, b_{xy_i} is the effect of the gene expression level on the outcome using SNPi, r_{ci} is the LD correlation between the causal variant and SNPi, h is calculated as 2f(1-f) with f being allele frequency.

Assessing if observing an association between gene expression and outcome is due to a linkage scenario is equivalent to testing whether there is a difference between b_{xy_c} and b_{xy_i}. If $d_i = b_{xy_i} - b_{xy_c}$, the test is equivalent to whether d_i is equal to zero for all the significant cis-eQTLs except for the causal variant. If we have:

$$\hat{d} = \{\hat{d}_1, \hat{d}_2, \dots, \hat{d}_n\},\ \hat{d} \sim \text{MVN (d, V)}$$

where n is total number of significant eQTLs in the cis-eQTL region, d is the mean vector, and V is the covariance matrix with the covariance between each pair of elements being calculated as:

$$cov(\hat{d}_i\, \hat{d}_j) = cov(\hat{b}_{xy_i}, \hat{b}_{xy_j}) - cov(\hat{b}_{xy_i}, \hat{b}_{xy_c}) - cov(\hat{b}_{xy_j}, \hat{b}_{xy_c}) + var(\hat{b}_{xy_c})$$

$$cov(\hat{b}_{xy_i}, \hat{b}_{xy_j}) = \frac{r_{ij}}{b_{zx_i} b_{zx_j}} \sqrt{var(\hat{b}_{zy_i}) var(\hat{b}_{zy_j})} + b_{xy_i} b_{xy_j} \left(\frac{r_{ij}}{z_{zx_i} z_{zx_j}} - \frac{1}{z^2_{zx_i} z^2_{zx_j}}\right)$$

where r_{ij} is the LD correlation between SNPi and SNPj.

Under the null hypothesis that there is no heterogeneity, where d = 0, a vector of standard normal variables was created, using z-score formula:

$$z_d = \{z_{d_1}, z_{d_2}, \dots, z_{d_m}\}, with\ z_{d_1} = \hat{d}_i / \sqrt{var(\hat{d}_i)} \text{ and } z_d \sim MVN(0, R)$$

Where R is the correlation matrix with the covariance between each pair of elements being calculated as:

$$r\left(z_{d_i}, z_{d_j}\right) = cov(\hat{d}_i, \hat{d}_j) / \sqrt{var(\hat{d}_i) var(\hat{d}_j)}$$

HEIDI test statistic was constructed to test against d = 0, being demonstrated as:

$$T_{HEIDI} = \sum_i^n z^2_{d_i}$$

HEIDI test is a quadric form of standard normal variables, with the distribution can be estimated using the Satterthwaite or Saddlepoint method (249).

2.2.3 Genetic instrument selection

cis-eQTL, identified as genetic variants located within 1 Mb on either side of the gene and significantly associated with changes in gene expression in a particular tissue, were selected as the genetic instruments for the expression of the genes encoding BP-lowering drug

targets. The cis-eQTLs were obtained from the Genotype-Tissue Expression (GTEx) version 8 (285), using the p-value threshold of 5×10^{-8} to filter the significant cis-eQTL. The top significant cis-eQTL was selected as a genetic instrument for the expression level of the gene in a particular tissue.

For genetic validation and positive control, SMR analysis was conducted to test whether changes in gene expression reflect a BP-reducing association with antihypertensive drugs. Only genes whose expression was associated with SBP ($p < 0.05$) were included for further analysis.

2.2.4 Exposure data

Summary data for the cis-eQTL were extracted from Genotype-Tissue Expression (GTEx) version 8. GTEx is an extensive public database that provides insights into human gene expression and regulation as well as its association with genetic variation across multiple diverse tissues and individuals (286). The GTEx project collected samples from up to 54 non-diseased tissue sites, from nearly 1,000 deceased individuals. Whole Genome Sequencing (WGS) technique was used to genotype densely all individuals to evaluate genetic variation present within their genomes. RNA sequencing (bulk RNA-seq) was used to assess gene expression of each tissue (286).

cis-eQTLs were identified as genetic variants that had a significant correlation with changes in the expression of genes located nearby. The eQTL analysis of GTEx v8 data was performed on 15,201 RNA-sequencing samples across 49 tissues of 838 post-mortem donors (285). These 49 tissues had at least 70 individuals, each with both RNA sequence and genotype data derived from whole genome sequencing. Out of the 838 donors, 715 (representing 85.3%) were of European American descent, 103 (12.3%) were African American, and 12(1.4%) were Asian American, with 16 (1.95) identified as Hispanic or Latino ethnicity. In terms of gender distribution, 557 (66.4%) donors were male and 281 (33.5%) were female (285). The analysis of the v8 GTEx data provides a catalogue of genetic regulatory variants affecting gene expression and splicing in cis and trans across 49 tissues and describes patterns and mechanisms of tissue- and cell-type specificity of genetic regulatory effects.

Cis-eQTL data from GTEx v8 in SMR binary format were extracted from yanglab.westlake.edu.cn/data/SMR/GTEx_V8_cis_eqtl_summary.html (249).

2.2.5 Outcome data

The outcomes used in the SMR analysis are identical to those in the two-sample MR analysis, including atrial fibrillation, heart failure, coronary artery disease, myocardial infarction, ischemic stroke, chronic kidney disease, and type 2 diabetes.

2.2.6 Format dataset for SMR

The SMR method requires a reference sample in PLINK format for linkage disequilibrium estimation, GWAS summary-level data for the outcome, and summary-level data from an eQTL study in binary format.

Genotype data from the 1000 Genomes samples of European ancestry (n = 503) (in PLINK binary format) were used as a reference for linkage disequilibrium estimation (287). The 1000 Genomes Project has been set to provide a comprehensive catalogue of human genetic variation, achieved through the application of whole-genome sequencing across a varied cohort of individuals from multiple populations. The final phase of the 1000 Genomes Project was phase 3, wherein the genomes of 2,504 individuals from 26 populations were reconstructed. These populations were categorised into five groups: African, East Asian, European, South Asian, and admixed American (287). The genotype dataset from the 1000 Genomes phase 3 was downloaded from PLINK 2.0 (https://www.cog-genomics.org/plink/2.0/resources) (288).

GWAS summary-level data for the outcome include the following columns: SNP rsID (SNP), the effect allele (A1), the other allele (A2), frequency of the allele (freq), effect size (se), standard error (se), p-value (p), and sample size (n). Sample size information is not used for SMR as well as the HEIDI test, so it can be replaced by NA if not available. SNPs with discrepant allele frequencies between two datasets will be removed, based on the allele frequency information (249).

eQTL summary-level data are in binary format, including three separate files: SNP information, probe information, and eQTL summary statistics (249).

2.2.7 Perform SMR

SMR and HEIDI tests were performed using the SMR software version 1.3.1 (249). I use the following default settings: cis-eQTLs selected using $p_{eQTL} < 5 \times 10^{-8}$; minor allele frequency (MAF) > 0.01; eQTLs included for HEIDI test have a minimum p-value of 1.57×10^{-3} (equivalent to a chi-squared value of 10); SNPs in the cis-eQTL region with $P_{eQTL} > 1.6 \times 10^{-3}$ (equivalent to $\chi_1^2 < 10$) were also removed from the HEDI test to avoid weak instrumental variables (249); eQTLs in very high or low LD ($r^2 > 0.9$ or $r^2 < 0.05$) with the top associated

eQTL were excluded in the HEIDI test to avoid the issue of collinearity; the minimum number of eQTLs in the HEIDI test is 3 and the maximum number of eQTLs is 20; cis-eQTLs were selected within 1 MB window on either side of the probe.

2.2.8 Interpret result

SMR estimates were presented as a beta coefficient with a 95% confidence interval (95% CI) per 1 unit increase in the expression of the genes in a particular tissue. The p-value threshold was adjusted using Bonferroni correction for multiple testing. A HEIDI test p-value less than 0.01 is indicative of an association due to a linkage scenario (120, 249).

2.2.9 Replication analysis

For the significant SMR associations, the SMR analysis was replicated using summary statistics obtained from the eQTLGen consortium (289). The eQTLGen consortium aims to explore the genetic architecture of blood gene expression. The consortium performed a large-scale meta-analysis on 31,684 human blood samples drawn from 37 cohorts. Most of the participants (~95%) were of European descent. Gene expression data were adjusted for up to 25 principal components that are not linked to genetic factors, in order to account for unrecognized technical confounders and biological confounders. Cis-eQTL analysis was conducted on the 31,684 blood samples, with cis-eQTL being determined as variants within 1 MB of the gene centre. The consortium identified cis-eQTL for 16,987 genes, which account for 88.2% of the genes expressed in blood. The false discovery rate of 0.05 was used to determine the cis-eQTL effect. The authors reported that 92% of lead cis-eQTL SNPs were located within 100 kb of the gene, with stronger cis-eQTL being physically closer to the gene (289). The cis-eQTL results in SMR binary format from eQTLGen Consortium phase I were downloaded from https://www.eqtlgen.org/cis-eqtls.html (289).

Colocalisation analysis was conducted to investigate if the observed association is due to a linkage scenario or a shared causal variant.

2.2.10 Example: Antihypertensive drug classes and psychiatric disorders

Background: Observational studies have reported associations of antihypertensive drug classes with psychiatric disorders and suggest that any effect could be independent of their BP-lowering effect. As observational studies are subjected to confounding bias and reverse causation, further study is needed to understand if the associations observed are due to causality. Chauquet et al. conducted an SMR study to investigate this research question (120).

Methods:

Two-sample SMR study was performed to examine the causal association of antihypertensive drug classes with schizophrenia, bipolar disorder, and major depressive disorder. The expression levels of the drug target genes in blood were used as proxies for the drug classes. The top associated cis-eQTL for the drug target genes were used as instrumental variables for the SMR analysis. The authors obtained summary data for the outcomes from the largest GWAS consortia at the time of publication (schizophrenia: 40,675 cases and 64,643 controls; bipolar disorder: 20,352 cases and 31,358 controls; major depressive disorder: 135,358 cases and 344,901 controls). Summary data for the gene expression levels in blood were obtained from the eQTLGen consortium (n = 31,684). Most participants in those studies predominantly have European ancestry.

MR assumptions:

(i) The relevance assumption: The authors selected cis-eQTL for the gene that has highest p_{eQTL} value. The p_{eQTL} threshold for the selection is not mentioned in the article. F-statistics were used to evaluate the strength of the instruments.

(ii) The independence assumption: The association of the selected cis-eQTL with potential confounders was not examined, which is one of the limitations of a two-sample MR setting.

(iii) The exclusion restriction assumption: the HEIDI test was conducted to test whether the observed association was due to a linkage scenario. The $P_{HEIDI} \geq 0.01$ was used to rule out the linkage.

To correct for multiple testing, the authors applied a Bonferroni-corrected p_{SMR} to discover a significant association, while a p_{SMR} of 0.05 was considered a nominal association. A range of sensitivity analyses was applied to examine the robustness of the SMR findings, including colocalisation analysis, simulation of the effect of selection bias, ancestry-specific analyses, and conducting SMR analysis on additional datasets (if applicable). The study also investigated the association of SBP with the disease outcomes.

Results:

The study found that a decreased *ACE* gene expression in blood was associated with lower SBP but increased risk of schizophrenia disorder. They did not detect an association between genetically predicted SBP and the risk of schizophrenia. The study suggested that the ACE-schizophrenia association was likely independent of its effect on SBP.

Strengths and limitations:

Strengths:

- With a 2-sample MR setting, the study investigated the effects of different antihypertensive drug classes with multiple outcomes.
- Summary data from large-scale GWAS and eQTL studies can increase the statistical power.
- Using cis-eQTL as genetic instruments for gene expression can minimise the potential for horizontal pleiotropy.
- Interpreting their findings in a broader context: the authors compared their findings with previous observational and genetic studies.
- The authors conducted several sensitivity analyses to assess the robustness of the findings.

Limitations:

- It was assumed that alterations in gene expression would reflect alterations in protein levels and activity, but this may not always hold true.

The effect sizes obtained from SMR may not be equivalent to those of drug effects. This is due to differences in the duration of exposure, the relevance of tissues under the study, and the pharmacokinetic-pharmacodynamic profile of the drug.

2.3 Colocalisation analysis

For any significant association between gene expression and the outcome, the coloc method (255) was conducted to examine whether the two traits shared a causal variant in the genomic region. The coloc method uses the Bayesian framework to systematically explore putative causal configurations, where each configuration can be assigned to one of five hypotheses:

H_0: No causal variants for either trait

H_1: Causal variant for trait 1, not for trait 2

H_2: Causal variant for trait 2, not for trait 1

H_3: Distinct causal variants for trait 1 and trait 2

H_4: One shared causal variant for trait 1 and trait 2

In this framework, the test for colocalisation can be re-formulated as evaluating the support for all configurations in hypothesis H_4. The coloc method works under the assumption that there is at most one causal variant for each trait. It also requires the definition of prior probabilities, which are the prior probability for a genetic variant being associated with trait 1 (p_1), the prior probability for a genetic variant being associated with trait 2 (p_2), and the prior probability for a variant being associated with both traits (p_{12}). The posterior probability for a shared causal variant is calculated as:

$$PP_4 = P(H_4|D) = \frac{\frac{P(H_4|D)}{P(H_0|D)}}{1 + \sum_{i=1}^{4} \frac{P(H_i|D)}{P(H_0|D)}}$$

$$\frac{P(H_1|D)}{P(H_0|D)} = p_1 \times \sum_j ABF_j^{(1)}$$

$$\frac{P(H_2|D)}{P(H_0|D)} = p_1 \times \sum_j ABF_j^{(2)}$$

$$\frac{P(H_3|D)}{P(H_0|D)} = p_1 p_2 \times \sum_{j \neq k} ABF_j^{(1)} ABF_k^{(2)}$$

$$\frac{P(H_4|D)}{P(H_0|D)} = p_{12} \times \sum_j ABF_j^{(1)} ABF_j^{(2)}$$

Where ABF indicates a Bayes factor for each SNP and each trait 1 and 2 (e.g. $ABF_j^{(1)}$ indicates Bayes factor for SNP_j and trait 1). The Bayes factor can be computed for each variant from p-value, or estimated regression coefficient and its' variance, using the Approximate Bayes Factor formula proposed by Wakefield (290).

2.3.1 Selecting SNPs for the colocalisation analysis

For a gene whose expression level in a particular tissue was associated with the outcome, a 1 MB window on either side of the gene was used to select SNPs for the analysis. SNPs must have summary data available in both the eQTL study and GWAS study for the outcome. All identified SNPs will be included in the analysis without any exclusions based on p-value, MAF or other filtering criteria. This is because the coloc method systematically assesses every potential combination of shared and non-shared causal variants to determine their respective levels of support. Without the availability of all SNPs, this comprehensive enumeration cannot be performed (255).

2.3.2 Coloc dataset

A minimum coloc dataset requires the estimated effect size and its uncertainty (i.e. the standard error squared). It also requires a prior on the true effect sizes at causal variants, which can be estimated using the trait standard deviation. Coloc can estimate the trait standard deviation using MAF and sample size information (255).

For each dataset, SNP rsID, MAF, sample size, beta, standard error squared, the trait standard deviation (if available), and type of the trait (quantitative or case-control) were supplied for the coloc analysis.

2.3.3 Colocalisation analysis

SNPs between datasets were matched between two datasets using their rsIDs. Coloc v5.2.1 R package (255) was then performed on the two datasets, by using function coloc.abf(). Default priors were used for the analysis: $p_1 = 10^{-4}$, $p_2 = 10^{-4}$, $p_{12} = 10^{-5}$. In which, p_1 is the prior probability for a variant being associated with trait 1, p_2 is the prior probability for a variant being associated with trait 2, and p_{12} is the prior probability of a variant being associated with both traits.

2.3.4 Interpret the result

The analysis provides the posterior probabilities for 5 scenarios (H_{0-4}). A large posterior probability for hypothesis 4 ($PP_4 > \sim0.8$) supports the model with a shared causal variant affecting both traits. Whereas a large posterior probability for hypothesis 3 ($PP_3 > \sim0.8$) supports the model with two independent causal SNPs associated with each trait, indicating confounding by linkage disequilibrium. A moderate PP_4 but low PP_3 would indicate a lack of power to detect colocalisation, but no evidence for the linkage scenario.

Chapter 3 ANTIHYPERTENSIVE GENES AND CARDIOVASCULAR OUTCOMES

3.1 Introduction

Elevated blood pressure (BP) is identified as the most common modifiable risk factor for cardiovascular disease (CVDs) and related impairments. Globally, high BP affects over a billion individuals, contributing to over a fifth of CVD cases (291). According to the World Health Organization (WHO), the number of individuals with hypertension (BP of 140/90 mmHg or higher, or being on medication for hypertension) increased twofold between 1990 and 2019, rising from 650 million to 1.3 billion (292). Hypertension imposes a substantial health impact across a wide range of diseases, including coronary heart disease, stroke, myocardial infarction, heart failure, peripheral artery disease, and renal failure (17, 18). A meta-analysis of individual data for 1 million adults in 61 prospective studies shows that the mortality risk from stroke, ischemic heart disease, or other vascular disease is increased by two-fold per 20 mmHg increase in SBP or 10 mmHg increase in DBP (5). Despite the prevalence of hypertension and the health burden exerted by it, nearly half of those affected remain unaware of their condition, underscoring the need for effective management strategies to mitigate the resultant health complications (292).

Evidence from RCTs underscores the efficacy of BP management; a reduction of 10 mmHg in SBP significantly lowers the risk of major CVD events, including CAD (relative risk reduction of 5% to 20%) and stroke (relative risk reduction of 25% to 30%) (30). The pharmacological treatment of hypertension is grounded in robust evidence, supported by a large number of outcome-oriented RCTs. Angiotensin-converting enzyme inhibitors (ACEi), angiotensin II receptor blockers (ARBs), calcium channel blockers (CCBs), thiazides, and beta-blockers (BBs) are major antihypertensive drug classes whose effectiveness in controlling BP has been proven in RCTs. Furthermore, their efficiency in reducing the risk of CVDs and hypertension-related mortality has also been reported in RCTs (40). These findings emphasize BP management as a critical intervention to reduce the CVD burden. However, despite the proven effectiveness of the antihypertensive drug classes in lowering BP, gaps remain in our understanding of their broader health impacts. Particular drugs could potentially have adverse drug effects on outcomes that have not been thoroughly explored in RCTs (120). For example, aliskiren, a renin inhibitor, has a potent and long-acting effect in reducing BP, whether used as a standalone treatment or in conjunction with a diuretic or a CCB. However, two RCTs, which were conducted to investigate the effect of aliskiren on hypertension-mediated organ damage and cardiovascular outcomes, were prematurely

terminated due to a rise in the incidence of adverse events (141). Sitaxentan, an endothelin receptor antagonist, was approved for the treatment of pulmonary arterial hypertension. Several years later this drug was withdrawn from the market worldwide by the licence holder due to fatal cases associated with hepatic injury (142). Other antihypertensive drugs that have been withdrawn from the market or have their use restricted because of safety concerns include debrisoquine (143), guanethidine (144), mibefradil (145), reserpine (146), methyldopa (147), doxazosin (148), fenoldopam (149), and omapatrilat (150). On the other hand, some antihypertensive drug classes may exhibit beneficial effects (19) that extend beyond their primary function of reducing BP, which offers opportunities for drug repurposing. Evidence from clinical trials suggests that ACEi and ARBs may have a beneficial effect on reducing the risk of new-onset T2D. A meta-analysis of 12 RCTs on 116,220 patients showed that ACEi and ARBs were associated with a reduced risk of newly diagnosed diabetes by 27% (95% CI, 16% - 37%) and 23% (95% CI, 17% - 29%), respectively. The pooled effect analysis (either by ACEi or ARB) resulted in a 25% risk reduction (95% CI, 18% - 31%) (293). In a RCT involving 9,306 patients with impaired glucose tolerance and either established CVD or cardiovascular risk factors, a 5-year treatment with valsartan significantly reduced the risk of newly diagnosed diabetes, when compared with placebo (294). In another RCT, which involves 5269 patients with impaired fasting glucose levels or glucose tolerance, using ramipril for 3 years significantly increases regression to normoglycemia (42). One possible mechanism behind the protective effect of ACEi and ARBs on T2D is that these medications may improve insulin sensitivity and blood flow to skeletal muscles, which are pivotal in insulin-mediated glucose uptake and serve as the endpoint determinant of this process. By expanding the vascular interface with skeletal muscle tissue, these drugs can effectively increase the area for glucose exchange, thereby contributing to managing blood glucose levels (295). Current guidelines recommend the use of ACEi or ARBs for the management of hypertension in patients with T2D (296). These medications are also recommended for patients with diabetic nephropathy, as part of a broader strategy to delay the progression of the kidney disease in diabetic patients.

While RCTs are invaluable in establishing the efficacy of antihypertensive medications, they have certain limitations. RCTs are resource-intensive and often have limited follow-up periods, which restricts their capacity to evaluate the long-term safety and efficacy of antihypertensive drugs (297) (298). Also, RCTs typically involve selective patient populations such as individuals with certain comorbidities or demographic characteristics, leading to the constraint in extrapolating the findings to the broader and more diverse hypertensive population. As RCTs are typically designed to test a particular hypothesis about drug efficacy and safety, they may overlook the beneficial effects of antihypertensive drugs on

outcomes other than BP (297) (298). Moreover, RCTs may not have sufficient statistical power to uncover rare but clinically significant adverse effects due to their relatively smaller sample sizes. Treatment guidelines frequently rely on extrapolated evidence from RCTs, which may not fully capture the potential benefits or adverse effects of antihypertensive therapies. Therefore, the broader implications of antihypertensive therapy, including potential beneficial effects beyond BP reduction and unexplored adverse outcomes, remain under investigation (19).

Observational studies have sought to address these gaps, yet their susceptibility to confounding, reverse causation, and measurement errors limits their utility (157). In contrast, MR, leveraging genetic variations as natural experiments, offers a novel approach to circumvent these limitations (157). MR is a form of instrumental variable analysis, which uses genetic variants as instruments to investigate the causal association between a modifiable risk factor and an outcome (157). When the exposure is a drug target, MR can be applied to evaluate the effect of the drug target across a range of outcomes by leveraging genetic variants located near the gene encoding the drug target (221). As MR leverages large-scale genetic data to infer the lifelong impact of genetically proxied drug targets, it can elucidate the long-term safety and rare adverse effects of antihypertensive drugs. MR can be applied to diverse populations, thereby providing insights into the drug effects across different ethnic and demographic groups. Furthermore, by investigating the effect of genetic proxies of drug targets on a wide range of outcomes, MR has the potential to illuminate the unintended effects of drugs, which can help to identify potential repurposing opportunities or unexpected beneficial effects of antihypertensive drugs on outcomes other than BP, such as diabetes risk, cognitive decline, or cancer risk (120-122).

MR has been applied in the context of lipid-lowering therapies. Increased risk of T2D is one of the adverse effects of statins. A meta-analysis of RCTs including 129,170 individuals followed up for 4.2 years demonstrated the causal effect of statins on the increased risk of new-onset T2D. MR studies used LDL-associated genetic variants that are located near the gene encoding the HMGCR protein, a protein target of statins, as genetic proxies for the effect of the drug. Genetically proxies for the inhibition of HMGCR were associated with lower LDL levels, higher body weight, higher plasma insulin and higher plasma glucose concentration. This finding underscores the role of HMGCR inhibition in increasing T2D risk (225).

In the context of BP-lowering treatments, the MR approach has been applied to explore the unintended effects of antihypertensive medications, utilising genetic proxies for SBP and antihypertensive drug classes. Gill D. et al (122) conducted a MR study to estimate the effects of three major antihypertensive drug classes (ACEi, BBs, and CCBs) on CAD and

stroke risk, leveraging genetic variants in the genes encoding protein targets for the corresponding drug class. Genetically proxied CCBs showed a protective effect against both stroke and CAD. Genetically proxied BBs demonstrated a protective effect against CAD, while genetically proxied ACEi were protective against stroke. MR estimates in their study were comparable to the results from RCTs, demonstrating the potential of using genetic variants to complement evidence from RCTs and observational studies (122). Using a similar MR approach, other MR studies were conducted to explore the unintended effects of antihypertensive drug classes, mainly focusing on first-line antihypertensive drugs. Genetically predicted CCBs were associated with a reduced risk of HF and AF, but increased risk of diverticulosis (244, 299). Genetically predicted BBs were also associated with a reduced risk of HF and AF (244, 299), but an increased risk of T2D (245). By contrast, genetically proxied ACEi showed protective effects against T2D (245) but might have a negative impact on schizophrenia risk (120) and colorectal cancer (121).

This chapter aims to utilise the MR approach to systematically explore the causal effects of genetically proxied systolic blood pressure and antihypertensive drug classes on a broad range of cardiovascular, diabetic, and renal outcomes. By including both established and newly developed drugs, as well as first-line and non-first-line therapies, this research seeks to uncover potential drug repurposing opportunities and highlight adverse effects not previously detected in RCTs. Through this comprehensive analysis, I aim to refine hypertension treatment guidelines and improve patient care by providing a clearer understanding of the full spectrum of antihypertensive drug effects.

3.2 Method

Study Overview: I designed my study to explore the causal relationships between antihypertensive drug classes, decreased SBP, and a range of cardiovascular, diabetic, and renal outcomes. I utilised publicly available summary-level GWAS data, ensuring all datasets adhered to ethical approval and participant consent standards.

Data Selection: GWAS datasets were selected based on the largest sample size and relevance to the outcomes of interest. I prioritised datasets with comprehensive phenotype definitions and those that provided extensive coverage of the European-ancestry population, from which the genetic instruments for SBP and drug targets were primarily derived.

3.2.1 Study design

The study design is illustrated in **Figure 3.1**. The antihypertensive drug classes and their active ingredients were identified using the British National Formulary (BNF) (269) and the DrugBank database (https://go.drugbank.com/; version 5.1.10). Using summary statistics

from GWAS, my study conducted MR analysis to investigate the causal effects of genetically decreased SBP and genetically proxied antihypertensive drug classes on a range of outcomes related to hypertension: CAD, MI, HF, AF, ischemic stroke, CKD, and T2D (17, 300-302). The MR analysis was also conducted on glomerular filtration rate estimated by serum creatinine (eGFRcrea), glomerular filtration rate estimated by cystatin C (eGFRcys), glycated haemoglobin (HbA1c), and fasting glucose, as sensitivity analysis for CKD and T2D.

A Bonferroni-corrected p-value less than 0.0007 that accounted for multiple testing (0.05 / 11 (drug classes & SBP) × 7 (outcomes)) was used to indicate statistical significance in the MR analyses. Whereas a p-value less than 0.05 but greater than 0.0007 was considered suggestive evidence for a significant association between exposure and outcome.

3.2.2 Identification of Drug Target Genes

Following an exhaustive review of the BNF (269), I identified 17 antihypertensive drug classes and their active ingredients. Criteria for drug class selection included clinical relevance, diversity of mechanism of action, and availability of genetic data related to their targets. 17 antihypertensive drug classes include angiotensin-converting enzyme inhibitors (ACEi), angiotensin II receptor blockers (ARBs), calcium channel blockers (CCBs), beta-adrenoceptor blockers (BBs), thiazides and related diuretics (TZDs), alpha-adrenoceptor blockers, adrenergic neuron blocking drugs, centrally acting antihypertensive drugs, loop diuretics, potassium-sparing diuretics (PSDs), aldosterone antagonists, renin inhibitors, endothelin receptor antagonists (ERAs), phosphodiesterase type 5 (PDE5) inhibitors, soluble guanylate cyclase (sGC) stimulators, and vasodilator antihypertensives (**Table 3.1**). I paid special attention to emerging therapies such as Zilebesiran (60), an investigational RNA interference agent that inhibits the production of hepatic angiotensinogen, due to its novel mechanism and potential for hypertension management. Using the DrugBank database, pharmacologically active protein targets of each drug and the corresponding genes were identified. I obtained 57 pharmacologically active drug target genes (**Table 3.1**).

Figure 3.1 Flowchart illustrating the study design.

SBP, systolic blood pressure; GWAS, genome-wide association study; SNP, single nucleotide polymorphism;; MR, Mendelian randomization; CAD, coronary artery disease; MI, myocardial infarction; AF, atrial fibrillation; HF, heart failure; CKD, chronic kidney disease; T2D, type 2 diabetes; FG, fasting glucose; IVW, inverse variant weighted; (*) indicates a blood pressure GWAS meta-analysis (115).

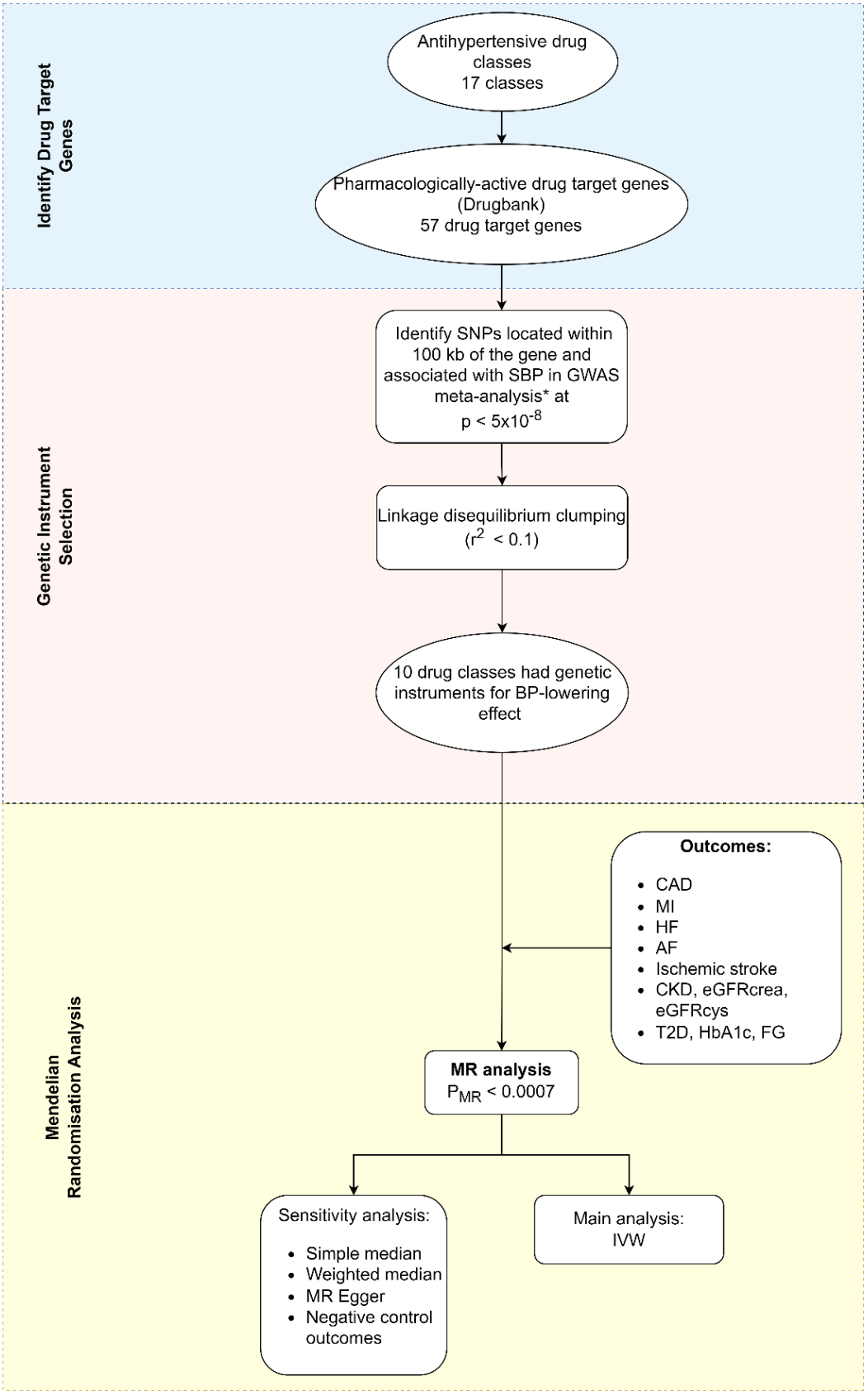


Table 3.1 Antihypertensive Drug Classes, Drug Names, and Corresponding Drug Target Genes.

AGTi, angiotensinogen inhibitors; ACEi, angiotensin-converting enzyme inhibitors; ARBs, angiotensin II receptor blockers; CCBs, calcium channel blockers; β-blockers, β-adrenoceptor blockers; α-blockers, alpha-adrenoceptor blockers; TZDs, thiazides and related diuretics; ANBs, adrenergic neuron blocking drugs; CAAs, centrally acting antihypertensive drugs; PSDs, potassium-sparing diuretics; ERAs, endothelin receptor antagonists; PDE5 inhibitors, phosphodiesterase-5 inhibitors; sGC stimulators, soluble guanylate cyclase stimulators. Empty columns indicate the absence of DrugBank data regarding the target gene for the specified drug.

Drug class	Drug name	Target gene
AGTi	zilebesiran	*AGT*
ACEi	captopril	*ACE*
	enalapril	*ACE*
	lisinopril	*ACE*
	perindopril	*ACE*
	ramipril	*ACE*
	quinapril	*ACE*
	cilazapril	*ACE*
	fosinopril	*ACE*
	trandolapril	*ACE*
	moexipril	*ACE*
	imidapril	
ARBs	losartan	*AGTR1*
	eprosartan	*AGTR1*
	valsartan	*AGTR1*
	irbesartan	*AGTR1*
	candesartan	*AGTR1*
	telmisartan	*AGTR1, PPARG*
	olmesartan medoxomil	*AGTR1*
	azilsartan medoxomil	*AGTR1*
CCBs	amlodipine	*CACNA1C, CACNA1I*
	diltiazem	*CACNA1C, CACNG1*
	felodipine	*CACNA1C, CACNA2D1, CACNB2, CACNA1D, CACNA1S*
	isradipine	*CACNA1C, CACNA2D1, CACNB2, CACNA1H, CACNA2D2, CACNA1D, CACNA1S*
	lacidipine	*CACNA1C, CACNA1D, CACNA1F, CACNA1S, CACNB1, CACNB2, CACNB3, CACNB4, CACNA1A*
	lercanidipine	*CACNG1*
	nicardipine	*CACNA1C, CACNB2, CACNA2D1, CACNA1D*
	nifedipine	*CACNA1C, CACNA1D, CACNB2, CACNA1S*
	nimodipine	*CACNA1C, CACNA1D, CACNA1F, CACNB1, CACNA1S, CACNB2, CACNB3, CACNB4*
	nisoldipine	*CACNA1C, CACNA2D1, CACNB2, CACNA1D, CACNA1S*

	verapamil	*CACNA1C, CACNA1B, CACNA1A, CACNA1G, CACNA1H, CACNA1S, CACNA1D, CACNA1F, CACNB1, CACNB2, CACNB3, CACNB4*
β-blockers	acebutolol	*ADRB1*
	atenolol	*ADRB1*
	bisoprolol	*ADRB1*
	celiprolol	*ADRB1, ADRB2*
	esmolol	*ADRB1*
	landiolol	
	metoprolol	*ADRB1*
	nadolol	*ADRB1*
	nebivolol	*ADRB1*
	oxprenolol	*ADRB1*
	pindolol	*ADRB1, ADRB2*
	propranolol	*ADRB1*
	sotalol	*ADRB1, ADRB2, KCNH2*
	timolol	*ADRB1, ADRB2*
β- and α-blockers	labetalol	*ADRB1, ADRB2, ADRA1A, ADRA1B, ADRA1D*
	carvedilol	*ADRA1A, ADRA1D, ADRA1B, ADRB1*
TZDs	bendroflumethiazide	*SLC12A3, KCNMA1*
	chlorothiazide	*SLC12A3, CA1, CA2*
	chlortalidone	*SLC12A1, CA1*
	cyclopenthiazide	
	hydrochlorothiazide	*SLC12A3, KCNMA1*
	indapamide	*SLC12A3*
	metolazone	*SLC12A3*
	polythiazide	*SLC12A3*
	xipamide	
α-blockers	prazosin	*ADRA1A, ADRA1D, ADRA1B*
	indoramin	*ADRA1A*
	doxazosin	*ADRA1A, ADRA1D, ADRA1B*
	phenotalamine	*ADRA1A, ADRA2A, ADRA1B, ADRA1D*
	terazosin	*ADRA1A, ADRA1B, ADRA1D*
	phenoxybenzamine	*ADRA1A, ADRA2A*
ANBs	guanethidine monosulfate	*SLC6A2*
CAAs	clonidine	*ADRA2A, ADRA2B, ADRA2C*
	guanfacine	*ADRA2A, ADRA2B*
	moxonidine	*ADRA2A, NISCH*
	methyldopa	*ADRA2A*
Loop diuretics	furosemide	*SLC12A1*
	bumetanide	*SLC12A1, SLC12A2, SLC12A4, SLC12A5*
	torasemide	*SLC12A1, SLC12A2*

Aldosterone antagonists	spironolactone	*NR3C2*
	eplerenone	*NR3C2*
PSDs	amiloride	*SCNN1A, SCNN1B, SCNN1G, SCNN1D*
	triamterene	*SCNN1G, SCNN1A, SCNN1B*
Renin inhibitors	aliskiren	*REN*
ERAs	ambrisentan	*EDNRA*
	bosentan	*EDNRA, EDNRB*
	macitentan	*EDNRA*
	sitaxentan	*EDNRA*
PDE5 inhibitors	sildenafil	*PDE5A, NOS3*
	tadalafil	*PDE5A*
sGC stimulators	riociguat	*GUCY1A2, GUCY1A3*
	vericiguat	*GUCY1B1*
Vasodilator antihypertensives	hydralazine	*AOC3*
	minoxidil	*KCNJ1*
	Iloprost	*PTGIR, PTGER1*

3.2.3 Genetic Variants Selection

3.2.3.1 Genetic Instruments for Drug Class Effect Analysis

After obtaining drug target genes for each antihypertensive drug class, I identified genetic instruments to proxy the effect of lowering BP through the drug target as SNPs that were located within 100 kb on either side of the corresponding gene and associated with SBP at genome-wide significance level ($p < 5\times10^{-8}$). Genetic association estimates for SBP were extracted from a GWAS meta-analysis of 757,601 European-ancestry individuals from the UK Biobank and the International Consortium of Blood Pressure (ICPP) (115). For antihypertensive drug classes that have protein targets encoded by several different genes, such as CCBs, PDE5 inhibitors, and loop diuretics, I selected genetic variants related to the individual genes and combined their data to examine the effect of the drug target perturbation (122). This approach intended to capture the most relevant genetic influences on drug target activity and subsequent BP modulation. The selection was further refined by linkage disequilibrium clumping to ensure instrument independence. SNPs were clumped to a LD threshold of $r^2 < 0.1$ using the 1000 Genomes European reference panel (272). The R^2 and F-statistics were calculated to estimate the proportion of variation in SBP explained by the genetic variant and the strength of each variant, respectively (280). Genetic instruments were conventionally considered to be sufficient if the corresponding F-statistic was greater than 10.

3.2.3.2 Genetic Instruments for Blood Pressure

For a broader analysis of BP effects, SNPs associated with SBP across the genome were selected, adhering to stringent linkage disequilibrium criteria to minimize bias. Instrument strength was quantified using R^2 and F-statistics, with a conventional threshold for sufficiency set at F>10, aligning with established MR methodologies. MR analysis with genetically predicted SBP as the exposure was performed to see if BP changes potentially mediate the significant causal association of the drug target perturbation on the outcome. Genetic instruments for SBP were selected as SNPs that were associated with SBP at p-value < 5×10^{-8} throughout the whole genome in the GWAS meta-analysis of 757,601 European-ancestry individuals from the UK Biobank and the ICPP (115). Given the larger number of SNPs identified, a more stringent LD threshold $r^2 < 0.001$ was applied to clump the SNPs to reduce potential bias from linkage disequilibrium. The R^2 and F-statistics were calculated to estimate the instrument strength.

3.2.4 Outcome Data

I obtained GWAS summary data for the outcomes from the largest and most comprehensive studies available, focusing on those with substantial European ancestry representation to align with the exposure data. Outcomes were chosen based on their clinical significance and relevance to hypertension and antihypertensive drug effects, encompassing both primary diseases of interest and relevant biomarkers for sensitivity analysis. The primary outcomes include AF (60,620 cases and 970,216 controls of European ancestry) (275), HF (47,309 cases and 930,014 controls of European descent) (276), CAD (181,522 cases and 984,168 controls, with more than 95% of European ancestry) (273), MI (43,676 cases and 128,199 controls, 77% of European ancestry) (274), ischemic stroke (62,100 cases and 1,234,808 controls of European descent) (277), CKD (41,395 cases and 439,303 controls of European ancestry) (303), and T2D (80,154 cases and 853,816 controls of European descent) (279). Secondary outcomes for sensitivity analysis include eGFRcrea (n = 1,004,040; European ancestry) (278), eGFRcys (n = 460,826; European ancestry) (278), HbA1c (n = 146,806; European ancestry) (304), and fasting glucose (n = 200,622; European ancestry) (305) (**Table 3.2**). Further details on population characteristics, trait definitions, and GWA analyses were in the cited original publications.

Table 3.2 Genome-wide association studies used in Mendelian randomisation analyses.

Trait	Population	Sample size	Unit	Adjustments
Systolic blood pressure (115)	European	757,601	mmHg	Age, age squared, sex, BMI, genotyping chips, antihypertensive medication
Atrial fibrillation (275)	European	1,030,836 (60620 cases)	log-transformed odds	Age, sex, study-specific covariates, principal components
Heart failure (276)	European	977,323 (47309 cases)	log-transformed odds	Age, sex, principal components
Coronary artery disease (273)	> 95% European	1,165,690 (181,522 cases)	log-transformed odds	Age, sex, study-specific covariates, principal components
Myocardial infarction (274)	77% European	171,875 (43676 cases)	log-transformed odds	Age, sex, study-specific covariates, principal components
Ischemic stroke (277)	European	1,296,908 (62100 cases)	log-transformed odds	Age, sex, study-specific covariates, principal components
Type 2 diabetes (279)	European	933,970 (80154 cases)	log-transformed odds	Age, sex, study-specific covariates, principal components
Chronic kidney disease (303)	European	480,698 (41395 cases)	log-transformed odds	Age, sex, study-specific covariates, principal components
eGFRcrea (278)	European	1,004,040	log(ml/min/1.73 m^2)	Age, sex, study-specific covariates, principal components
eGFRcys (278)	European	460,826	log(ml/min/1.73 m^2)	Age, sex, study-specific covariates, principal components
HbA1c (304)	European	146,806	%	Study-specific covariates, principal components
Fasting glucose (304)	European	200,622	mmol l^{-1}	BMI, study-specific covariates, principal components

3.2.5 Statistical Analysis

3.2.5.1 Two-sample MR

MR is a methodological approach that assesses the potential causal effect of an exposure on an outcome by utilising genetic variants as instrumental variables. Given that genetic variants are allocated randomly at conception, MR can circumvent biases arising from

confounding variables and reverse causality (157). MR relies on three main instrumental variable assumptions: the genetic variant is strongly associated with the exposure; the genetic variant is not associated with any confounders in the exposure-outcome association; the genetic variant affects the outcomes only via the exposure (306).

Variant-exposure association was harmonised with the variant-outcome association by aligning the effect allele. Palindromic variants were not omitted from the study. Proxies were not employed to maintain uniformity in the genetic variants utilised across various MR analyses. Primary and Sensitivity Analyses: The inverse-variance weighted (IVW) method served as the primary analysis technique, complemented by weighted median, simple median, and MR-Egger methods as sensitivity analyses. These were chosen for their capacity to test the robustness of my findings against possible pleiotropy and other biases. Additional negative control outcome analyses were conducted to further assess potential confounding influences. IVW MR was implemented as the primary analysis to investigate the causal effects of genetically proxied antihypertensive drug classes on the outcomes (179). MR estimates were presented as odds ratio (OR) for each binary outcome or beta for each continuous outcome, with a 95% confidence interval (95% CI) per 10 mmHg decrease in SBP, orientating to the BP-lowering effect of antihypertensive drug classes. The simple median, weighted median and MR Egger were performed as the sensitivity analysis to assess the robustness of the main result to the potential pleiotropic effects of genetic instruments (192, 195). In the weighted median method, the MR estimates from variants are ordered by their magnitude weighted for their precision. The overall MR estimate is the median of the variant-specific estimates, with the standard error derived by bootstrapping. The weighted median method can give a robust MR estimate when valid genetic variants account for at least 50% of the total variants (192). The manifestation of horizontal pleiotropy, which violates the exclusion restriction assumption of MR (186), can be detected using the MR Egger intercept test (195). Moreover, the MR Egger method can provide a consistent MR estimate of the causal effect under a less stringent assumption than the instrumental variable assumptions (the InSIDE assumption). It assumes that any direct effects of the genetic variants on the outcome are uncorrelated with the effects of these variants on the exposure (195). Furthermore, negative control outcome MR analysis (307) was conducted on the following outcomes: low hand grip strength (308) (48,596 European ancestry cases, 207,927 European ancestry controls), myopia (309) (36,623 European ancestry cases, 419,031 European ancestry controls), Parkinson disease (310) (15,056 European ancestry cases, 18,618 European ancestry proxy cases, 449,056 European ancestry controls), and heel bone mineral density (311) (583,314 European ancestry individuals). These outcomes are not known to share any pathological pathways with

hypertension. The negative control outcome analysis aims to detect potential confounding of the genetic variants and the exposure or outcome, which cannot be identified through the standard pleiotropy robust methods (307). As collider bias may be introduced in the MR analysis using summary statistics from the GWAS meta-analysis study by Evangelou et al. (115) due to the adjustment for BMI in the regression analysis, I performed sensitivity analysis with BP genetic estimates obtained from a GWAS conducted on UK Biobank data (312). The summary data were extracted from MR-Base (MR-Base ID: ukb-a-360) (205).

For the observed significant associations of antihypertensive drug classes with the outcomes, the propagation of error method was performed to examine whether the MR estimates for the drug target perturbation significantly differ from those obtained for genetically proxied SBP-lowering effect more generally (section 2.1.7).

IVW, simple median, weighted median, and MR Egger analyses were conducted using the MendelianRandomization package (version 0.6.0) in R (281). LD clumping was performed using PLINK (205).

3.3 Results

3.3.1 Mendelian randomization analysis of genetically proxied systolic blood pressure

Across the genome, 454 uncorrelated SNPs ($r^2 < 0.001$) were identified as genetic instruments for SBP. All of the SNPs had F-statistics greater than 10, indicating a low risk of weak instrument bias. The total variance in SBP explained by the genetic instruments is 4.7%.

Genetically predicted lower SBP was significantly associated with all the main outcomes of interest. Every 10 mmHg reduction in genetically proxied SBP was associated with a decreased risk of all the outcomes (**Table 3.3**), with ORs for the outcomes ranging from 0.71 to 0.87. The MR estimates obtained from the sensitivity analysis (simple median, weighted median, and MR Egger) were overall consistent with the primary analysis for all the outcomes except for T2D, which had evidence of horizontal pleiotropy from the MR Egger intercept test (MR Egger OR per 10 mmHg decrease in SBP, 1.08, 95% CI, 0.92 - 1.26, p = 0.35) (**Table 3.3**). MR results from the analysis on the outcomes eGFRcrea, eGFRcys, and HbA1c were not significant (**Table 3.4**). Genetically proxied SBP was nominally associated with a decrease in fasting glucose (Beta estimate per 10 mmHg decrease in SBP, -0.017, 95% CI, (-0.027) - (-0.007), p = 0.0009) (**Table 3.4**).

Table 3.3 MR estimates of the effect of genetically predicted SBP on main outcomes.

SBP, systolic blood pressure; CAD indicates coronary artery disease; MI, myocardial infarction; AF, atrial fibrillation; HF, heart failure; CKD, chronic kidney disease; T2D, type 2 diabetes; IVW, inverse-variant weighted method; nSNPs, number of single nucleotide polymorphisms; OR, odds ratio per 10 mmHg decrease in SBP; OR_lci95, lower bound of 95% confidence interval of odds ratio ; OR_uci95, upper bound of 95% confidence interval of odds ratio; pval, p-value; NA, not applicable.

Method	Exposure	Outcome	nSNPs	OR	OR_lci95	OR_uci95	p
Simple median	SBP	AF	454	0.839	0.805	0.873	1.57E-17
Weighted median	SBP	AF	454	0.833	0.8	0.868	2.15E-18
IVW	SBP	AF	454	0.851	0.817	0.887	9.49E-15
MR-Egger	SBP	AF	454	0.851	0.767	0.944	0.002
(intercept)	SBP	AF	454	NA	NA	NA	0.994
Simple median	SBP	CAD	454	0.714	0.691	0.739	4.85E-87
Weighted median	SBP	CAD	454	0.704	0.681	0.729	3.46E-90
IVW	SBP	CAD	454	0.711	0.678	0.745	8.14E-47
MR-Egger	SBP	CAD	454	0.715	0.635	0.806	3.44E-08
(intercept)	SBP	CAD	454	NA	NA	NA	0.912
Simple median	SBP	MI	454	0.738	0.693	0.786	1.89E-21
Weighted median	SBP	MI	454	0.725	0.68	0.773	5.94E-23
IVW	SBP	MI	454	0.745	0.704	0.789	5.00E-24
MR-Egger	SBP	MI	454	0.732	0.633	0.846	2.51E-05
(intercept)	SBP	MI	454	NA	NA	NA	0.793
Simple median	SBP	HF	454	0.79	0.754	0.828	2.96E-23
Weighted median	SBP	HF	454	0.789	0.752	0.828	9.58E-22
IVW	SBP	HF	454	0.808	0.777	0.839	1.53E-27
MR-Egger	SBP	HF	454	0.801	0.726	0.884	1.01E-05
(intercept)	SBP	HF	454	NA	NA	NA	0.865
Simple median	SBP	Ischemic stroke	454	0.764	0.73	0.798	1.12E-32
Weighted median	SBP	Ischemic stroke	454	0.748	0.716	0.782	4.57E-38

IVW	SBP	Ischemic stroke	454	0.754	0.726	0.782	2.18E-50
MR-Egger	SBP	Ischemic stroke	454	0.69	0.628	0.758	1.04E-14
(intercept)	SBP	Ischemic stroke	454	NA	NA	NA	0.044
Simple median	SBP	CKD	454	0.871	0.825	0.92	7.54E-07
Weighted median	SBP	CKD	454	0.86	0.814	0.909	7.83E-08
IVW	SBP	CKD	454	0.864	0.817	0.914	3.76E-07
MR-Egger	SBP	CKD	454	0.836	0.724	0.966	0.015
(intercept)	SBP	CKD	454	NA	NA	NA	0.626
Simple median	SBP	T2D	454	0.858	0.823	0.895	8.11E-13
Weighted median	SBP	T2D	454	0.896	0.858	0.935	4.92E-07
IVW	SBP	T2D	454	0.83	0.779	0.883	4.56E-09
MR-Egger	SBP	T2D	454	1.078	0.921	1.262	0.35
(intercept)	SBP	T2D	454	NA	NA	NA	0.0004

Table 3.4 MR estimates of the effect of genetically lower SBP on eGFR, HbA1c, and fasting glucose outcomes.

SBP, systolic blood pressure; eGFRcrea, creatinine-estimated glomerular filtration rate; eGFRcys, eGFR based on cystatin C; FG, fasting glucose; IVW, inverse-variant weighted method; nSNPs, number of single nucleotide polymorphisms; b, effect size per 10 mmHg decrease in SBP; lo_ci, lower bound of 95% confidence interval of b ; up_ci, upper bound of 95% confidence interval of b; pval, p-value; NA, not applicable.

Method	**Exposure**	**Outcome**	**nSNPs**	**b**	**lo_ci**	**up_ci**	**p**
Simple median	SBP	eGFRcrea	454	0.003	0.001	0.005	0.003
Weighted median	SBP	eGFRcrea	454	0.003	0.001	0.004	0.011
IVW	SBP	eGFRcrea	454	0.001	-0.002	0.005	0.513
MR-Egger	SBP	eGFRcrea	454	0.005	-0.004	0.015	0.253
(intercept)	SBP	eGFRcrea	454	NA	NA	NA	0.336
Simple median	SBP	eGFRcys	454	0.000	-0.003	0.003	1
Weighted median	SBP	eGFRcys	454	0.000	-0.003	0.003	1

IVW	SBP	eGFRcys	454	0.003	-0.002	0.007	0.306
MR-Egger	SBP	eGFRcys	454	0.006	-0.006	0.018	0.339
(intercept)	SBP	eGFRcys	454	NA	NA	NA	0.547
Simple median	SBP	HbA1c	454	-0.001	-0.009	0.006	0.718
Weighted median	SBP	HbA1c	454	-0.000	-0.008	0.008	1
IVW	SBP	HbA1c	454	-0.001	-0.009	0.006	0.724
MR-Egger	SBP	HbA1c	454	0.014	-0.005	0.033	0.141
(intercept)	SBP	HbA1c	454	NA	NA	NA	0.079
Simple median	SBP	FG	454	-0.015	-0.025	-0.004	0.006
Weighted median	SBP	FG	454	-0.009	-0.02	0.002	0.102
IVW	SBP	FG	454	-0.017	-0.027	-0.007	0.001
MR-Egger	SBP	FG	454	0.007	-0.018	0.032	0.58
(intercept)	SBP	FG	454	NA	NA	NA	0.041

3.3.2 MR analysis of genetically proxied antihypertensive drug classes

By using predefined selection criteria, I found genetic instruments for 10 drug classes. There were 2 SNPs for AGTi, 1 SNP for ACEi, 29 SNPs for CCB, 7 SNPs for beta-blockers, 1 SNP for alpha-blockers, 4 SNPs for loop diuretics, and 1 SNP for other PSDs (PSDs that are not aldosterone antagonists). For the vasodilator antihypertensive drug classes, there was 1 SNP for ERAs, 5 SNPs for PDE5 inhibitors, and 6 SNPs for sGC stimulators (**Table 3.5)**. I could not identify any instruments for the remaining drug classes, including ARBs, TZDs and aldosterone antagonists. The F-statistics for the SNPs ranged from 30.48 to 352.48, indicating a low risk of weak instrument bias.

Table 3.5 Genetic proxies identified for each antihypertensive drug class.

AGTi, angiotensinogen inhibitor; ACEi, angiotensin-converting enzyme inhibitor; CCB, calcium channel blocker; β-blocker, beta-adrenoceptor blocker; α-blocker, alpha-adrenoceptor blocker; PSD, potassium-sparing diuretic; ERA, endothelin receptor antagonist; PDE5 inhibitor, phosphodiesterase-5 inhibitor; sGC stimulator, soluble guanylate cyclase stimulator; sbp_ea, effect allele in the GWAS for SBP; sbp_nea, non-effect allele; sbp_eaf; effect allele frequency; sbp_se; standard error; sbp_p, p-value; sbp_n, sample size.

Drug class	Target gene	SNP	chr	chr_pos	sbp_ea	sbp_nea	sbp_eaf	sbp_b	sbp_se	sbp_p	sbp_n	R^2	F_stat
ACEi	*ACE*	rs8077276	17	61547562	A	G	0.62	-0.295	0.031	5.15E-21	743708	0.0001	88.204
AGTi	*AGT*	rs61751101	1	230836734	A	G	0.93	-0.404	0.061	2.73E-11	738168	0.0001	44.356
	AGT	rs699	1	230845794	A	G	0.593	-0.375	0.031	5.59E-34	721189	0.0002	148.08
α-blocker	*ADRA1D*	rs1741344	20	4101800	T	C	0.636	0.204	0.031	6.94E-11	744815	0.0001	42.604
β-blocker	*ADRB1*	rs11196549	10	115707298	A	G	0.043	0.688	0.078	1.58E-18	738169	0.0001	77.099
	ADRB1	rs460718	10	115721364	A	G	0.327	-0.276	0.032	1.36E-17	738169	0.0001	72.775
	ADRB1	rs11196597	10	115788094	A	G	0.133	0.286	0.046	4.23E-10	737164	0.0001	38.94
	ADRB1	rs17875473	10	115800294	T	C	0.087	0.328	0.055	2.66E-09	738170	0.0000	35.372
	ADRB1	rs1801253	10	115805056	C	G	0.734	0.463	0.034	2.84E-41	738169	0.0002	180.84
	ADRB1	rs68122733	10	115831533	A	G	0.828	0.334	0.041	1.69E-16	738167	0.0001	67.798
	ADRB1	rs143854972	10	115843445	A	G	0.058	0.431	0.066	8.89E-11	738168	0.0001	42.054
CCB	*CACNA1C*	rs2239046	12	2434419	A	G	0.682	0.208	0.032	9.58E-11	745818	0.0001	41.807

CACNA1C	rs714277	12	2514270	T	C	0.283	0.199	0.033	2.38E-09	745820	0.0000	35.569
CACNB2	rs1891392	10	18336421	T	C	0.287	0.228	0.033	9.02E-12	738170	0.0001	46.436
CACNB2	rs2488152	10	18355456	A	G	0.896	-0.377	0.05	2.54E-14	738167	0.0001	57.975
CACNB2	rs75699707	10	18359294	A	G	0.026	-0.576	0.103	2.10E-08	730149	0.0000	31.406
CACNB2	rs1888693	10	18440444	A	G	0.345	0.386	0.032	4.69E-34	736050	0.0002	148.117
CACNB2	rs16916914	10	18457722	T	C	0.963	-0.564	0.081	2.72E-12	737424	0.0001	48.896
CACNB2	rs7076319	10	18459450	A	G	0.734	-0.321	0.034	5.07E-21	737054	0.0001	88.614
CACNB2	rs61278674	10	18481737	A	G	0.906	-0.33	0.054	1.03E-09	737163	0.0001	37.3
CACNB2	rs1779209	10	18514561	T	C	0.288	0.274	0.034	4.23E-16	729448	0.0001	66.306
CACNB2	rs10828399	10	18553968	A	G	0.522	-0.195	0.03	1.10E-10	738168	0.0001	41.564
CACNB2	rs10828452	10	18592450	A	T	0.793	0.305	0.039	4.20E-15	737164	0.0001	61.63
CACNB2	rs10828542	10	18627285	A	G	0.614	0.182	0.031	5.18E-09	738170	0.0000	34.134
CACNB2	rs12780039	10	18678987	C	G	0.121	0.285	0.047	1.26E-09	738167	0.0000	36.822
CACNB2	rs112133583	10	18695681	T	C	0.03	-0.555	0.097	1.18E-08	737169	0.0000	32.489
CACNB2	rs11014170	10	18710991	A	G	0.021	-0.67	0.115	5.61E-09	732148	0.0000	33.953
CACNB2	rs7923191	10	18727901	A	G	0.792	-0.369	0.038	1.10E-22	737054	0.0001	96.311
CACNB2	rs12258967	10	18727959	C	G	0.705	0.633	0.034	1.08E-78	737165	0.0005	352.48
CACNB2	rs72786098	10	18729855	A	G	0.032	-0.503	0.088	1.18E-08	737055	0.0000	32.489
CACNB2	rs1998822	10	18755664	A	G	0.723	-0.196	0.034	1.15E-08	727331	0.0000	32.586
CACNB2	rs4748474	10	18790727	A	G	0.521	0.195	0.03	1.61E-10	729908	0.0001	40.977
CACNA1D	rs114718455	3	53464055	A	G	0.966	-0.51	0.091	1.72E-08	730732	0.0000	31.782
CACNA1D	rs3821843	3	53558012	A	G	0.681	0.337	0.034	6.56E-24	736049	0.0001	101.378
CACNA1D	rs114987861	3	53605712	A	G	0.028	0.529	0.096	3.36E-08	737054	0.0000	30.48
CACNA1D	rs113210396	3	53612327	T	G	0.045	-0.434	0.077	1.76E-08	737164	0.0000	31.739
CACNA1D	rs7340705	3	53734443	T	C	0.673	-0.243	0.032	4.87E-14	738169	0.0001	56.717

	CACNA1D	rs62250937	3	53870318	T	C	0.911	0.324	0.056	7.28E-09	736050	0.0000	33.433
	CACNA1H	rs11641374	16	1347717	A	C	0.6	-0.194	0.031	3.26E-10	741588	0.0001	39.539
	CACNB3	rs150857355	12	49209340	C	G	0.022	0.941	0.112	5.20E-17	731300	0.0001	70.279
ERA	*EDNRA*	rs13143677	4	148419040	A	G	0.293	0.224	0.033	1.82E-11	738169	0.0001	45.128
Loop diuretic	*SLC12A1*	rs36029774	15	48693034	T	C	0.104	-0.376	0.051	1.23E-13	745818	0.0001	54.871
	SLC12A2	rs6860245	5	127367998	C	G	0.244	-0.211	0.035	2.10E-09	738169	0.0000	35.796
	SLC12A2	rs17676242	5	127589648	A	G	0.092	0.416	0.053	6.62E-15	736050	0.0001	60.799
	SLC12A2	rs17676396	5	127616544	T	C	0.705	0.21	0.033	2.22E-10	737054	0.0001	40.162
Other PSDs	*SCNN1D*	rs307359	1	1280014	A	G	0.069	-0.379	0.065	5.18E-09	672751	0.0001	34.136
PDE5 inhibitor	*PDE5A*	rs4834792	4	120555696	A	T	0.48	0.197	0.03	7.24E-11	735152	0.0001	42.4
	NOS3	rs740956	7	150598440	T	C	0.567	0.2	0.031	5.21E-11	736110	0.0001	43.042
	NOS3	rs3918226	7	150690176	T	C	0.081	0.664	0.058	8.46E-31	731379	0.0002	133.352
	NOS3	rs891511	7	150704843	A	G	0.334	-0.351	0.033	6.13E-26	735050	0.0002	110.85
	NOS3	rs34581141	7	150810073	A	G	0.684	-0.188	0.033	9.08E-09	738169	0.0000	33.018
sGC stimulator	*GUCY1A3*	rs1483040	4	156503399	C	G	0.094	0.346	0.053	5.22E-11	738168	0.0001	43.055
	GUCY1A3	rs990619	4	156507678	C	G	0.476	-0.297	0.03	4.68E-23	738170	0.0001	97.622
	GUCY1A3	rs10010626	4	156519815	A	T	0.072	0.369	0.063	4.12E-09	719168	0.0000	34.541
	GUCY1A3	rs28531888	4	156613960	T	C	0.199	-0.226	0.04	1.17E-08	737164	0.0000	32.455
	GUCY1A3	rs12643599	4	156639846	A	G	0.64	0.313	0.031	1.23E-23	738170	0.0001	100.256
	GUCY1A3	rs78975989	4	156718205	A	G	0.109	-0.32	0.05	1.67E-10	738167	0.0001	40.771

3.3.2.1 AGT inhibitors

Genetically proxied AGT inhibition represented by SBP reduction was associated with a decreased risk of CAD (OR per 10 mmHg decrease in SBP, 0.65, 95% CI, 0.52 - 0.81, p = 0.0002), and ischemic stroke (OR, 0.49, 95% CI, 0.33 - 0.71, p = 0.0002), and nominally associated with a decreased risk of MI (OR, 0.47, 95% CI, 0.29 - 0.76, p = 0.002) but not with other outcomes (**Fig. 3.2 - 3.3**, **Table 3.6**).

3.3.2.2 ACE inhibitors

Genetically proxied ACE inhibition represented by SBP reduction showed suggestive evidence ($p < 0.05$) for a protective effect on CAD risk (OR per 10 mmHg decrease in SBP, 0.65, 95% CI, 0.46 - 0.91, p = 0.001) and T2D risk (OR, 0.51, 95% CI, 0.33 - 0.80, p = 0.003). However, the p-values did not reach the Bonferroni-corrected threshold (**Fig. 3.2 - 3.3, Table 3.6**). Genetically proxied ACE inhibition was associated with decreased HbA1c level (effect size, per 10 mmHg decrease in SBP, -0.102, 95% CI, -0.195 - (-0.009), p = 0.03), and decreased fasting glucose level (-0.21, 95% CI, (-0.337 - (-0.084), p = 0.001 (**Fig. 3.4**).

3.3.2.3 Calcium channel blockers

For CCBs, MR analysis found significant causal associations with reduced AF risk (OR, 0.64, 95% CI, 0.57 - 0.70, $p = 5.35 \times 10^{-18}$), CAD risk (OR, 0.75, 95% CI, 0.66 - 0.84, $p = 1.1 \times 10^{-6}$), ischemic stroke risk (OR, 0.82, 95% CI, 0.74 - 0.92, p = 0.0006), and HF risk (OR, 0.73, 95% CI, 0.65 - 0.83, $p = 3.8 \times 10^{-7}$). Suggestive evidence was also found for the causal association of genetically predicted CCBs with reduced MI risk (OR, 0.78, 95% CI, 0.66 - 0.92, p = 0.003). The IVW estimates were consistent with the MR estimates from the weighted median, the simple median and the MR Egger methods. No evidence of directional pleiotropy was found in the MR Egger intercept tests (intercept p-value > 0.05) (**Fig. 3.2 - 3.3, Table 3.6**).

3.3.2.4 Beta-blockers

Using genetic instruments for beta-blockers, MR results showed protective effects on AF (OR, 0.65, 95% CI, 0.54 - 0.79, $p = 1.37 \times 10^{-5}$), CAD (OR, 0.67, 95% CI, 0.58 - 0.77, $p = 2.39 \times 10^{-8}$), and MI (0.57, 95% CI, 0.42 - 0.76, p = 0.0002), and suggestive protective effects on HF (0.70, 95% CI, 0.56 - 0.88, p = 0.002), but not with other outcomes. MR results from the sensitivity analysis were consistent in magnitude and direction with the IVW results. No evidence of directional pleiotropy was detected (**Fig. 3.2 - 3.3, Table 3.6**).

Figure 3.2 MR estimates for the effect of genetically lower blood pressure through the drug target genes of five drug classes on the risk of the outcomes.
Data were represented as odds ratio (OR) with 95% confidence interval (CI) of the outcome per 10 mmHg decrease in systolic blood pressure with risk for the outcomes of interest. OR of less than 1 suggests a decreased risk of disease associated with blood pressure lowering treatments, and vice versa. CAD indicates coronary artery disease; MI, myocardial infarction; AF, atrial fibrillation; HF, heart failure; CKD, chronic kidney disease; T2D, type 2 diabetes; AGTi, Angiotensinogen inhibitor; ACEi, angiotensin-converting enzyme inhibitors; CCBs, calcium channel blockers; SBP, systolic blood pressure.

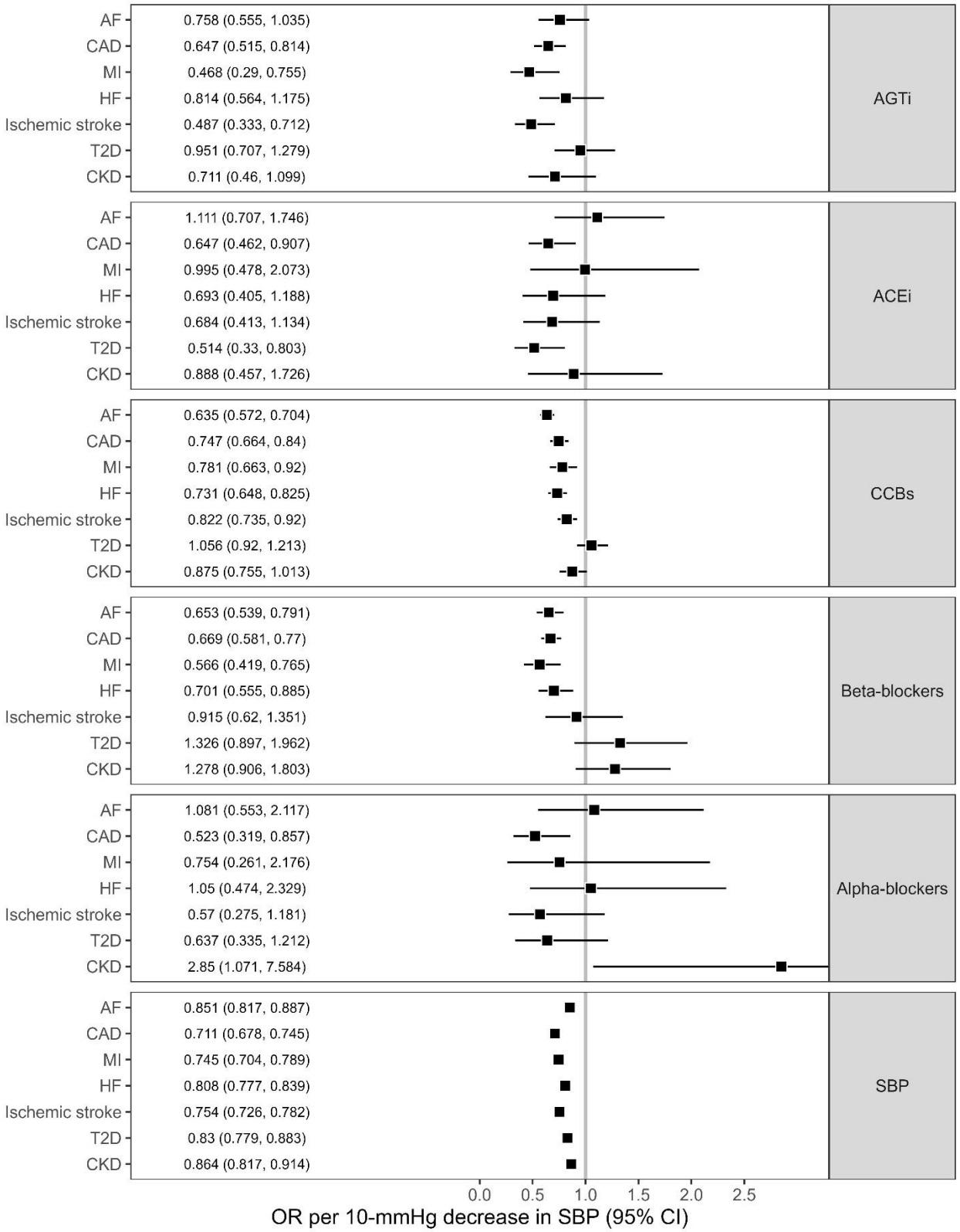


Figure 3.3 Drug target Mendelian randomization estimates estimating the association of genetically lower blood pressure through drug target perturbation with the risk of the outcomes.

MR estimate was presented as an odds ratio for each outcome per 10 mmHg decrease in systolic blood pressure. The rows depict the seven outcomes; the columns represent the genetically predicted antihypertensive drug classes and systolic blood pressure. The heatmap cells were coloured based on the direction of the effect times -$\log_{10}$ (p-value). A negative direction indicates the effect of decreasing outcome risk and vice versa. Odds ratios were provided for the significant association considering the conventional p-value threshold. The asterisk indicates the significant association after adjusting for multiple testing. CAD indicates coronary artery disease; MI, myocardial infarction; AF, atrial fibrillation; HF, heart failure; CKD, chronic kidney disease; T2D, type 2 diabetes; AGTi, angiotensinogen inhibitor; ACEi, angiotensin-converting enzyme inhibitors; CCBs, calcium channel blockers; Other PSDs, potassium-sparing diuretics (not include aldosterone antagonists); ERAs, endothelin receptor antagonists; PDE5 inhibitors, phosphodiesterase type 5; sGC; soluble guanylate cyclase; SBP, systolic blood pressure.

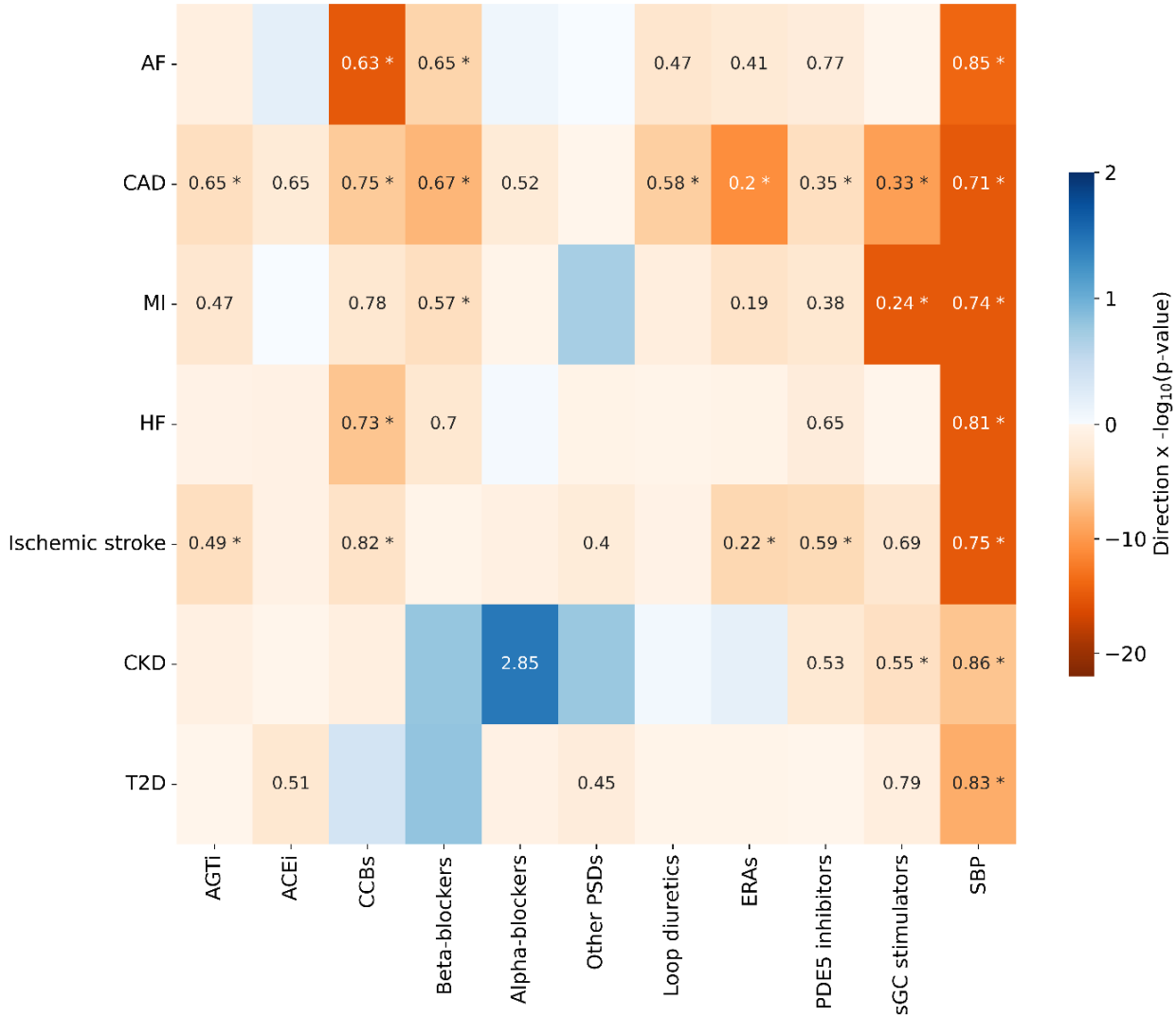
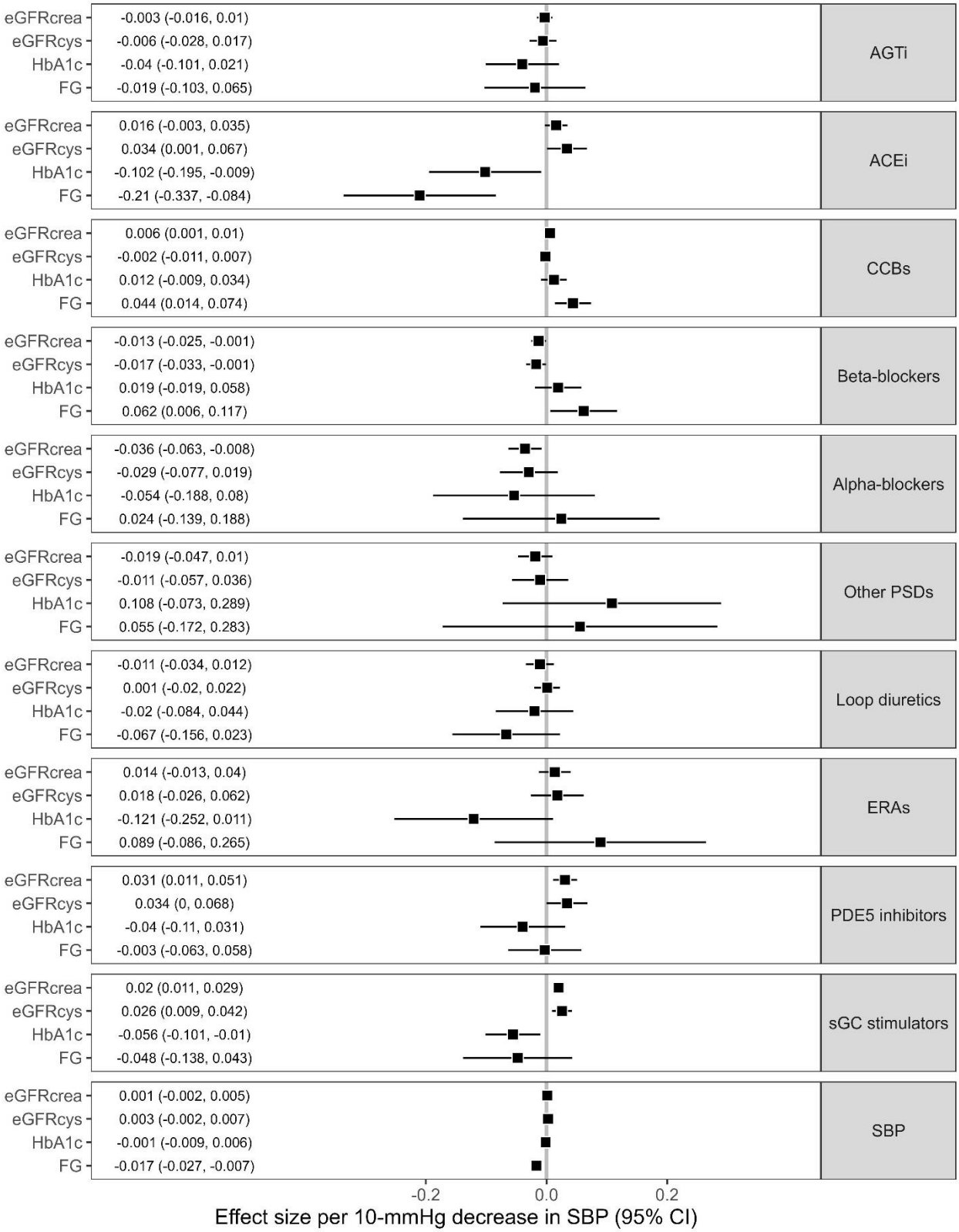


Figure 3.4 MR estimates of the effect of genetically lower SBP generally and genetically lower SBP through drug target perturbation on outcomes in the sensitivity analysis.
Data were represented as effect size with 95% confidence interval (CI) of the outcome per 10 mmHg decrease in systolic blood pressure with risk for the outcomes of interest. FG, fasting glucose; AGTi, Angiotensinogen inhibitor; ACEi, angiotensin-converting enzyme inhibitors; CCBs, calcium channel blockers; Other PSDs, potassium-sparing diuretics (not include aldosterone antagonists); ERAs, endothelin receptor antagonists; PDE5 inhibitors, phosphodiesterase type 5; sGC; soluble guanylate cyclase; SBP, systolic blood pressure.



3.3.2.5 Alpha-blockers

For alpha-blockers, MR analysis did not find any causal signal that crossed the Bonferroni threshold. However, I found nominally significant association of genetically proxied alpha-blockers with reduced CAD risk (OR, 0.52, 95% CI, 0.32 - 0.86, p = 0.01), and increased CKD risk (OR, 2.85, 95% CI, 1.07 - 7.58, p = 0.04). ((**Fig. 3.2 - 3.3, Table 3.6)**. Sensitivity analysis found significant association between genetically proxied alpha-blockers and decreased eGFRcrea (Effect size, -0.04, 95% CI, -0.06 - (-0.01), p = 0.01) (**Fig. 3.4).**

3.3.2.6 Loop Diuretics

For loop diuretics, MR results showed a protective effect on CAD risk (OR, 0.58, 95% CI, 0.46 - 0.73, $p = 2.4 \times 10^{-6}$) and a suggestive protective effect on AF risk (OR, 0.47, 95% CI, 0.29 - 0.75, p = 0.002). MR results from the sensitivity analysis were consistent with the IVW results. No evidence of directional pleiotropy was found in the MR Egger intercept tests (intercept p-value > 0.05) (**Fig. 3.3, Fig. 3.5, Table 3.6**). Using the genetic instrument for potassium-sparing diuretics (not including aldosterone antagonists), I found suggestive evidence for a causal association of PSDs with reduced ischemic stroke risk (OR 0.4, 95% CI, 0.18 - 0.86, p = 0.02) and decreased T2D risk (OR 0.45, 95% CI, 0.23 - 0.91, p = 0.03) (**Fig. 3.3, Fig. 3.5, Table 3.6)**. However, sensitivity analysis on HbA1c and fasting glucose was not significant (**Fig. 3.4).**

3.3.2.7 Endothelin receptor antagonists

Genetically proxied endothelin receptor antagonism was significantly associated with decreased CAD risk (OR 0.2, 95% CI, 0.12 - 0.32, $p = 1.19 \times 10^{-11}$), and decreased ischemic stroke risk (OR 0.22, 95% CI, 0.11 - 0.44, $p = 2.56 \times 10^{-5}$); and was and nominally associated with decreased MI risk (OR, 0.19, 95% CI, 0.07 - 0.50, p = 0.0007), and decreased AF risk (OR, 0.41, 95% CI, 0.22 - 0.78, p = 0.006) (**Fig. 3.3, Fig. 3.5, Table 3.6)**.

3.3.2.8 PDE5 inhibitors

Genetically predicted PDE5 inhibition was significantly associated with a lower risk of CAD (OR 0.35, 95% CI, 0.20 - 0.61, p = 0.0002) and a decreased risk of ischemic stroke (OR, 0.59, 0.45 - 0.76, $p = 6.55 \times 10^{-5}$), and nominally associated with a reduced risk of MI (OR 0.38, 95% CI, 0.20 - 0.73, p = 0.003), HF (OR 0.65, 95% CI, 0.45 - 0.94, p = 0.02), AF (OR, 0.77, 95% CI, 0.61 - 0.96, p = 0.02), and CKD (OR, 0.53, 95% CI, 0.35 - 0.82, p = 0.004). MR results from the sensitivity analysis were consistent with the IVW results (**Fig. 3.3, Fig. 3.5, Table 3.6)**. Sensitivity MR analyses showed significant association of genetically proxied PDE5 inhibition with increased eGFRcrea (effect size, per 10 mmHg decrease in

SBP, 0.03, 95% CI, 0.01 - 0.05, p = 0.003) and increased eGFRcys (effect size, 0.03, 95% CI, 0 - 0.07, p = 0.048) (**Fig. 3.4)**.

3.3.2.9 sGC stimulators

Genetically proxied sGC stimulation was significantly associated with decreased CAD risk (OR, 0.33, 95% CI, 0.24 - 0.47, $p = 3.45 \times 10^{-10}$), decreased MI risk (OR, 0.24, 95% CI, 0.17 - 0.34, $p = 5.68 \times 10^{-16}$), and decreased CKD risk (OR, 0.55, 95% CI, 0.40 - 0.76, p = 0.003), and nominally associated with a decreased risk of ischemic stroke (OR, 0.69, 0.54 - 0.87, p = 0.002) and T2D (OR, 0.80, 95% CI, 0.63 - 1, p = 0.05). The sensitivity analysis showed consistent results with the IVW results, with no directional pleiotropy found in the MR Egger intercept tests (**Fig. 3.3, Fig. 3.5, Table 3.6).** Sensitivity analysis found significant association with increased eGFRcrea (Effect size, 0.02, 95% CI, 0.01 - 0.03, $p = 2.79 \times 10^{-5}$), and increased eGFRcys (Effect size, 0.03, 95% CI, 0.01 - 0.04, p = 0.002) (**Fig. 3.4).**

Negative outcome analyses on low hand grip strength, myopia, Parkinson disease, and heel bone mineral density did not show significant associations (**Table 3.7**).

3.3.2.10 Replication analyses

In the MR analyses that used genetic associations with SBP from another BP GWAS conducted on UK Biobank data, the associations between antihypertensive drug classes and the main outcomes were consistent with those from the main MR analyses (**Table 3.8**).

3.3.2.11 Effect difference

Results from the propagation of error method showed a significant difference between MR estimates for SBP lowering effect more generally and those obtained for the following drug targets: genetically proxied AGT inhibition (on ischemic stroke), CCBs (on AF), beta-blockers (on AF), loop diuretics (on AF), ERAs (on CAD and ischemic stroke), PDE5 inhibitors (on CAD), and sGC stimulators (on CAD, MI, and CKD) (**Table 3.9**).

Figure 3.5 MR estimates for the effect of genetically lower blood pressure through the drug target genes of five drug classes on the risk of the outcomes.
Data were represented as odds ratio (OR) with 95% confidence interval (CI) of the outcome per 10-mmHg decrease in systolic blood pressure with risk for the outcomes of interest. OR of less than 1 suggests a decreased risk of disease associated with blood pressure lowering treatments, and vice versa.

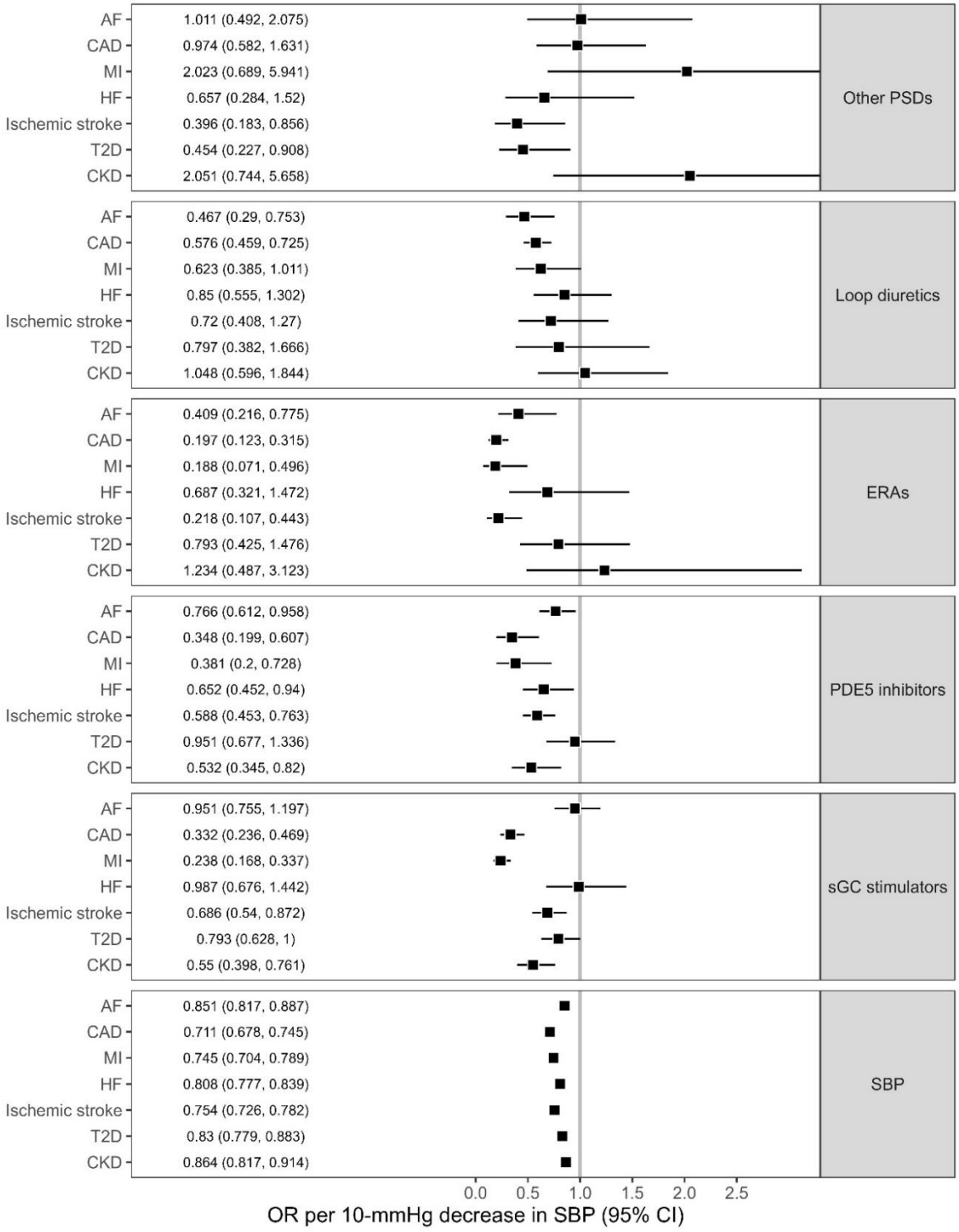


Table 3.6 MR analysis results for genetically proxied drug classes on the main outcomes.

SBP, systolic blood pressure; CAD indicates coronary artery disease; MI, myocardial infarction; AF, atrial fibrillation; HF, heart failure; CKD, chronic kidney disease; T2D, type 2 diabetes; IVW, inverse-variant weighted method; nSNP, number of single nucleotide polymorphisms; OR, odds ratio per 10 mmHg decrease in SBP; OR_lci, lower bound of 95% confidence interval of odds ratio ; OR_uci, upper bound of 95% confidence interval of odds ratio; pval, p-value; NA, not applicable; AGTi, Angiotensinogen inhibitor; ACEi, angiotensin-converting enzyme inhibitors; CCBs, calcium channel blockers; Other PSDs, potassium-sparing diuretics (not include aldosterone antagonists); ERAs, endothelin receptor antagonists; PDE5 inhibitors, phosphodiesterase type 5; sGC; soluble guanylate cyclase

Drug class	Exposure	Outcome	Method	n_SNP	OR	OR_lci	OR_uci	p
AGTi	SBP	AF	IVW	2	0.758	0.555	1.035	0.081
		CAD		2	0.647	0.515	0.814	0.0002
		MI		2	0.468	0.29	0.755	0.002
		HF		2	0.814	0.564	1.175	0.271
		Ischemic stroke		2	0.487	0.333	0.712	0.0002
		CKD		2	0.711	0.46	1.099	0.125
		T2D		2	0.951	0.707	1.279	0.74
ACEi	SBP	AF	IVW	1	1.111	0.707	1.746	0.648
		CAD		1	0.647	0.462	0.907	0.012
		MI		1	0.995	0.478	2.073	0.99
		HF		1	0.693	0.405	1.188	0.182
		Ischemic stroke		1	0.684	0.413	1.134	0.141
		CKD		1	0.888	0.457	1.726	0.726
		T2D		1	0.514	0.33	0.803	0.003
CCBs	SBP	AF	Simple median	29	0.609	0.52	0.714	8.79E-10
			Weighted median	29	0.616	0.525	0.723	3.17E-09
			IVW	29	0.635	0.572	0.704	5.35E-18
			MR-Egger	29	0.69	0.531	0.896	0.005
			(intercept)	29	NA	NA	NA	0.498
CCBs	SBP	CAD	Simple median	29	0.725	0.642	0.82	2.68E-07
			Weighted median	29	0.725	0.648	0.812	2.34E-08
			IVW	29	0.747	0.664	0.84	1.10E-06
			MR-Egger	29	0.634	0.472	0.852	0.002
			(intercept)	29	NA	NA	NA	0.236
CCBs	SBP	MI	Simple median	29	0.781	0.609	1.001	0.051
			Weighted median	29	0.731	0.576	0.928	0.01
			IVW	29	0.781	0.663	0.92	0.003
			MR-Egger	29	0.707	0.461	1.083	0.111
			(intercept)	29	NA	NA	NA	0.62
CCBs	SBP	HF	Simple median	29	0.766	0.635	0.923	0.005

			Weighted median	29	0.706	0.593	0.842	0.0001
			IVW	29	0.731	0.648	0.825	3.80E-07
			MR-Egger	29	0.681	0.502	0.924	0.014
			(intercept)	29	NA	NA	NA	0.617
CCBs	SBP	Ischemic stroke	Simple median	29	0.841	0.707	0.999	0.049
			Weighted median	29	0.793	0.674	0.932	0.005
			IVW	29	0.822	0.735	0.92	0.001
			MR-Egger	29	0.87	0.654	1.156	0.337
			(intercept)	29	NA	NA	NA	0.675
CCBs	SBP	CKD	Simple median	29	0.855	0.683	1.071	0.173
			Weighted median	29	0.837	0.674	1.04	0.108
			IVW	29	0.875	0.755	1.013	0.073
			MR-Egger	29	0.682	0.47	0.989	0.044
			(intercept)	29	NA	NA	NA	0.153
CCBs	SBP	T2D	Simple median	29	1.108	0.945	1.3	0.206
			Weighted median	29	1.131	0.968	1.321	0.121
			IVW	29	1.056	0.92	1.213	0.439
			MR-Egger	29	1.137	0.8	1.615	0.475
			(intercept)	29	NA	NA	NA	0.655
β-blockers	SBP	AF	Simple median	7	0.65	0.489	0.865	0.003
			Weighted median	7	0.674	0.533	0.852	0.001
			IVW	7	0.653	0.539	0.791	1.37E-05
			MR-Egger	7	0.615	0.298	1.269	0.188
			(intercept)	7	NA	NA	NA	0.868
β-blockers	SBP	CAD	Simple median	7	0.639	0.523	0.78	1.08E-05
			Weighted median	7	0.665	0.56	0.789	3.22E-06
			IVW	7	0.669	0.581	0.77	2.39E-08
			MR-Egger	7	0.694	0.413	1.167	0.168
			(intercept)	7	NA	NA	NA	0.885
β-blockers	SBP	MI	Simple median	7	0.601	0.39	0.925	0.021
			Weighted median	7	0.604	0.416	0.876	0.008
			IVW	7	0.566	0.419	0.765	0.0002
			MR-Egger	7	0.65	0.211	2.001	0.453
			(intercept)	7	NA	NA	NA	0.801
β-blockers	SBP	HF	Simple median	7	0.633	0.446	0.897	0.01
			Weighted median	7	0.692	0.516	0.927	0.014
			IVW	7	0.701	0.555	0.885	0.003
			MR-Egger	7	0.769	0.298	1.98	0.586

			(intercept)	7	NA	NA	NA	0.842
β-blockers	SBP	Ischemic stroke	Simple median	7	0.838	0.573	1.225	0.361
			Weighted median	7	0.744	0.538	1.028	0.073
			IVW	7	0.915	0.62	1.351	0.654
			MR-Egger	7	0.791	0.174	3.599	0.761
			(intercept)	7	NA	NA	NA	0.844
β-blockers	SBP	CKD	Simple median	7	1.424	0.925	2.19	0.108
			Weighted median	7	1.451	1.012	2.081	0.043
			IVW	7	1.278	0.906	1.803	0.163
			MR-Egger	7	1.454	0.366	5.78	0.595
			(intercept)	7	NA	NA	NA	0.848
β-blockers	SBP	T2D	Simple median	7	1.267	0.873	1.838	0.214
			Weighted median	7	1.398	1.049	1.863	0.022
			IVW	7	1.326	0.897	1.962	0.158
			MR-Egger	7	1.104	0.225	5.412	0.903
			(intercept)	7	NA	NA	NA	0.814
α-blockers	SBP	AF	IVW	1	1.081	0.553	2.117	0.819
		CAD		1	0.523	0.319	0.857	0.01
		MI		1	0.754	0.261	2.176	0.602
		HF		1	1.05	0.474	2.329	0.904
		Ischemic stroke		1	0.57	0.275	1.181	0.13
		CKD		1	2.85	1.071	7.584	0.036
		T2D		1	0.637	0.335	1.212	0.17
Loop diuretics	SBP	AF	Simple median	4	0.419	0.27	0.649	0.0001
			Weighted median	4	0.425	0.273	0.662	0.0002
			IVW	4	0.467	0.29	0.753	0.002
			MR-Egger	4	1.085	0.251	4.691	0.913
			(intercept)	4	NA	NA	NA	0.236
Loop diuretics	SBP	CAD	Simple median	4	0.573	0.431	0.763	0.0001
			Weighted median	4	0.584	0.437	0.78	0.0003
			IVW	4	0.576	0.459	0.725	2.40E-06
			MR-Egger	4	0.61	0.288	1.289	0.195
			(intercept)	4	NA	NA	NA	0.878
Loop diuretics	SBP	MI	Simple median	4	0.652	0.369	1.152	0.141
			Weighted median	4	0.644	0.361	1.147	0.135
			IVW	4	0.623	0.385	1.011	0.055
			MR-Egger	4	0.67	0.138	3.252	0.619
			(intercept)	4	NA	NA	NA	0.926

Loop diuretics	SBP	HF	Simple median	4	0.902	0.568	1.433	0.663
			Weighted median	4	0.95	0.595	1.519	0.832
			IVW	4	0.85	0.555	1.302	0.456
			MR-Egger	4	1.474	0.325	6.688	0.615
			(intercept)	4	NA	NA	NA	0.454
Loop diuretics	SBP	Ischemic stroke	Simple median	4	0.566	0.362	0.884	0.012
			Weighted median	4	0.611	0.387	0.965	0.035
			IVW	4	0.72	0.408	1.27	0.256
			MR-Egger	4	2.201	0.44	11.021	0.337
			(intercept)	4	NA	NA	NA	0.153
Loop diuretics	SBP	CKD	Simple median	4	1.058	0.612	1.828	0.839
			Weighted median	4	1.022	0.584	1.787	0.94
			IVW	4	1.048	0.596	1.844	0.87
			MR-Egger	4	0.673	0.076	5.972	0.722
			(intercept)	4	NA	NA	NA	0.676
Loop diuretics	SBP	T2D	Simple median	4	0.954	0.641	1.42	0.817
			Weighted median	4	1.009	0.683	1.491	0.964
			IVW	4	0.797	0.382	1.666	0.547
			MR-Egger	4	1.725	0.111	26.816	0.697
			(intercept)	4	NA	NA	NA	0.562
Other PSDs	SBP	AF	IVW	1	1.011	0.492	2.075	0.977
		CAD		1	0.974	0.582	1.631	0.92
		MI		1	2.023	0.689	5.941	0.2
		HF		1	0.657	0.284	1.52	0.326
		Ischemic stroke		1	0.396	0.183	0.856	0.018
		CKD		1	2.051	0.744	5.658	0.165
		T2D		1	0.454	0.227	0.908	0.026
ERAs	SBP	AF	IVW	1	0.409	0.216	0.775	0.006
		CAD		1	0.197	0.123	0.315	1.19E-11
		MI		1	0.188	0.071	0.496	0.001
		HF		1	0.687	0.321	1.472	0.334
		Ischemic stroke		1	0.218	0.107	0.443	2.56E-05
		CKD		1	1.234	0.487	3.123	0.657
		T2D		1	0.793	0.425	1.476	0.464
PDE5 inhibitors	SBP	AF	Simple median	5	0.772	0.57	1.047	0.096
			Weighted median	5	0.785	0.598	1.032	0.083
			IVW	5	0.766	0.612	0.958	0.019

			MR-Egger	5	0.823	0.494	1.371	0.454
			(intercept)	5	NA	NA	NA	0.759
PDE5 inhibitors	SBP	CAD	Simple median	5	0.393	0.241	0.64	0.0002
			Weighted median	5	0.338	0.228	0.5	5.76E-08
			IVW	5	0.348	0.199	0.607	0.0002
			MR-Egger	5	0.197	0.054	0.714	0.013
			(intercept)	5	NA	NA	NA	0.336
PDE5 inhibitors	SBP	MI	Simple median	5	0.33	0.162	0.67	0.002
			Weighted median	5	0.33	0.177	0.618	0.001
			IVW	5	0.381	0.2	0.728	0.003
			MR-Egger	5	0.231	0.041	1.295	0.096
			(intercept)	5	NA	NA	NA	0.533
PDE5 inhibitors	SBP	HF	Simple median	5	0.812	0.526	1.256	0.35
			Weighted median	5	0.755	0.513	1.111	0.154
			IVW	5	0.652	0.452	0.94	0.022
			MR-Egger	5	0.362	0.199	0.659	0.001
			(intercept)	5	NA	NA	NA	0.032
PDE5 inhibitors	SBP	Ischemic stroke	Simple median	5	0.663	0.466	0.944	0.023
			Weighted median	5	0.561	0.406	0.777	0.001
			IVW	5	0.588	0.453	0.763	6.55E-05
			MR-Egger	5	0.417	0.218	0.801	0.009
			(intercept)	5	NA	NA	NA	0.26
PDE5 inhibitors	SBP	CKD	Simple median	5	0.73	0.435	1.225	0.234
			Weighted median	5	0.534	0.344	0.831	0.005
			IVW	5	0.532	0.345	0.82	0.004
			MR-Egger	5	0.236	0.118	0.47	3.96E-05
			(intercept)	5	NA	NA	NA	0.01
PDE5 inhibitors	SBP	T2D	Simple median	5	0.951	0.643	1.407	0.802
			Weighted median	5	0.808	0.613	1.066	0.132
			IVW	5	0.951	0.677	1.336	0.772
			MR-Egger	5	0.536	0.315	0.912	0.022
			(intercept)	5	NA	NA	NA	0.019
sGC stimulators	SBP	AF	Simple median	6	0.924	0.696	1.226	0.583
			Weighted median	6	0.938	0.715	1.231	0.644
			IVW	6	0.951	0.755	1.197	0.669
			MR-Egger	6	2.152	0.352	13.159	0.407
			(intercept)	6	NA	NA	NA	0.373
	SBP	CAD	Simple median	6	0.319	0.239	0.426	1.01E-14

			Weighted median	6	0.325	0.243	0.434	2.77E-14
sGC stimulators			IVW	6	0.332	0.236	0.469	3.45E-10
			MR-Egger	6	1.281	0.091	18.074	0.855
			(intercept)	6	NA	NA	NA	0.314
sGC stimulators	SBP	MI	Simple median	6	0.212	0.129	0.348	7.80E-10
			Weighted median	6	0.208	0.129	0.335	1.29E-10
			IVW	6	0.238	0.168	0.337	5.68E-16
			MR-Egger	6	0.378	0.024	5.843	0.486
			(intercept)	6	NA	NA	NA	0.739
sGC stimulators	SBP	HF	Simple median	6	1.131	0.794	1.611	0.495
			Weighted median	6	1.002	0.708	1.418	0.99
			IVW	6	0.987	0.676	1.442	0.947
			MR-Egger	6	18.748	2.222	158.217	0.007
			(intercept)	6	NA	NA	NA	0.006
sGC stimulators	SBP	Ischemic stroke	Simple median	6	0.626	0.458	0.857	0.003
			Weighted median	6	0.629	0.463	0.854	0.003
			IVW	6	0.686	0.54	0.872	0.002
			MR-Egger	6	2.272	0.341	15.137	0.396
			(intercept)	6	NA	NA	NA	0.212
sGC stimulators	SBP	CKD	Simple median	6	0.567	0.373	0.86	0.008
			Weighted median	6	0.585	0.396	0.865	0.007
			IVW	6	0.55	0.398	0.761	0.0003
			MR-Egger	6	0.618	0.045	8.506	0.719
			(intercept)	6	NA	NA	NA	0.93
sGC stimulators	SBP	T2D	Simple median	6	0.715	0.531	0.962	0.027
			Weighted median	6	0.722	0.542	0.962	0.026
			IVW	6	0.793	0.628	1	0.05
			MR-Egger	6	2.022	0.34	12.036	0.439
			(intercept)	6	NA	NA	NA	0.299

Table 3.7 MR analyses on negative control outcomes, using inverse-variant weighted method.

SBP, systolic blood pressure; nSNPs, number of single nucleotide polymorphisms; beta, effect size per 10 mmHg decrease in SBP; lo_ci, lower bound of 95% confidence interval of beta ; up_ci, upper bound of 95% confidence interval of beta; pval, p-value; NA, not applicable; AGTi, Angiotensinogen inhibitor; ACEi, angiotensin-converting enzyme inhibitors; CCBs, calcium channel blockers; Other PSDs, potassium-sparing diuretics (not include aldosterone antagonists); ERAs, endothelin receptor antagonists; PDE5i, phosphodiesterase type 5 inhibitors; sGC; soluble guanylate cyclase

Drug class	nSNP	exposure	outcome	beta	lo_ci	up_ci	pval
ACEi	1	SBP	low hand grip strength	0.046	-0.005	0.098	0.075

ACEi	1	SBP	heel bone mineral density	-0.002	-0.014	0.011	0.82
ACEi	1	SBP	myopia	-0.021	-0.075	0.033	0.444
ACEi	1	SBP	Parkinson disease	0.014	-0.117	0.145	0.835
AGTi	2	SBP	low hand grip strength	-0.014	-0.048	0.02	0.425
AGTi	2	SBP	heel bone mineral density	0.009	0.000	0.017	0.053
AGTi	2	SBP	myopia	0.008	-0.028	0.045	0.666
AGTi	2	SBP	Parkinson disease	-0.063	-0.144	0.017	0.123
α-blockers	1	SBP	low hand grip strength	0.045	-0.029	0.119	0.232
α-blockers	1	SBP	heel bone mineral density	0.018	-0.001	0.036	0.064
α-blockers	1	SBP	myopia	0.067	-0.010	0.145	0.089
α-blockers	1	SBP	Parkinson disease	-0.17	-0.338	-0.002	0.047
β-blockers	7	SBP	low hand grip strength	-0.013	-0.035	0.008	0.23
β-blockers	7	SBP	heel bone mineral density	-0.002	-0.008	0.004	0.493
β-blockers	7	SBP	myopia	0.023	-0.006	0.051	0.117
β-blockers	7	SBP	Parkinson disease	0.025	-0.025	0.076	0.322
PDE5i	5	SBP	low hand grip strength	0.009	-0.015	0.034	0.457
PDE5i	5	SBP	heel bone mineral density	0.013	0.003	0.022	0.009
PDE5i	5	SBP	myopia	-0.031	-0.057	-0.004	0.022
PDE5i	5	SBP	Parkinson disease	0.038	-0.023	0.1	0.223
CCBs	28	SBP	low hand grip strength	0.004	-0.008	0.016	0.517
CCBs	28	SBP	heel bone mineral density	-0.001	-0.005	0.004	0.729
CCBs	28	SBP	myopia	0.003	-0.009	0.015	0.641
CCBs	28	SBP	Parkinson disease	-0.003	-0.038	0.032	0.871
Loop diuretics	4	SBP	low hand grip strength	0.031	-0.004	0.065	0.079
Loop diuretics	4	SBP	heel bone mineral density	-0.008	-0.052	0.037	0.738
Loop diuretics	4	SBP	myopia	-0.049	-0.094	-0.004	0.032
Loop diuretics	4	SBP	Parkinson disease	0.035	-0.068	0.139	0.506
Other PSDs	1	SBP	low hand grip strength	-0.029	-0.106	0.047	0.453
Other PSDs	1	SBP	heel bone mineral density	0.023	0.004	0.042	0.018
Other PSDs	1	SBP	myopia	0.039	-0.042	0.121	0.344
Other PSDs	1	SBP	Parkinson disease	0.119	-0.057	0.294	0.185
ERAs	1	SBP	low hand grip strength	-0.06	-0.131	0.011	0.098
ERAs	1	SBP	heel bone mineral density	0.02	0.002	0.038	0.029
ERAs	1	SBP	myopia	0.031	-0.044	0.107	0.416
ERAs	1	SBP	Parkinson disease	-0.111	-0.298	0.077	0.247
sGC stimulators	6	SBP	low hand grip strength	-0.012	-0.039	0.015	0.38
sGC stimulators	6	SBP	heel bone mineral density	0.009	0.003	0.016	0.005
sGC stimulators	6	SBP	myopia	-0.007	-0.034	0.02	0.618

sGC stimulators	6	SBP	Parkinson disease	-0.05	-0.124	0.023	0.18

Table 3.8 MR analysis results for genetically proxied drug classes on the main outcome, using blood pressure estimates from UKB GWAS.

SBP, systolic blood pressure; CAD indicates coronary artery disease; MI, myocardial infarction; AF, atrial fibrillation; HF, heart failure; CKD, chronic kidney disease; T2D, type 2 diabetes; IVW, inverse-variant weighted method; nSNPs, number of single nucleotide polymorphisms; beta, effect size per 1 SD increase in SBP; lo_ci, lower bound of 95% confidence interval of beta ; up_ci, upper bound of 95% confidence interval of beta; pval, p-value; NA, not applicable; AGTi, Angiotensinogen inhibitor; ACEi, angiotensin-converting enzyme inhibitors; CCBs, calcium channel blockers; Other PSDs, potassium-sparing diuretics (not include aldosterone antagonists); ERAs, endothelin receptor antagonists; PDE5 inhibitors, phosphodiesterase type 5; sGC; soluble guanylate cyclase.

Drug class	nSNPs	exposure	outcome	beta	lo_ci	up_ci	pval
AGTi	2	SBP	AF	0.228	-0.485	0.94	0.532
	2	SBP	CAD	1.045	0.527	1.563	7.73E-05
	2	SBP	MI	1.764	0.717	2.811	0.0009
	2	SBP	HF	-0.12	-0.973	0.733	0.783
	2	SBP	Ischemic stroke	1.155	0.397	1.913	0.003
	2	SBP	CKD	-0.019	-1.002	0.965	0.97
	2	SBP	T2D	-0.371	-1.057	0.315	0.289
ACEi	1	SBP	AF	-0.339	-1.798	1.12	0.648
	1	SBP	CAD	1.404	0.314	2.494	0.012
	1	SBP	MI	0.015	-2.354	2.383	0.99
	1	SBP	HF	1.182	-0.556	2.92	0.182
	1	SBP	Ischemic stroke	1.226	-0.405	2.856	0.141
	1	SBP	CKD	0.383	-1.762	2.528	0.726
	1	SBP	T2D	2.145	0.708	3.583	0.003
CCBs	30	SBP	AF	0.942	0.733	1.151	9.92E-19
	30	SBP	CAD	0.634	0.407	0.861	4.44E-08
	30	SBP	MI	0.568	0.234	0.901	0.001
	30	SBP	HF	0.749	0.501	0.996	3.01E-09
	30	SBP	Ischemic stroke	0.36	0.133	0.587	0.002
	30	SBP	CKD	0.317	0.019	0.615	0.037
	30	SBP	T2D	-0.125	-0.393	0.144	0.363
Beta-blockers	9	SBP	AF	0.807	0.397	1.218	0.0001
	9	SBP	CAD	0.811	0.51	1.111	1.21E-07
	9	SBP	MI	1.104	0.473	1.736	0.001
	9	SBP	HF	0.599	0.07	1.127	0.026
	9	SBP	Ischemic stroke	-0.294	-1.118	0.53	0.484
	9	SBP	CKD	-0.237	-0.874	0.4	0.466
	9	SBP	T2D	-0.662	-1.485	0.16	0.115
Alpha-blockers	1	SBP	AF	-0.17	-1.623	1.284	0.819
	1	SBP	CAD	1.404	0.335	2.474	0.01
	1	SBP	MI	0.61	-1.683	2.904	0.602
	1	SBP	HF	-0.106	-1.83	1.618	0.904

	1	SBP	Ischemic stroke	1.219	-0.36	2.797	0.13
	1	SBP	CKD	-2.268	-4.386	-0.149	0.036
	1	SBP	T2D	0.975	-0.417	2.366	0.17
Loop diuretics	4	SBP	AF	1.749	1.098	2.401	1.40E-07
	4	SBP	CAD	1.201	0.722	1.679	8.58E-07
	4	SBP	MI	0.905	-0.111	1.92	0.081
	4	SBP	HF	0.495	-0.27	1.26	0.205
	4	SBP	Ischemic stroke	0.745	-0.291	1.78	0.159
	4	SBP	CKD	-0.081	-0.999	0.837	0.863
	4	SBP	T2D	0.318	-1.313	1.948	0.703
Other PSDs	1	SBP	AF	-0.105	-1.128	0.917	0.84
	1	SBP	CAD	0.054	-0.668	0.776	0.884
	1	SBP	MI	-0.874	-2.365	0.617	0.251
	1	SBP	HF	0.691	-0.722	2.103	0.338
	1	SBP	Ischemic stroke	1.4	0.319	2.481	0.011
	1	SBP	CKD	-0.657	-2.187	0.873	0.4
	1	SBP	T2D	1.085	0.106	2.063	0.03
ERAs	1	SBP	AF	1.516	0.431	2.6	0.006
	1	SBP	CAD	2.754	1.958	3.55	1.19E-11
	1	SBP	MI	2.837	1.188	4.485	0.001
	1	SBP	HF	0.637	-0.656	1.929	0.334
	1	SBP	Ischemic stroke	2.584	1.381	3.787	2.56E-05
	1	SBP	CKD	-0.356	-1.931	1.218	0.657
	1	SBP	T2D	0.394	-0.661	1.449	0.464
PDE5 inhibitors	5	SBP	AF	0.663	0.135	1.191	0.014
	5	SBP	CAD	2.218	0.525	3.912	0.01
	5	SBP	MI	2.009	0.359	3.66	0.017
	5	SBP	HF	0.858	-0.18	1.896	0.105
	5	SBP	Ischemic stroke	1.26	0.671	1.85	2.76E-05
	5	SBP	CKD	1.434	0.211	2.658	0.022
	5	SBP	T2D	0.047	-0.76	0.854	0.909
sGC stimulators	6	SBP	AF	0.118	-0.352	0.589	0.622
	6	SBP	CAD	2.102	1.181	3.023	7.75E-06
	6	SBP	MI	2.758	1.825	3.69	6.90E-09
	6	SBP	HF	0.044	-0.685	0.774	0.905
	6	SBP	Ischemic stroke	0.725	0.17	1.279	0.01
	6	SBP	CKD	1.168	0.508	1.829	0.001
	6	SBP	T2D	0.382	-0.143	0.907	0.154

Table 3.9 Difference between MR estimates for genetically proxied drug target perturbation and those for genetically proxied SBP lowering generally.

$Beta_{diff}$ indicates difference between MR estimates for SBP and those for drug target; SE_{diff}, standard error of difference; lo_ci, 95% confidence interval lower bound; up_ci, 95% confidence interval upper bound; z, z-score; p, p-value.

Drug class	Exposure	Outcome	$Beta_{diff}$	SE_{diff}	lo_ci	up_ci	z	p
AGTi	SBP	CAD	-0.009	0.012	-0.033	0.014	-0.784	0.433
AGTi	SBP	Ischemic stroke	-0.044	0.02	-0.082	-0.006	-2.245	0.025
CCBs	SBP	CAD	0.005	0.006	-0.008	0.018	0.774	0.439
CCBs	SBP	Ischemic stroke	0.009	0.006	-0.003	0.021	1.448	0.148
CCBs	SBP	HF	-0.01	0.006	-0.023	0.003	-1.536	0.125
CCBs	SBP	AF	-0.029	0.006	-0.04	-0.018	-5.195	2.04E-07
BBs	SBP	CAD	-0.006	0.008	-0.021	0.009	-0.801	0.423
BBs	SBP	MI	-0.028	0.016	-0.058	0.003	-1.76	0.078
BBs	SBP	AF	-0.027	0.01	-0.046	-0.007	-2.649	0.008
LDs	SBP	CAD	-0.021	0.012	-0.044	0.002	-1.756	0.079
LDs	SBP	AF	-0.06	0.024	-0.108	-0.012	-2.458	0.014
ERAs	SBP	CAD	-0.128	0.024	-0.175	-0.081	-5.329	9.86E-08
ERAs	SBP	Ischemic stroke	-0.124	0.036	-0.195	-0.053	-3.425	0.001
PDE5i	SBP	CAD	-0.071	0.028	-0.127	-0.016	-2.509	0.012
PDE5i	SBP	Ischemic stroke	-0.025	0.013	-0.051	0.001	-1.849	0.064
sGCs	SBP	CAD	-0.076	0.018	-0.111	-0.041	-4.291	0.00002
sGCs	SBP	MI	-0.114	0.018	-0.149	-0.079	-6.349	2.16E-10
sGCs	SBP	CKD	-0.045	0.017	-0.078	-0.012	-2.691	0.007

3.4 Discussion

This chapter utilised MR to investigate the causal associations of antihypertensive drug classes with the risk of cardiovascular, diabetic, and renal outcomes. Through large-scale GWAS data, I explored the impact of genetically proxied BP reduction and specific antihypertensive classes on these health outcomes.

Genetically proxied CCBs and BBs demonstrated protective effects against CAD, HF, and AF. CCBs were also protective against ischemic stroke. Genetically proxied ACEi demonstrate protective effects against CAD and T2D at nominally significant level. These findings are in line with existing RCT meta-analyses and previous MR studies, underscoring the potential of these drug classes in mitigating CVD risk (30, 40, 42, 122, 244, 245, 299, 313, 314). I also discover new associations between genetically predicted drug classes and the outcomes: AGT inhibition and a decreased risk of CAD and ischemic stroke; loop diuretic and a decreased risk of CAD; ERA and a decreased risk of CAD and ischemic stroke; PDE5 inhibition and a lower risk of CAD, ischemic stroke, and CKD; sGC stimulation and a

decreased risk of CAD, MI, and CKD. **Table 3.10** illustrates the comparison of my MR findings with results from RCTs and previous MR studies.

3.4.1 Genetic proxies for BP-lowering

The study observed that genetically decreased SBP is universally associated with reduced risk across all main outcomes, reinforcing the notion that BP management is crucial for mitigating cardiovascular and metabolic disease risks. Evidence from systematic reviews and meta-analyses demonstrated the protective effects of lowering BP on major CVD events, including stroke, HF, and CHD, across various baseline blood pressure levels and comorbidities (30, 31). A systematic review and meta-analysis of 123 RCTs with 613,815 participants showed that a reduction in SBP significantly reduced the risk of major cardiovascular events, CHD, stroke, and HF (30). Another meta-analysis of 68 RCTs with 245,885 individuals demonstrated that the risks of stroke, HF, CHD, cardiovascular and all-cause mortality were significantly reduced by BP lowering (31). The association between hypertension and AF were demonstrated in observational studies (315), however, the initiation of BP-lowering therapy did not show a clear reduction in the burden of AF (316). Previous MR studies found increases in genetically proxied SBP to be associated with increased odds of AF, which are complementary to my current study (299, 317). My MR study shows a causal association of genetically increased SBP with an increased risk of T2D, although the MR Egger intercept test detected horizontal pleiotropy. Sensitivity analysis showed a causal association of genetically increased SBP with increased fasting glucose, but did not detect a causal association with HbA1c. The research studies on the association between SBP and T2D yielded inconsistent results. Previous observational studies and a two-sample MR study, which used an older T2D GWAS, reported a significant association (318). Whereas, another MR study utilising data from the UK Biobank did not provide conclusive evidence on the causal association (319). This thesis detected an association between genetically proxied SBP and T2D risk, using the most updated GWAS for T2D and SBP. The potential causal association can be due to elevated BP causing vasoconstriction, which impairs blood flow and glucose disposal in peripheral tissue such as skeletal muscles or adipose, potentially leading to T2D (319). High BP has long been recognized as a risk factor for kidney function decline based on observational studies. However, the impact of intensive blood pressure control on kidney health remains uncertain when considering the results from RCTs (320, 321). In current MR study, I detected causal associations of SBP with CKD risk. However, I did not detect significant association of SBP with eGFRcrea and eGFRcys, which is consistent with a previous bidirectional MR study that investigated causal effects of blood pressure on kidney function (322).

There were some drug classes whose MR estimates differ from those for genetically proxied SBP lowering, including genetically proxied AGT inhibition (on ischemic stroke), CCBs (on AF), beta-blockers (on AF), loop diuretics (on AF), ERAs (on CAD and ischemic stroke), PDE5 inhibitors (on CAD), and sGC stimulators (on CAD, MI, and CKD). These heterogeneous effects suggest that the benefits of these drug targets might not be solely due to lowering SBP but also additional mechanisms. Whereas, for the remaining drug classes, their associations with the outcomes appear to be mediated through their SBP-lowering effect.

3.4.2 Genetic proxies for calcium channel blockers

The chapter demonstrated the causal associations of genetically proxied CCBs with a decreased risk of CAD, HF, ischemic stroke, and AF. CCBs demonstrated superiority over other first-line antihypertensive medications in terms of stroke prevention in meta-analysis studies that compared head-to-head first-line antihypertensive drug classes (30, 41). The protective effects of CCBs against stroke (all types of stroke) were demonstrated in a previous systematic review and meta-analysis of 10 RCTs with 30,359 patients, with risk ratio (RR) of 0.66 (0.58 – 0.75) per 6.3 mmHg decrease in SBP (40). Additionally, in that study, they conducted sensitivity meta-analysis of 4 RCTs, in which they enrolled hypertensive patients with no or minimal background therapy and then found significant evidence for the protective effects of CCBs against CAD (RR, 0.74 (0.58–0.94), per 6.3 mmHg decrease in SBP) and HF (RR, 0.68 (0.49–0.95), per 6.3 mmHg decrease in SBP). The causal effects of CCBs on CAD, HF, AF, and ischemic stroke were demonstrated in previous 2-sample MR studies (122, 244, 299, 313, 314). In comparison with the previous MR studies, my current MR study integrated data from newer GWAS with larger sample size, particularly for CAD and ischemic stroke. Genetic instruments in the previous MR studies were selected as SNPs located in the genomic regions corresponding to the drug target genes and their regulatory regions and strongly associated with BP trait. The effect estimates of genetically predicted CCBs on the outcomes closely resembled those obtained from my MR study. The alignment with previous RCTs and MR studies reinforces the robustness of my findings and underscores the relevance of genetic predictors in understanding drug effects.

Table 3.10 Comparison of results from my MR study with those from RCT Meta-Analysis and previous MR studies.

SBP, systolic blood pressure; CAD indicates coronary artery disease; MI, myocardial infarction; AF, atrial fibrillation; HF, heart failure; CKD, chronic kidney disease; T2D, type 2 diabetes; RCT, randomisation controlled trial; SNP, single nucleotide polymorphism; OR, odd ratio; RR, relative risk; HR, hazard ratio; n, sample size; AGTi, Angiotensinogen inhibitor; ACEi, angiotensin-converting enzyme inhibitors; CCBs, calcium channel blockers; Other PSDs, potassium-sparing diuretics (not include aldosterone antagonists); ERAs, endothelin receptor antagonists; PDE5 inhibitors, phosphodiesterase type 5; sGC; soluble guanylate cyclase; (*) retrospective observational cohort study.

Drug	Outcomes	Meta-analysis/ RCT results	Previous MR	This current MR
Beta-blockers	CAD	• n = 18,724 (40) • CAD-RR: 0.88 (0.77–1.01), per -10.55 mmHg	• SBP GWAS (n = 757,601), CAD GWAS (60,801 cases; 123,504 controls) (122) • 6 SNPs for BBs. • CAD-RR, 0.62 (0.47–0.81); P=4×10^{-4}; per - 9.51 mmHg SBP	• SBP GWAS (n = 757,601), CAD GWAS (181,522 cases; 984,168 controls) • 7 SNPs for BBs • CAD-OR, 0.67 (0.58 - 0.77); p = 2.39 × 10^{-8}; per -10 mmHg SBP
	Ischemic stroke	• n = 18,724 (40) • Stroke-RR: 0.77 (0.61–0.97), per -10.5 mmHg SBP	• SBP GWAS (n = 757,601), ischemic stroke (60,341 cases; 454,450 controls) (313) • 8 SNPs for BBs • Ischemic stroke-OR, 1.02 (0.86 - 1.22); p = 0.795	• SBP GWAS (n = 757,601), ischemic stroke GWAS (62,100 cases; 1,234,808 controls) • 7 SNPs for BBs • Ischemic stroke-OR, 0.92 (0.62 – 1.35); p = 0.65; per -10 mmHg SBP
			• SBP GWAS (n = 757,601), stroke GWAS (67,162 cases; 454,450 controls) (122) • 6 SNPs for BBs • Stroke-RR, 0.91 (0.73–1.14); P=0.41; per - 9.51 mmHg SBP	
	HF	• n = 18,724 (40) • HF-RR: 0.57 (0.35–0.91), per -14.8 mmHg SBP	• SBP GWAS (n = 757,601), HF GWAS (47,309 cases; 930,014 controls) (244) • 5 SNPs for BBs • HF-OR, 0.62 (0.45–0.84); p = 0.002; per -10 mmHg SBP	• SBP GWAS (n = 757,601); HF GWAS (47,309 cases; 930,014 controls) • 7 SNPs for BBs

		• SBP GWAS (n = 757,601), HF GWAS (47,309 cases; 930,014 controls) (314) • 19 SNPs for BBs • HF-OR, 0.71 (0.62 to 0.82); p < .001; per -10 mmHg SBP	• HF-OR, 0.70 (0.56 - 0.89); p = 0.003; per -10 mmHg SBP
	MI		• SBP GWAS (n = 757,601); MI GWAS (43,676 cases; 128,199 controls) • 7 SNPs for BBs • MI-OR, 0.57 (0.42 - 0.77), p = 0.0002, per -10mmHg in SBP
	AF	• SBP GWAS (n = 757,601), AF GWAS (65,446 cases; 522,744 controls; 84.2% European) (299) • 5 SNPs for BBs • AF-OR, 0.61 (0.46–0.81); P=6×10^{-4}; per -10 mmHg SBP	• SBP GWAS (n = 757,601); AF GWAS (60,620 cases; 970,216 controls) • 7 SNPs for BBs • AF-OR, 0.65 (0.54 - 0.79), p = 1.37×10^{-5}
	CKD		• SBP GWAS (n = 757,601); CKD GWAS (41,395 cases; 439,303 controls) • 7 SNPs for BBs • CKD-OR, 1.28 (0.91 - 1.80), p = 0.16; per -10 mmHg SBP
	T2D	• SBP GWAS (n = 757,601), T2D GWAS (74,124 cases; 824,006 controls) (245) • 2 SNPs for BBs • T2D-OR 1.05 (1.02–1.09), p = $8.7\text{x}10^{-4}$, per mmHg lower in SBP	• SBP GWAS (n = 757,601); T2D GWAS (80,154 cases; 853,816 controls) • 7 SNPs for BBs • T2D-OR, 1.33 (0.90-1.96), p = 0.16; per -10 mmHg SBP

CCBs	CAD	• n = 30,359 (40) • CAD-RR: 0.83 (0.65-1.05), per -5.9 mmHg SBP. • CAD-RR: 0.74 (0.58–0.94), per -6.3 mmHg SBP, in the sensitivity analysis	• SBP GWAS (n = 757,601), CAD GWAS (60,801 cases; 123,504 controls) (122) • 24 SNPs identified for CCBs. • CAD-RR, 0.73 (0.64–0.84); $P=6\times10^{-6}$, per -8.9 mmHg SBP	• SBP GWAS (n = 757,601), CAD GWAS (181,522 cases; 984,168 controls) • 29 SNPs for CCBs • CAD-OR, 0.75 (0.66 - 0.84), p = 1.1×10^{-6}; per -10 mmHg SBP
	Ischaemic stroke	• n =30,359 (40) • Stroke-RR 0.66 (0.58-0.75), per -6.1 mmHg SBP. • Stroke-RR 0.63 (0.53–0.76), per -6.5 mmHg SBP, in sensitivity analysis	• SBP GWAS (n = 757,601), ischemic stroke (60,341 cases; 454,450 controls) (313) • 60 SNPs identified for CCBs • Ischemic stroke-OR 0.71 (0.65-0.76), p = $4.3\text{x}10^{-20}$ per -10 mmHg SBP	• SBP GWAS (n = 757,601), ischemic stroke GWAS (62,100 cases; 1,234,808 controls) • 29 SNPs for CCBs • Ischemic stroke-OR, 0.82 (0.74 - 0.92), p = 0.0006; per -10 mmHg SBP
			• SBP GWAS (n = 757,601), stroke GWAS (67,162 cases; 454,450 controls) (122) • 24 SNPs identified for CCBs • Stroke-RR, 0.75 (0.66–0.84); $P=1\times10^{-6}$, per -8.9 mmHg SBP	
	HF	• n = 30,359 (40) • HF-RR 0.81 (0.60-1.10), per -5.8 mmHg SBP. • HF-RR 0.68 (0.49–0.95), per -6.3 mmHg SBP, in sensitivity analysis	• SBP GWAS (n = 757,601), HF GWAS (47,309 cases; 930,014 controls) (244) • 20 SNPs identified for CCBs • HF-OR 0.730 (0.625–0.851), p = $6.23\text{x}10^{-5}$, per -10 mmHg SBP	• SBP GWAS (n = 757,601); HF GWAS (47,309 cases; 930,014 controls) • 29 SNPs for CCBs • HF-OR, 0.73 (0.65 - 0.83), p = 3.8×10^{-7}; per -10 mmHg SBP
			• SBP GWAS (n = 757,601), HF GWAS (47,309 cases; 930,014 controls) (314) • 57 SNPs identified for CCBs • HF-OR 0.71 (0.65 to 0.78), p < .001, per -10 mmHg SBP	

	MI			• SBP GWAS (n = 757,601); MI GWAS (43,676 cases; 128,199 controls) • 29 SNPs for CCBs • MI-OR, 0.78 (0.66 - 0.92), p = 0.003, per -10 mmHg SBP
	AF		• SBP GWAS (n = 757,601), AF GWAS (65,446 cases; 522,744 controls; 84.2% European) (299) • 20 SNPs identified for CCBs • AF-OR, 0.66 (0.57–0.76), p = 8×10^{-9}; per -10 mmHg SBP	• SBP GWAS (n = 757,601); AF GWAS (60,620 cases; 970,216 controls) • 29 SNPs for CCBs • AF-OR, 0.64 (0.57 - 0.70), p = 5.35×10^{-18}; per -10 mmHg SBP
	CKD			• SBP GWAS (n = 757,601); CKD GWAS (41,395 cases; 439,303 controls) • 29 SNPs for CCBs • CKD-OR, 0.88 (0.76 - 1.01), p = 0.07; per -10 mmHg SBP
	T2D		• SBP GWAS (n = 757,601), T2D GWAS (74,124 cases; 824,006 controls) (245) • 8 SNPs identified for CCBs • T2D-OR, 0.99 (0.97 - 1.02), p = 0.57, per -10 mmHg SBP	• SBP GWAS (n = 757,601); T2D GWAS (80,154 cases; 853,816 controls) • 29 SNPs for CCBs • T2D-OR, 1.06 (0.92 - 1.21); p = 0.44; per -10 mmHg SBP
ACE inhibitors	CAD	• n = 35707 (40) • CAD-RR 0.87 (0.79–0.97), per -4.2 mmHg SBP	• SBP GWAS (n = 757,601), CAD GWAS (60,801 cases; 123,504 controls) (122) • 1 SNP identified for ACEi. • CAD-RR, 0.67 (0.16–2.56); P=0.58; per -21.14 mmHg SBP	• SBP GWAS (n = 757,601), CAD GWAS (181,522 cases; 984,168 controls) • 1 SNP for ACEi • CAD-OR, 0.65 (0.46 - 0.91); p = 0.01; per -10 mmHg SBP

	Ischemic stroke	• n = 35707 (40) • Stroke-RR 0.80 (0.69–0.93), per -4.2 mmHg SBP	• SBP GWAS (n = 757,601), stroke GWAS (67,162 cases; 454,450 controls) (122) • 1 SNP identified for ACEi. • Stroke-RR, 0.21 (0.06–0.72); P=0.01; per -21.14 mmHg SBP	• SBP GWAS (n = 757,601), ischemic stroke GWAS (62,100 cases; 1,234,808 controls) • 1 SNP for ACEi • ischemic stroke-OR, 0.68 (0.41 - 1.13); p = 0.14; per -10 mmHg SBP
	HF	• n = 35707 (40) • HF-RR 0.79 (0.66–0.93), per -3.5 mmHg SBP	• SBP GWAS (n = 757,601), HF GWAS (47,309 cases; 930,014 controls) (314) • 1 SNP identified for ACEi. • HF-OR, 0.69 (0.40 to 1.19), p = 0.182; per -10 mmHg SBP	• SBP GWAS (n = 757,601); HF GWAS (47,309 cases; 930,014 controls) • 1 SNP for ACEi • HF-OR, 0.69 (0.41 - 1.19); p = 0.18; per -10 mmHg SBP
	MI			• SBP GWAS (n = 757,601); MI GWAS (43,676 cases; 128,199 controls) • 1 SNP for ACEi • MI-OR, 0.99 (0.48 - 2.07); p = 0.99; per -10 mmHg SBP
	AF			• SBP GWAS (n = 757,601); AF GWAS (60,620 cases; 970,216 controls) • 1 SNP for ACEi • AF-OR, 1.11 (0.71 - 1.75), p = 0.65; per -10 mmHg SBP
	CKD			• SBP GWAS (n = 757,601); CKD GWAS (41,395 cases; 439,303 controls) • 1 SNP for ACEi • CKD-OR, 0.89 (0.46 - 1.73); p = 0.72; per -10 mmHg SBP

	T2D	• n = 5269 (42) • T2D or death HR : 0.91 (0.81 - 1.03); p = 0.15 • Regression to normoglycemia HR: 1.16 (1.07 - 1.27); p = 0.001	• SBP GWAS (n = 757,601), T2D GWAS (74,124 cases; 824,006 controls) (245) • 1 SNP identified for ACEi. • T2D-OR 0.85 (0.78–0.93), p = $2.8x10^{-4}$; per -1 mmHg SBP	• SBP GWAS (n = 757,601); T2D GWAS (80,154 cases; 853,816 controls) • 1 SNP for ACEi • T2D-OR, 0.51 (0.33 - 0.80), p = 0.003, per -10 mmHg SBP
Alpha-blockers	CAD	ALLHAT RCT (323) on 33,357 participants, the doxazosin arm was terminated early due to 2-fold higher incidence of congestive heart failure in patients with doxazosin than patients with chlorthalidone.		CAD-OR, 0.523 (0.319 - 0.857), p = 0.01, per -10 mmHg SBP
	Ischemic stroke			Ischemic stroke-OR, 0.57 (0.275 - 1.181), p = 0.13, per -10 mmHg SBP
	HF			HF-OR, 1.05 (0.474 - 2.329), p = 0.9; per -10 mmHg SBP
	MI			MI-OR, 0.754 (0.261 - 2.176), p = 0.6; per -10 mmHg SBP
	AF			AF-OR, 1.081 (0.553 - 2.117), p = 0.81; per -10 mmHg SBP
	CKD			CKD-OR, 2.85 (1.071 - 7.584), p = 0.04, per -10 mmHg SBP
	T2D			T2D-OR, 0.637 (0.335 - 1.212), p = 0.17; per -10 mmHg SBP
Loop diuretics	CAD			CAD-OR, 0.576 (0.459 - 0.725), p = $2.4x10^{-06}$ per -10 mmHg SBP
	Ischemic stroke			Ischemic stroke-OR, 0.72 (0.408 - 1.27), p = 0.26; per -10 mmHg SBP
	HF			HF-OR, 0.85 (0.555 - 1.302), p = 0.46; per -10 mmHg SBP

	MI		MI-OR, 0.623 (0.385 - 1.011), p = 0.06; per -10 mmHg SBP
	AF		AF-OR, 0.467 (0.29 - 0.753), p = 0.001; per -10 mmHg SBP
	CKD		CKD-OR, 1.048 (0.596 - 1.844), p = 0.87; per -10 mmHg SBP
	T2D		T2D-OR, 0.797 (0.382 - 1.666), p = 0.55; per -10 mmHg SBP
Other PSDs	CAD		CAD-OR, 0.974 (0.582 - 1.631); p = 0.92; per -10 mmmHg SBP
	Ischemic stroke		Ischemic stroke-OR, 0.396 (0.183 - 0.856); p = 0.02; per -10 mmHg SBP
	HF		HF-OR, 0.657 (0.284 - 1.52); p = 0.33; per -10 mmHg SBP
	MI		MI-OR, 2.023 (0.689 - 5.941); p = 0.2; per -10 mmHg SBP
	AF		AF-OR, 1.011 (0.492 - 2.075); p = 0.98; per -10 mmHg SBP
	CKD		CKD-OR, 2.051 (0.744 - 5.658); p = 0.17; per -10 mmHg SBP
	T2D		T2D-OR, 0.454 (0.227 - 0.908); p = 0.03; per -10 mmHg SBP
ERAs	CAD	Phase 3 trial of aprocitentan (102), in patients with resistant	CAD-OR, 0.197 (0.123 - 0.315), p = $1.19x10^{-11}$ per -10 mmHg SBP

	Ischemic stroke	hypertension, aprocitentan was well tolerated and superior to placebo in lowering SBP at week 4 and sustained effect at week 40.		Ischemic stroke-OR, 0.218 (0.107 - 0.443), p = $2.56x10^{-05}$, per -10 mmHg SBP
	HF			HF-OR, 0.687 (0.321 - 1.472), p = 0.33; per -10 mmHg SBP
	MI			MI-OR, 0.188 (0.071 - 0.496), p = 0.0007; per -10 mmHg SBP
	AF			AF-OR, 0.409 (0.216 - 0.775), p = 0.006; per -10 mmHg SBP
	CKD			CKD-OR, 1.234 (0.487 - 3.123), p = 0.66; per -10 mmHg SBP
	T2D			T2D-OR, 0.793 (0.425 - 1.476), p = 0.46; per -10 mHg SBP
PDE5 inhibitors	CAD			CAD-OR, 0.348 (0.199 - 0.607); p = 0.0002; per -10 mmHg SBP
	Ischemic stroke	n = 72,948 men (*) (324) Stroke HR, 0.92 (0.79 - 1.08); p = 0.318		Ischemic stroke-OR, 0.588 (0.453 - 0.763); p = $6.55x10^{-05}$; per -10 mmhg SBP
	HF	n = 72,948 men (*) (324) HF-HR, 0.85 (0.73 -1); p = 0.05		HF-OR, 0.652 (0.452 - 0.94); p = 0.02; per -10 mmHg SBP
	MI	n = 72,948 men (*) (324) MI-HR, 0.95 (0.81 - 1.11), p = 0.529		MI-OR, 0.381 (0.2 - 0.728); p = 0.003; per -10 mmHg SBP

	AF		AF-OR, 0.766 (0.612 - 0.958); p = 0.02; per -10 mmHg SBP
	CKD		CKD-OR, 0.532 (0.345 - 0.82);p = 0.004; per -10 mmHg SBP
	T2D		T2D-OR, 0.951 (0.677 - 1.336), p = 0.77; per -10 mmHg SBP
sGC stimulators	CAD		CAD-OR, 0.332 (0.236 - 0.469), p = $3.45x10^{-10}$ per -10 mmHg SBP
	Ischemic stroke		Ischemic stroke-OR, 0.686 (0.54 - 0.872); p = 0.002; per -10 mmHg SBP
	HF	Phase 3 RCT (89) of vericiguat on 5050 patients with chronic heart failure and an ejection fraction of < 45%, comparing vericiguat (10 mg once daily) vs placebo. The primary outcome is a composite of death from cardiovascular causes or first hospitalization for heart failure. HR: 0.90 (0.82 - 0.98); p = 0.02	HF-OR, 0.987 (0.676 - 1.442); p = 0.94; per -10 mmHg SBP
	MI		MI-OR, 0.238 (0.168 - 0.337); p = $5.68x10^{-16}$; per -10 mmHg SBP
	AF		AF-OR, 0.951 (0.755, 1.197); p = 0.67; per -10 mmHg SBP

	CKD			CKD-OR, 0.55 (0.398 - 0.761); p = 0.0003, per -10 mmHg SBP
	T2D			T2D-OR, 0.793 (0.628 - 1); p = 0.05, per -10 mmHg SBP
AGTi	CAD	Phase 1 trial (60) on 107 patients showed that dose-dependent decreases in serum angiotensinogen levels and 24-hour ABPM were sustained for up to 24 weeks after a single subcutaneous dose of zilebesiran of 200 mg or more.		CAD-OR, 0.647 (0.515 - 0.814), p = 0.0001; per -10 mmHg SBP
	Ischemic stroke			Ischemic stroke-OR, 0.487 (0.333 - 0.712), p = 0.0002; per -10 mmHg SBP
	HF			HF-OR, 0.814 (0.564 - 1.175); p =0.27; per -10 mmHg SBP
	MI			MI-OR, 0.468 (0.29 - 0.755); p = 0.002; per -10 mmHg SBP
	AF			AF-OR, 0.758 (0.555 - 1.035), p = 0.08; per -10 mmHg SBP
	CKD			CKD-OR, 0.711 (0.46 - 1.099), p = 0.124; per -10 mmHg SBP
	T2D			T2D-OR, 0.951 (0.707 - 1.279), p = 0.74; per -10 mmHg SBP

3.4.3 Genetic proxies for beta-blockers

Genetically proxied beta-blockers demonstrates protective effects against AF, CAD, and HF. Consistent with my findings, prior MR investigations have also detected such associations (122, 299). Meta-analysis of 18,724 patients from 5 RCTs showed causal effects of beta-blockers with a reduced risk of stroke (all types of stroke) and HF (40). When compared with other antihypertensive drugs, beta-blockers were inferior in the prevention of stroke (30, 41). My study did not uncover a significant association between genetically proxied beta-blockers and ischemic stroke. Consistent with my results, earlier MR investigations that explored the effects of blood pressure-lowering drugs on stroke risk subtypes demonstrated that CCB - but not beta-blockers - were linked to a reduced risk of stroke and ischemic stroke (313).

3.4.4 Genetically proxies for ACE inhibitors

The protective effect of ACEi on T2D, highlighted through their association with lower HbA1c levels and decreased fasting glucose, aligns with RCT evidence suggesting their favourable metabolic profile. However, the observed nominal associations and the challenges posed by pleiotropy emphasize the need for cautious interpretation and further investigation. In line with my findings, a two-sample MR study, which integrated an older T2D GWAS meta-analysis with a smaller sample size, detected such protective effect of ACEi on T2D (245). In RCTs, ACEi exhibit favourable effects on blood glucose levels, potentially reducing risk of diabetes (42). As RCTs are usually conducted on patients with hyperglycaemia and normally have short follow-up periods (42), evidence from MR analysis highlights that the benefits of ACEi on diabetes may also apply to the broader general population. Genetically proxied ACEi were nominally associated with a decreased risk of CAD. Meta-analysis of 12 RCTs (with 35,707 patients) demonstrated that ACEi significantly reduced the risk of CAD, stroke, and HF. My study detected association of ACEi and CAD risk, while the associations with ischemic stroke and HF were not significant. As only one genetic instrument was identified for ACE inhibition, the MR analysis may not have sufficient statistical power to detect significant associations with the other CVD outcomes.

3.4.5 Genetic proxies for AGT inhibition

Angiotensinogen (AGT), a hepatic peptide hormone, serves as the exclusive precursor for all angiotensin peptides. Zilebesiran (60), an investigational RNA interference therapeutic agent, exhibits strong binding affinity to the hepatic asialoglycoprotein receptor. Its specific mechanism aims to reduce hepatic angiotensinogen messenger RNA (mRNA) levels, leading to decreased angiotensinogen production and subsequent inhibition of the renin-angiotensin-aldosterone system. In phase 1 study, zilebesiran demonstrated potential to

modulate blood pressure (60). Dose-dependent reductions in serum angiotensinogen levels persisted, alongside sustained improvements in 24-hour ABPM. These effects endured for up to 24 weeks following administration of a single subcutaneous dose of 200 mg or higher. Currently, the phase 2 study is ongoing, aiming to further elucidate the therapeutic potential of zilebesiran (325). In my current study, I provide additional evidence supporting the investigation of AGT inhibitors. Specifically, I detected protective effects of genetically proxied AGT inhibition on CAD and ischemic stroke. These findings underscore the potential benefits of targeting AGT as a therapeutic strategy for improving cardiovascular outcomes.

3.4.6 Genetic proxies for loop diuretic

This thesis found a protective effect of loop diuretics against CAD and AF risk. The effect on CAD appears to be mediated through lowering SBP, whereas, the effect on AF might involve not only lowering SBP but also other mechanisms. Nonetheless, the interpretation of protective effects needs to be cautious and further investigation is required due to potential pleiotropy bias in MR. Currently, there is no evidence from clinical trials about the protective effects of loop diuretics on CVDs. A recent clinical trial highlights the role of UMOD genotype-guided use of loop diuretics in hypertension management. The *UMOD* gene, expressed exclusively in the thick ascending limb of the loop of Henle, is involved in BP regulation through its interaction with NKCC2, a target of loop diuretics. Patients with uncontrolled hypertension who have the *UMOD*-increasing AA genotype showed a greater response to the drug compared to those with the AG/GG genotypes (326).

3.4.7 Genetically proxies for phosphodiesterase-5 inhibitors

This research found that genetically proxied PDE5i drugs exhibit associations with CAD and ischemic stroke. Additionally, there is a nominal association with decreased risks of MI, HF, AF, and CKD. PDE5 inhibitors (e.g.sildenafil) have traditionally been used for the treatment of erectile dysfunction and pulmonary arterial hypertension. Sildenafil was initially developed for treating angina. In preclinical studies, sildenafil demonstrated vasodilatory effects, prevented platelet aggregation, and inhibited thrombus reformation in a damaged carotid artery. However, sildenafil appeared to be less promising due to its relatively short half-life and interaction with nitrates. One common side effect of sildenafil was to induce penile erections, which led to its development as a treatment for erectile dysfunction (327). PDE5 inhibitors exert their effects by blocking the degradative action of cGMP-specific phosphodiesterase type 5, which results in elevated cGMP levels and prolonged signalling actions of NO. This pharmacological mechanism promotes vasodilation and enhances blood flow not only in the vasculature of the genitalia but also in the vasculature of all body system

(328, 329). Additionally, previous studies unveiled an interaction between PDE5 inhibitors and endothelial nitric oxide synthase (NOS3) (330). The spatial proximity of PDE5A to caveolae facilitates a feedback loop between endothelial PDE5A and NOS3. Treatment of human endothelium with PDE5 inhibitors enhanced NOS3 activity, while PDE5A overexpression reduced NOS3 activity and endothelium-dependent vasodilation. Beyond their established role in sexual health, PDE5 inhibitors hold promise in influencing endothelial function and potentially impacting cardiovascular outcomes. Indeed, emerging evidence from both animal models and observational studies suggests that these drugs may offer cardioprotective effects (327, 331, 332). In a retrospective nationwide cohort study (324) involving a large population of men with erectile dysfunction (23,186 cases and 48,682 controls), PDE5 inhibitors were linked to a lower incidence of coronary revascularisation, HF, unstable angina, and cardiovascular death. Men without pre-existing CAD but with cardiovascular risk factors at baseline also demonstrated a similar pattern of benefits. In lines with these investigations, my current study supports the protective effects of genetically proxied PDE5i on CVD outcomes. Additionally, I also so detected a nominal association with reduced risk of CKD, increased eGFRcrea, and increased eGFRcys. PDE5 is highly expressed in various renal components, including glomeruli, mesangial cells, cortical tubules, and inner medullary collecting duct. Therefore, it is proposed that PDE5 inhibition can lead to the enhanced function of the NO-cGMP pathway in the kidney, alleviating the NO deficit associated with CKD (333). PDE5 inhibitors exert kidney protective effects in experimental models of CKD and diabetic kidney disease (334). In 5/6 nephrectomized rats, initiating 8 weeks of treatment with a PDE5 inhibitor immediately after nephrectomy prevented hypertension development, ameliorated kidney injury, and reduced proteinuria (335). This robust kidney protection diminished if the treatment with PDE5 inhibition was delayed until 4 weeks after nephrectomy, by which time proteinuria was already evident. Their kidney protective effects were thought to extend far beyond their antihypertensive effects, suggesting renoprotective properties (334). PDE5 inhibitors exhibit potent antiproliferative effects that prevent mesangial cell proliferation and extracellular matrix expansion - effects not achievable through comparable blood pressure reduction using hydralazine (336). Furthermore, studies highlight the beneficial impact of PDE5 inhibition on renal cell apoptosis (337), oxidative stress (338), and inflammation (339) in CKD models. A phase II trial, which involved 256 patients with T2D and detected nephropathy, found that once-daily treatment with a long-acting PDE5 inhibitor for 12 weeks significantly reduced albuminuria (340). This kidney-protective effect persisted despite concurrent administration with RAAS blockers and remained unaffected by any changes in blood pressure or glomerular filtration rate.

3.4.8 Genetic proxies for endothelin receptor antagonists

This research demonstrates a significant association between genetically proxied endothelin receptor antagonism (ERA) and a decreased risk of CAD, and ischemic stroke. Moreover, there is a nominal association with a decreased risk of MI, and AF. Endothelin-1 (ET-1), a 21-amino acid peptide, possesses both mitogenic and vasoconstricting properties. It is elevated in patients with atherosclerosis and in the coronary circulation of those with early atherosclerosis and coronary endothelial dysfunction. ET-1 plays an important role in vascular tone regulation through two primary receptors: endothelin-A (ETA) receptors and endothelin-B (ETB) receptors. ETA receptors, situated on vascular smooth muscle, play a crucial role in mediating vasoconstriction. ETB receptors mediate dilation via nitric oxide release in endothelial cells and contribute to constriction in smooth muscle cells. To date ERAs have been used to treat pulmonary arterial hypertension, however, in recent years translational research has renewed interest in the ET system as a therapeutic target across various CVDs including resistant hypertension, microvascular angina, and post-coronavirus disease 2019 conditions (93, 102, 341, 342). A multicentre, phase 3 trial (the PRECISION trial), investigated the efficacy of aprocitentan, a dual ETA/ETB receptor antagonist, in patients with resistant hypertension. After a 4-week double-blind phase, the treatment arm experienced a significant reduction in SBP. This improvement was maintained throughout a 32-week non-placebo-controlled phase with 25 mg aprocitentan and was reversed after a 12-week withdrawal phase where patients were re-randomised to receive either 25 mg aprocitentan or placebo (102).

3.4.9 Genetic proxies for sGC stimulators

Genetically proxied sGC stimulation is associated with a decreased risk of CAD, MI, and CKD, and nominally associated with a decreased risk of ischemic stroke and T2D. There is also a nominal association with increased eGFRcrea and eGFRcys in the sensitivity analysis. sGC stimulators are a class of small molecule drugs that synergistically enhance sGC enzyme activity NO. sGC, an enzyme present in the cardiopulmonary system, catalyzes the conversion of guanosine-5′-triphosphate (GTP) to cGMP. When NO binds to the prosthetic heme group of sGC, it induces a conformational change, leading to enzymatic activation and increased cGMP production. cGMP plays central roles in cellular and physiological responses including but not limited to vasodilation. A couple of sGC stimulators have been approved for the treatment of pulmonary hypertension (riociguat (87)) and chronic HFrEF (vericiguat (89)). Novel sGC stimulators and activators might be a potential new approach for the treatment of other cardiovascular and kidney disorders (343, 344). As deficiency of cGMP is implicated in the pathogenesis of CKD, sGC stimulators and activators

are being investigated in preclinical models under conditions where nitric oxide is deficient. In preclinical models with different etiologies of CKD, treatment with sGC stimulators and activators has demonstrated consistently kidney-protective effects (345). These favourable effects seem to be largely independent of any blood pressure reduction (344).

3.4.10 Strengths and limitations

This study has several strengths. Genetic variants were used as instruments to investigate the effect of antihypertensive drugs leveraging existing data obtained from the latest large-scale GWAS studies. A range of sensitivity analyses were applied to exclude pleiotropic effects and other biases. This approach can avoid the time and resource-intensive nature of conducting RCTs and circumvent the limitations of potential confounding and reverse causation that can arise from conventional observational methods. While previous studies (122, 244-246, 299, 313) have utilised a comparable methodology to examine the impact of major antihypertensive drugs on CVDs, my study stands out as the first to carry out an exhaustive exploration of all antihypertensive medications across a broad spectrum of diseases, including cardiovascular, diabetes, and kidney diseases.

This study has limitations. Certain antihypertensive drug classes target specific tissue, e.g. Zilebesiran targets on hepatic production of angiotensinogen. In this chapter, the tissue-specific effects have not been accounted for but will be investigated in the next chapter using summary-based MR. Genetic variants generally exert smaller effects that accumulate throughout the entire life course, whereas pharmacological interventions are typically administered later in life and tend to produce more pronounced effects. Thus, MR effect may differ from the effect of drugs administered at a specific time and for a shorter duration. Another limitation of MR is the validation of IV assumptions. While the first assumption can be validated, the second and third assumptions can be disproved but not fully verified, especially with two-sample MR. A less stringent LD criterion for clumping might have increased the number of genetic variants, however, this could also have decreased the sensitivity and specificity of the analysis due to the potential introduction of weak instrument bias and the inclusion of invalid instruments. Moreover, only one single genetic instrument identified for ACEi, alpha-blockers, and ERAs, which might not offer sufficient statistical power to detect significant associations. There is no genetic instrument for some drug classes such as thiazides, angiotensin II receptor blockers, and aldosterone antagonists. Future studies that include larger GWAS datasets for blood pressure could potentially identify such variants, and therefore could provide more profound understanding into the varying effects of different antihypertensive drug classes. Finally, GWAS data used in my research was restricted to individuals of European descent, potentially limiting the

generalizability of my results to this specific group. There is evidence for varying responses to different classes of antihypertensive drugs across ethnicities, an aspect not investigated in my current study. The availability of large-scale GWAS datasets from more ethnically diverse populations will allow future MR studies to investigate potential ethnic differences.

3.5 Conclusion

In conclusion, this research corroborates the established associations of major antihypertensive drugs (ACEi, beta-blockers, CCBs) with a range of CVDs, which were obtained from earlier RCTs and MR studies. This consistency strengthens the validity of my findings and aligns with the existing knowledge. The novel associations identified in this study, particularly with emerging drug targets like PDE5 inhibitors, ERAs, and sGC stimulators, provide a foundation for the development of new therapeutic strategies. My findings support the utility of MR as a complementary tool to RCTs and observational studies, offering a novel lens through which the efficacy and potential side effects of antihypertensive drugs can be evaluated.

Chapter 4 MR integrating GWAS and eQTL data to prioritise genes encoding BP-lowering drug targets for cardiovascular, diabetic, and renal diseases

4.1 Introduction

Thousands of genetic variants have been identified through GWAS (105). Large-scale GWAS for CVDs, such as CAD (273), MI (274), and stroke (277), have discovered genetic risk variants and prioritised candidate genes. Nonetheless, understanding the underlying mechanisms of the associations is challenging (346). The genes or functional DNA components responsible for the GWAS association signals through remain largely elusive. Several factors have complicated the mapping resolution of GWAS. Linkage disequilibrium (LD), which refers to the non-random association of alleles of different loci, poses a significant challenge in identifying causal variants (347). With the presence of LD, a significant association detected from a GWAS may result from two distinct scenarios. The first scenario occurs when the SNP, which affects a biological system and subsequently results in the disease outcome, is directly genotyped in the GWAS and discovered to have a statistical association with the outcome. The genotyped SNP in this scenario is referred to as a functional SNP. The second scenario occurs when the significant outcome-associated SNP in the GWAS does not affect the disease outcome but in high LD with the functional SNP. This significant disease-associated SNP is often referred to as tag SNP. Due to these scenarios, a significant association derived from GWAS does not directly specify the causal variant; further research is necessary to map the precise location of the functional SNP. Another challenge of post-GWAS is that more than 90% of GWAS-identified variants reside in non-coding regions of the genome, and many are located at a considerable distance from the nearest identified gene (348). Non-coding regions of the genome encompass elements like promoters and enhancers that regulate gene expression in intricate ways (134). Previous studies demonstrate the enrichment of disease-associated variants and the tag SNPs in strong LD with them in predicted transcriptional regulatory regions (136, 348). This finding suggests that loci, identified from GWAS as affecting the phenotype, could potentially exert their effects by modifying the genetic regulation of one or more target genes. Understanding these regulatory mechanisms is essential yet challenging, necessitating advanced analytical approaches.

Summary-data-based Mendelian randomization (SMR) (249) is one of the valuable tools for post-GWAS analysis, helping to prioritize genes underlying GWAS hits for subsequent

functional studies. SMR has been developed to examine if the effect of a genetic variant on a phenotype is mediated by transcription, by integrating summary data from expression quantitative trait loci (eQTL) studies and GWAS. SMR method uses eQTL, which are genetic variants that have effects on the expression levels of a particular gene, as proxies to evaluate the association between a trait and the expression level of a gene. The principles behind SMR are similar to those of MR. If a genetic variant has an influence on a particular gene-expression level, which is known as an eQTL, this will result in different gene expression levels among individuals with different genotypes of the genetic variant (e.g. AA, Aa, and aa). Then if the expression level of the gene influences a trait, phenotypic differences will be observed among the different genotype groups. In other words, the genetic variant will also have an effect on the trait (249). As eQTLs are stable across the life-course and generally not influenced by confounding from non-genetic factors, eQTLs can be used as proxies for investigating associations between gene expression and outcomes. SMR method integrates HEIDI test to distinguish pleiotropy from linkage. Pleiotropy refers to a scenario where there is a single genetic variant affecting both gene expression and phenotype. Whereas linkage refers to a scenario where multiple genetic variants, which are in linkage disequilibrium, independently affect gene expression and phenotype. Thus, the SMR method and the HEIDI test incorporate the ability to evaluate whether the association of the expression level of a gene and a trait is due to a shared causal variant (249).

SMR has shown their potential in identifying genes whose expression levels are associated with various phenotypes, including CVDs, inflammatory bowel disease, cancers, and educational attainment (250, 253, 349). A previous study applied the SMR method for three traits (BMI, height, and waist-to-hip ratio adjusted by BMI) and two diseases (rheumatoid arthritis and schizophrenia), using large-scale available GWAS data and eQTL data (250). 126 genes have been prioritised, including *TRAF1* and *ANKRD55* for rheumatoid arthritis, and *SNX19* and *NMRAL1* for schizophrenia. About 20% of these genes are novel candidates, which passed the SMR and HEIDI test and are situated more than 1 Mb away from the nearest GWAS hit. About 60% of the prioritised genes are not the closest annotated gene to the top associated GWAS SNP. These findings demonstrate the potential of the SMR approach in identifying new genetic candidates. In another study, the SMR method was applied to prioritize potential druggable targets for multiple sclerosis, leveraging eQTL data, methylation quantitative trait loci (mQTL) data, and multiple sclerosis GWAS data (252). This study has identified 15 druggable target genes in peripheral blood. These genes showed indications of a pathway linking methylation, expression, and multiple sclerosis. Out of these genes, five are targeted by existing medications and three were replicated in a smaller eQTL dataset. The SMR approach has also been applied to prioritize mitochondrial-

related genes for multiple cancer types (253). The research showed the causative link between mitochondrial dysfunction and various types of cancers. Furthermore, potential druggable targets have been identified for cancer prevention, including *FDPS* (for breast cancer), *NSUN4* (for breast cancer and prostate cancer), *VARS2* (for lung cancer).

The SMR method can be employed to examine the effects of drugs on disease phenotypes. eQTLs of a gene that encodes the protein target of a drug can be used as genetic instruments for the drug effect (120, 221). A previous study has used this approach to examine the association between antihypertensive drug target genes and the risk of schizophrenia, bipolar disorder, and major depressive disorder (120). The expression levels of BP-lowering drug target genes in blood were used as proxies for drug exposure, and eQTLs of these genes were used as genetic instruments for the SMR analysis (120). Their study demonstrated that lower expression of the angiotensin-converting enzyme (*ACE*) gene in blood was associated with an increased risk of schizophrenia, suggesting potential adverse effects of lower ACE mRNA on schizophrenia risk (120). Another study has used SMR to investigate the potential effect of lipid-lowering drug targets on non-alcoholic fatty liver disease (NAFLD). Genetically proxied lipoprotein lipase (LPL) enhancement was found to be associated with a decreased risk of NAFLD, indicating that LPL can be a promising candidate drug target in NAFLD. Moreover, there is no association detected between lipid traits (LDL, HDL, and triglycerides) and the risk of NAFLD; therefore, the impact of LPL on NAFLD may be independent of its lipid-lowering effects (254).

Another statistical approach that has been commonly used in post-GWAS analyses is colocalisation, which can be viewed as an extension of fine-mapping to multiple traits (260). Colocalisation has been proposed to examine whether genetic associations with two traits at a locus are explained by a single shared causal variant or distinct variants. The most popular framework for colocalisation methods is Bayesian framework that uses genetic summary statistics for both traits across multiple SNPs in a particular genetic region to assess whether two association signals are colocalised (255). Colocalisation has been integrated with MR as a powerful tool for elucidating the genetic architecture of complex traits and diseases, as well as for the discovery of potential therapeutic targets (260). For example, a phenome-wide association study was conducted to evaluate the effects of circulating levels of 1002 proteins on 225 phenotypes, using two-sample MR and colocalisation (264). The findings revealed that 68.5% of the associations identified through MR were corroborated by colocalisation. Furthermore, by evaluating data from historical drug development programs, they found that protein-phenotype associations, which received validation from both MR and colocalisation, exhibited a higher probability of being successful in drug development trials (264). Colocalisation has been applied to identify potential therapeutic targets for diseases such as

autoimmune diseases (350, 351), Alzheimer’s disease (352), and lung adenocarcinoma (353). In a study investigating the effect of ACE gene expression on Alzheimer's disease (AD) risk, colocalisation provided evidence for a shared causal variant for each pairwise combination: cortical ACE gene expression and AD liability; cortical ACE expression and SBP; SBP and AD liability. In MR analysis, lower cortical ACE expression was associated with lower SBP, and increased risk of AD; whereas, genetically predicted SBP was not associated with AD risk. These findings suggest an association between cerebral ACE gene expression and AD risk, and this association does not appear to be mediated through the effect on BP (352).

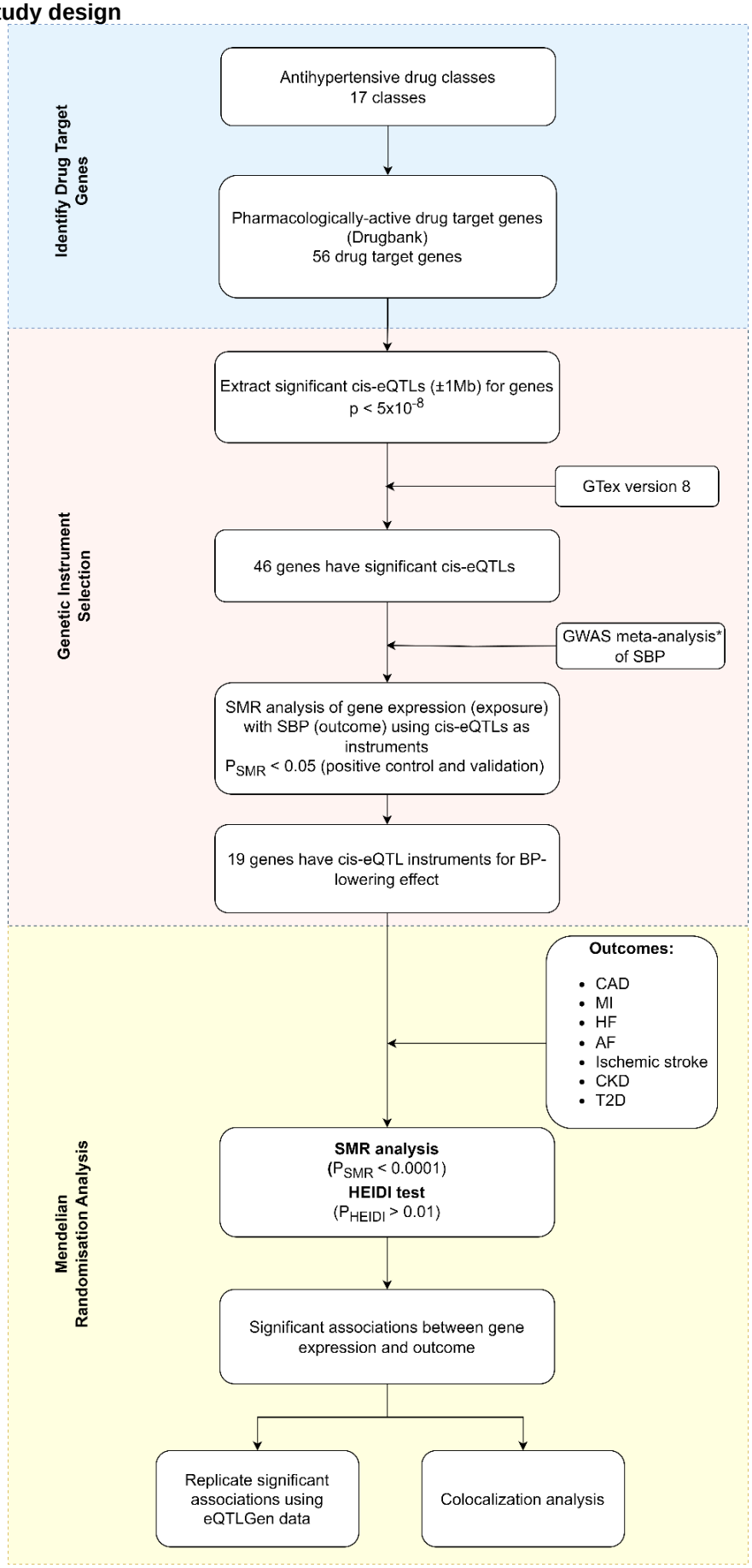
Building on findings from prior chapters where causal associations of BP-lowering drugs with cardiovascular, diabetic, and renal diseases were observed, this chapter aims to explore whether these effects are mediated through changes in the transcription levels of genes targeted by these drugs. I will utilise cis-eQTLs of antihypertensive drug target genes as instruments in my SMR analysis, employing large-scale public databases for robust validation and sensitivity checks using HEIDI and colocalisation analyses. This approach not only deepens my understanding of the genetic mechanisms behind drug effects but also tests the robustness of associations in a biologically informative way.

4.2 Methods

4.2.1 Study design

I performed a SMR approach to examine the association between the expression of genes encoding targets of the BP-lowering drugs and a range of cardiovascular, diabetic and renal outcomes. I used publicly available cis-eQTL data from the GTEx version 8 (285) (49 tissues, n = 70-838) to identify genetic instruments for the expression of the genes encoding BP-lowering drug targets in 29 tissues. The tissues were selected based on their relevance to the pathophysiology of the outcomes under the study. Subsequently, I performed SMR analysis, leveraging the eQTL data and the GWAS summary data of the disease outcomes. eQTL dataset from the eQTLGen consortium (whole blood, n = 31,684) (289) was used to perform the replication analysis for the significant associations between gene expression and the outcomes. The study design is illustrated in **Figure 4.1.**

Figure 4.1 Study design



4.2.2 Identification of genes encoding BP-lowering drug targets

In the chapter 3, following the exhaustive review of the BNF (269), I identified 17 BP-lowering drug classes and the corresponding active ingredients for each drug class. Genes encoding protein targeted by the active ingredients were obtained using the Drugbank database (https://go.drugbank.com/; version 5.1.10); 57 pharmacologically active drug target genes were identified.

4.2.3 Identification of genetic instruments for gene expression

Publicly available data from the GTEx v8 (285) were used to identify SNPs associated with the expression of drug target genes across multiple tissues.

GTEx is an extensive public database that provides insights into human gene expression and regulation as well as its association with genetic variation across multiple diverse tissues and individuals (286). The GTEx project collected samples from up to 54 non-diseased tissue sites, from nearly 1,000 deceased individuals. Whole Genome Sequencing (WGS) technique was used to genotype densely all individuals to evaluate genetic variation present within their genomes. RNA sequencing (bulk RNA-seq) was used to assess gene expression of each tissue (286).

Cis-expression quantitative loci (cis-eQTLs) were identified as genetic variants that had a significant correlation with changes in the expression of genes located nearby. The eQTL analysis of GTEx v8 data was performed on 15,201 RNA-sequencing samples across 49 tissues of 838 post-mortem donors (285). These 49 tissues had at least 70 individuals, each with both RNA sequence and genotype data derived from whole genome sequencing. Out of the 838 donors, 715 (representing 85.3%) were of European American descent, 103 (12.3%) were African American, and 12(1.4%) were Asian American, with 16 (1.95) identified as Hispanic or Latino ethnicity. In terms of gender distribution, 557 (66.4%) donors were male and 281 (33.5%) were female (285). The analysis of the v8 GTEx data provides a catalogue of genetic regulatory variants affecting gene expression and splicing in cis and trans across 49 tissues and describes patterns and mechanisms of tissue- and cell-type specificity of genetic regulatory effects.

Only cis-eQTLs, which are eQTLs situated within 1 Mb window on either side of the encoded gene (354), were selected as genetic instruments for drug target gene expression in each tissue type. Multiple papers have used this window to select genetic instruments for the SMR as well as colocalisation analysis (355-357). I performed SMR analysis on 29 tissues that are relevant to pathophysiology of the outcomes of interest and the mechanism of the

antihypertensive drugs. These tissues are adipose visceral omentum (n = 469), adipose subcutaneous (n = 581), adrenal gland (n = 233), aorta (n = 387), artery coronary (n = 213), artery tibial (n = 584), brain amygdala (n = 129), brain anterior cingulate cortex BA24 (n = 147), brain caudate basal ganglia (n = 194), brain cerebellar hemisphere (n = 175), brain cerebellum (n = 209), brain cortex (n = 205), brain frontal cortex BA9 (n = 175), brain hippocampus (n = 165), brain hypothalamus (n = 170), brain nucleus accumbens (basal ganglia) (n = 194), brain putamen (basal ganglia) (n = 170), brain spinal cord cervical c-1 (n = 126), brain substantia nigra (n = 114), heart atrial appendage (n = 372), heart left ventricle (n = 386), kidney cortex (n = 73), liver (n = 208), muscle skeletal (n = 706), nerve tibial (n = 532), pancreas (n = 305), pituitary (n = 237), thyroid (n = 574), and whole blood (670).

The eQTL data were extracted from the Version 8 release of the GTEx project. The eQTL data in SMR binary format can be downloaded at yanglab.westlake.edu.cn/data/SMR/GTEx_V8_cis_eqtl_summary.html. The effect size of eQTL is defined as the effect size of the effect allele relative to the reference allele in the human genome reference. The effect sizes in GTEx cohort are computed in a normalized space where magnitude has no direct biological interpretation. The eQTL data are on the scale of 1 unit change in the expression level of the gene for each additional effect allele.

For each gene, the top significant cis-eQTL was selected as a genetic instrument for gene expression in a particular tissue (each of the 29 tissues of interest). The p-value threshold to select the significant cis-eQTLs for SMR analysis is 5×10^{-8}. For genetic instrument validation, SMR analysis was conducted to test whether changes in gene expression reflect a BP-reducing association with antihypertensive drugs. Only genes whose expression was associated with SBP ($p < 0.05$) were included for further analysis.

4.2.4 Outcomes

GWAS summary data were obtained for the outcomes from the largest and most comprehensive studies available, focusing on those with substantial European ancestry representation to align with the exposure data. Outcomes were chosen based on their clinical significance and relevance to hypertension and antihypertensive drug effects. The outcomes include atrial fibrillation (60,620 cases and 970,216 controls of European ancestry) (275), heart failure (47,309 cases and 930,014 controls of European descent) (276), coronary artery disease (181,522 cases and 984,168 controls, with more than 95% of European ancestry) (273), myocardial infarction (43,676 cases and 128,199 controls, 77% of European ancestry) (274), ischemic stroke (62,100 cases and 1,234,808 controls of European descent) (277), chronic kidney disease (41,395 cases and 439,303 controls of European ancestry) (303), and type 2 diabetes (80,154 cases and 853,816 controls of

European descent) (279). SBP (n = 757,601, European ancestry (115)) was used as a positive control as well as for validation of genetic instruments. Further details on population characteristics, trait definitions, and GWA analyses were in the cited original publications.

4.2.5 Summary-based Mendelian randomization (SMR) analyses

SMR uses the principle of MR integrating GWAS and eQTL summary statistics to test for association between gene expression and outcome due to shared and potentially causal variants at a locus (249). In this SMR analysis, genetic instruments are cis-eQTLs that satisfied the predefined selection criteria; exposures are gene expression levels in particular tissues (of the 29 tissues of interest); outcomes are the cardiovascular, diabetic and renal diseases. For each of the genetic instruments used in SMR analysis, the cis-eQTL and GWAS summary statistics were combined and harmonised by aligning effect alleles. To maintain consistency in the genetic instruments used across various analyses, proxies were not used. SMR estimates were presented as a beta coefficient with a 95% confidence interval (95% CI) per 1 unit increase in the expression of the genes in a particular tissue.

The HEIDI test is performed to explore if the association between gene expression and the outcome is due to a linkage scenario instead of true causality (249). A linkage scenario occurs when a SNP that affects gene expression is in high linkage disequilibrium with another SNP that affects the outcome. A HEIDI p-value < 0.01 indicates that the observed association is most likely due to the linkage scenario (120).

I adopted the default settings in SMR: $P_{eQTL} < 5 \times 10^{-8}$; MAF > 0.01; eQTLs in very high or low LD ($r^2 > 0.9$ or $r^2 < 0.05$) with the top associated eQTL were excluded in the HEIDI test to avoid the issue of collinearity (358). SMR software version 1.3.1 was used to perform SMR analysis. The p-value threshold of 0.0001 (0.05/498 tests) was applied using Bonferroni correction for multiple testing (19 genes passed the positive control test with SBP; in total there were 498 SMR tests performed between these genes and the seven outcomes across a total of 21 tissues).

Sensitivity analysis was conducted for significant SMR associations using summary statistics obtained from the eQTLGen consortium (whole blood, n = 31,684) (289). The eQTLGen consortium explores the genetic architecture of blood gene expression. It integrates 37 datasets, encompassing 31,684 individuals, with approximately 95% of them being of European ancestry. The cis-eQTL results from eQTLGen data can be downloaded from https://www.eqtlgen.org/cis-eqtls.html.

4.2.6 Colocalisation analysis

For any significant association between the gene expression and the outcome, genetic colocalisation analyses (255) were conducted to estimate the posterior probability for a common causal variant between the two traits, within a 1MB window of the genes. Coloc v5.2.1 R package was used to perform colocalisation analysis. Coloc conducted a Bayesian test for colocalisation in a genomic locus between two traits, assuming at most one causal variant per trait in the genomic region. Coloc calculates five posterior probabilities for the competing models:

H_0: No variant associated with either trait in the region

H_1: A variant associated with trait 1, but not trait 2 in the region

H_2: A variant associated with trait 2, but not trait 1 in the region

H_3: Distinct causal variants for trait 1 and trait 2 in the region

H_4: A shared causal variant for both traits in the region

Default priors were used for the analysis: $p_1 = 10^{-4}$, $p_2 = 10^{-4}$, $p_{12} = 10^{-5}$ (p_1 is the prior probability for a variant being associated with trait 1, p_2 is the prior probability for a variant being associated with trait 2, p_{12} is the prior probability a variant being associated with both traits).

4.3 Results

4.3.1 Genetic instrument selection

Blood Pressure

Significant associations were found between the expression of 19 genes across 21 tissues and SBP. Colocalisation analysis supported a shared causal variant for key genes like *CACNB3* and *ACE* in multiple tissues, suggesting their role in BP regulation.

56 genes that encode the protein targets of BP-lowering drugs were identified. 55 of the 56 genes have eQTL data from the cis-eQTL analysis of the GTEx v8 project. 46 of the 55 genes have significant cis-eQTLs at the p-value threshold of 5×10^{-8} in at least one of the 29 tissues of interest. SMR analysis was performed to examine the association of the 46 genes with SBP. I observed 77 associations ($P_{SMR} < 0.05$) of the expression of 19 genes with SBP in 21 tissues (**Figure 4.2, Table 4.1**). These 19 genes were taken forward to the main SMR analyses, which investigate the potential causal association between gene expression and the outcomes of interest (**Table 4.1**).

Figure 4.2 SMR associations of antihypertensive target gene-expression levels with SBP.
The x-axis represents the β estimates from SMR analysis, and the y-axis represents the -log10 (p_{SMR}). The circle, which signifies the corresponding gene, was coloured according to the drug class that targets its protein. β_{SMR} were estimates from SMR analysis per 1 unit increase in the expression of the genes in a particular tissue. $\beta_{SMR} < 0$ suggests that an increase in gene expression levels in a particular tissue was associated with a decrease in SBP, and vice versa. The blue dashed lines indicate a p-value of 0.05. The p-values were truncated at 10^{-12} to ensure that genes with larger p-values were adequately represented; The identifiers for each SMR association circle are the relevant gene name along with the following abbreviations for tissue types: AVO, Adipose Visceral Omentum; ASc, Adipose Subcutaneous; AdrG, Adrenal Gland; AAo, Artery aorta; Atib, Artery tibial, BCaud, Brain Caudate basal ganglia; BCH, Brain Cerebellar Hemisphere; BC, Brain Cerebellum; BCor, Brain Cortex; BFCB9, Brain Frontal Cortex BA9; BNacc, Brain Nucleus accumbens basal ganglia; BPut, Brain Putamen basal ganglia; HAA, Heart Atrial Appendage; HLV, Heart Left Ventricle, KC, Kidney Cortex, MSk, Muscle Skeletal, NTib, Nerve Tibial; Panc, Pancreas; Thy, Thyroid; WB, Whole Blood.

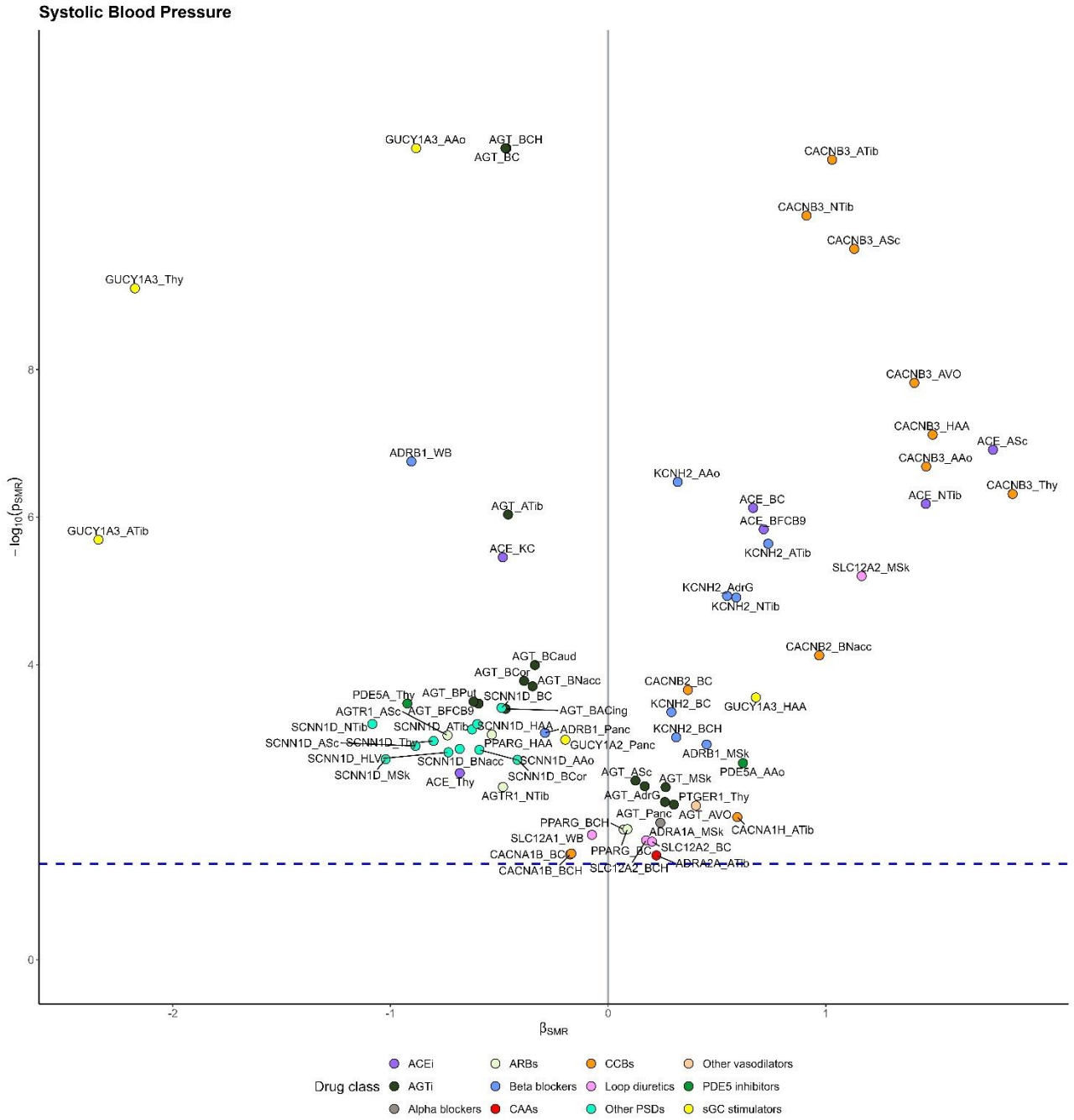


Table 4.1 SMR associations of antihypertensive target gene-expression levels with SBP.

b_{SMR}, SMR effect size per 1 SD increase in gene expression; se_{SMR}, standard error; A1, effect allele; A2, non-effect allele, Freq, frequency of A1.

Drug class	Gene	Outcome	Tissue	Top SNP	Chr	Pos_build37	A1	A2	Freq	b_{SMR}	se_{SMR}	P_{SMR}
ACEi	*ACE*	SBP	Adipose Subcutaneous	rs4308	17	61559625	G	A	0.633	1.768	0.334	1.22E-07
ACEi	*ACE*	SBP	Brain Cerebellum	rs4308	17	61559625	G	A	0.633	0.666	0.135	7.50E-07
ACEi	*ACE*	SBP	Brain Frontal Cortex BA9	rs4291	17	61554194	A	T	0.628	0.715	0.148	1.46E-06
ACEi	*ACE*	SBP	Kidney Cortex	rs4308	17	61559625	G	A	0.633	-0.484	0.104	3.50E-06
ACEi	*ACE*	SBP	Nerve Tibial	rs4292	17	61554341	T	C	0.632	1.459	0.293	6.61E-07
ACEi	*ACE*	SBP	Thyroid	rs12449782	17	61576249	A	G	0.419	-0.681	0.229	0.003
Alpha blockers	*ADRA1A*	SBP	Muscle Skeletal	rs17055703	8	26551416	C	A	0.672	0.24	0.098	0.014
Alpha blockers	*ADRA2A*	SBP	Artery tibial	rs1343449	10	112860526	G	A	0.303	0.222	0.107	0.038
CAA	*ADRA2A*	SBP	Artery tibial	rs1343449	10	112860526	G	A	0.303	0.222	0.107	0.038
β-blocker	*ADRB1*	SBP	Muscle Skeletal	rs7921036	10	115774142	T	C	0.407	0.452	0.14	0.001
β-blocker	*ADRB1*	SBP	Pancreas	rs2532827	10	115774519	C	G	0.408	-0.29	0.087	0.001
β-blocker	*ADRB1*	SBP	Whole Blood	rs4918889	10	115830718	G	C	0.166	-0.903	0.173	1.76E-07
AGTi	*AGT*	SBP	Adipose Subcutaneous	rs5050	1	230849886	G	T	0.178	0.126	0.043	0.004
AGTi	*AGT*	SBP	Adipose Visceral Omentum	rs5050	1	230849886	G	T	0.178	0.302	0.114	0.008
AGTi	*AGT*	SBP	Adrenal Gland	rs5050	1	230849886	G	T	0.178	0.168	0.059	0.004
AGTi	*AGT*	SBP	Artery tibial	rs1977413	1	230852533	A	G	0.129	-0.46	0.094	9.22E-07
AGTi	*AGT*	SBP	Brain Anterior cingulate cortex BA24	rs10864768	1	230825447	C	T	0.211	-0.471	0.133	0.0004
AGTi	*AGT*	SBP	Brain Caudate basal ganglia	rs3789649	1	230816460	T	C	0.211	-0.336	0.087	0.0001
AGTi	*AGT*	SBP	Brain Cerebellar Hemisphere	rs5051	1	230849872	T	C	0.413	-0.467	0.056	1.15E-16
AGTi	*AGT*	SBP	Brain Cerebellum	rs2493134	1	230849359	C	T	0.411	-0.472	0.054	4.47E-18
AGTi	*AGT*	SBP	Brain Cortex	rs10864769	1	230833927	G	A	0.209	-0.386	0.102	0.0002
AGTi	*AGT*	SBP	Brain Frontal Cortex BA9	rs3789649	1	230816460	T	C	0.211	-0.595	0.166	0.0003

AGTi	*AGT*	SBP	Brain Nucleus accumbens basal ganglia	rs10864768	1	230825447	C	T	0.211	-0.347	0.093	0.0002
AGTi	*AGT*	SBP	Brain Putamen basal ganglia	rs3789649	1	230816460	T	C	0.211	-0.618	0.172	0.0003
AGTi	*AGT*	SBP	Muscle Skeletal	rs5050	1	230849886	G	T	0.178	0.264	0.093	0.005
AGTi	*AGT*	SBP	Pancreas	rs5050	1	230849886	G	T	0.178	0.261	0.097	0.007
ARB	*AGTR1*	SBP	Adipose Subcutaneous	rs275677	3	148369183	A	G	0.196	-0.738	0.222	0.001
ARB	*AGTR1*	SBP	Nerve Tibial	rs275665	3	148390863	A	G	0.159	-0.483	0.17	0.005
CCB	*CACNA1B*	SBP	Brain Cerebellar Hemisphere	rs11137372	9	141004945	A	G	0.445	-0.175	0.084	0.037
CCB	*CACNA1B*	SBP	Brain Cerebellum	rs3750508	9	141017240	T	C	0.353	-0.169	0.081	0.036
CCB	*CACNA1H*	SBP	Artery tibial	rs35198836	16	1244631	T	C	0.257	0.593	0.235	0.012
CCB	*CACNB2*	SBP	Brain Cerebellum	rs10828906	10	18841795	T	C	0.75	0.367	0.099	0.0002
CCB	*CACNB2*	SBP	Brain Nucleus accumbens basal ganglia	rs10764318	10	18415963	A	G	0.122	0.97	0.245	7.48E-05
CCB	*CACNB3*	SBP	Adipose Subcutaneous	rs150857355	12	49209340	C	G	0.018	1.131	0.178	2.32E-10
CCB	*CACNB3*	SBP	Adipose Visceral Omentum	rs150857355	12	49209340	C	G	0.018	1.407	0.249	1.51E-08
CCB	*CACNB3*	SBP	Artery aorta	rs150857355	12	49209340	C	G	0.018	1.461	0.281	2.06E-07
CCB	*CACNB3*	SBP	Artery tibial	rs150857355	12	49209340	C	G	0.018	1.029	0.152	1.43E-11
CCB	*CACNB3*	SBP	Heart Atrial Appendage	rs150857355	12	49209340	C	G	0.018	1.491	0.277	7.63E-08
CCB	*CACNB3*	SBP	Nerve Tibial	rs150857355	12	49209340	C	G	0.018	0.911	0.14	8.20E-11
CCB	*CACNB3*	SBP	Thyroid	rs150857355	12	49209340	C	G	0.018	1.858	0.369	4.85E-07
sGC stimulator	*GUCY1A2*	SBP	Pancreas	rs10895931	11	106064055	A	G	0.776	-0.196	0.06	0.001
sGC stimulator	*GUCY1A3*	SBP	Artery aorta	rs17033041	4	156391307	G	A	0.186	-0.882	0.128	6.27E-12
sGC stimulator	*GUCY1A3*	SBP	Artery tibial	rs72687508	4	156628921	G	A	0.234	-2.341	0.493	2.02E-06
sGC stimulator	*GUCY1A3*	SBP	Heart Atrial Appendage	rs72679986	4	156469749	A	T	0.158	0.679	0.187	0.0003
sGC stimulator	*GUCY1A3*	SBP	Thyroid	rs4690974	4	156393641	C	T	0.508	-2.174	0.354	7.93E-10

β-blocker	*KCNH2*	SBP	Adrenal Gland	rs740955	7	150598479	C	T	0.452	0.547	0.125	1.17E-05
β-blocker	*KCNH2*	SBP	Artery aorta	rs887586	7	150652286	G	C	0.605	0.32	0.063	3.34E-07
β-blocker	*KCNH2*	SBP	Artery tibial	rs882156	7	150653336	C	T	0.605	0.735	0.156	2.29E-06
β-blocker	*KCNH2*	SBP	Brain Cerebellar Hemisphere	rs11771808	7	150675868	G	A	0.577	0.313	0.095	0.001
β-blocker	*KCNH2*	SBP	Brain Cerebellum	rs11771808	7	150675868	G	A	0.577	0.291	0.083	0.0004
β-blocker	*KCNH2*	SBP	Nerve Tibial	rs3807375	7	150667210	T	C	0.345	0.588	0.135	1.23E-05
PDE5i	*PDE5A*	SBP	Artery aorta	rs1393966	4	120798857	C	T	0.538	0.619	0.202	0.002
PDE5i	*PDE5A*	SBP	Thyroid	rs1987179	4	120519934	C	T	0.787	-0.923	0.257	0.0003
ARB	*PPARG*	SBP	Brain Cerebellar Hemisphere	rs73025230	3	12379911	C	A	0.096	0.072	0.03	0.017
ARB	*PPARG*	SBP	Brain Cerebellum	rs73025230	3	12379911	C	A	0.096	0.089	0.037	0.017
ARB	*PPARG*	SBP	Heart Atrial Appendage	rs2920502	3	12329195	C	G	0.316	-0.534	0.161	0.001
Vasodilator	*PTGER1*	SBP	Thyroid	rs3760702	19	14588237	A	G	0.317	0.403	0.152	0.008
PSD	*SCNN1D*	SBP	Adipose Subcutaneous	rs146421901	1	1218334	A	G	0.025	-0.885	0.275	0.001
PSD	*SCNN1D*	SBP	Artery aorta	rs12068885	1	1207215	G	A	0.032	-0.591	0.185	0.001
PSD	*SCNN1D*	SBP	Artery tibial	rs11260570	1	1201640	C	T	0.032	-0.625	0.185	0.001
PSD	*SCNN1D*	SBP	Brain Cerebellum	rs41307035	1	1190061	G	A	0.048	-0.49	0.138	0.0004
PSD	*SCNN1D*	SBP	Brain Cortex	rs112631554	1	1200192	G	C	0.032	-0.417	0.134	0.002
PSD	*SCNN1D*	SBP	Brain Nucleus accumbens basal ganglia	rs114964868	1	1210468	T	C	0.023	-0.682	0.213	0.001
PSD	*SCNN1D*	SBP	Heart Atrial Appendage	rs139183246	1	1202437	T	A	0.023	-0.601	0.176	0.001
PSD	*SCNN1D*	SBP	Heart Left Ventricle	rs146421901	1	1218334	A	G	0.025	-0.734	0.232	0.002
PSD	*SCNN1D*	SBP	Muscle Skeletal	rs112631554	1	1200192	G	C	0.032	-1.022	0.329	0.002
PSD	*SCNN1D*	SBP	Nerve Tibial	rs7519699	1	1218890	C	G	0.026	-1.083	0.317	0.001
PSD	*SCNN1D*	SBP	Thyroid	rs12068885	1	1207215	G	A	0.032	-0.802	0.245	0.001
TZD	*SLC12A1*	SBP	Whole Blood	rs74012001	15	48605208	T	A	0.17	-0.074	0.032	0.02
Loop diuretic	*SLC12A1*	SBP	Whole Blood	rs74012001	15	48605208	T	A	0.17	-0.074	0.032	0.02
Loop diuretic	*SLC12A2*	SBP	Brain Cerebellar Hemisphere	rs10478787	5	127294105	A	C	0.578	0.175	0.077	0.024

Loop diuretic	*SLC12A2*	SBP	Brain Cerebellum	rs10478787	5	127294105	A	C	0.578	0.201	0.09	0.025
Loop diuretic	*SLC12A2*	SBP	Muscle Skeletal	rs4836367	5	127482592	G	A	0.19	1.165	0.258	6.29E-06

For the significant associations detected between the gene expression and SBP, genetic colocalisation analyses were performed to estimate the posterior probability for a common causal variant between the two traits - SBP and gene expression level in corresponding tissue - within a 1MB window of the genes. The colocalisation analysis gave a high posterior probability for the model with a shared causal variant (H_4 from 0.999 to > 0.80) for the following genes: *CACNB3* (in nerve tibial, adipose subcutaneous, tibial artery, adipose visceral omentum, heart atrial appendage, aorta, and thyroid), *AGT* (in brain cerebellum and brain cerebellar hemisphere), *GUCY1A3* (in thyroid), *ACE* (in brain frontal cortex BA9, brain cerebellum, kidney cortex, and nerve tibial). The moderately high posterior probability for the shared model (H_4: 0.581 to 0.713) was obtained for *ACE* (in adipose subcutaneous), *GUCY1A3* (in tibial artery), and *AGT* (in brain frontal cortex BA9) (**Table 4.2**)

Table 4.2 Colocalisation analysis for SBP and gene expression levels

nSNP	PP.H0.abf	PP.H1.abf	PP.H2.abf	PP.H3.abf	PP.H4.abf	Trait 1	Trait 2
4394	8.21E-18	1.43E-13	6.63E-08	0.00016	0.99984	CACNB3_Nerve_Tibial	SBP
4408	3.26E-17	1.89E-13	2.63E-07	0.00053	0.99947	CACNB3_Adipose_Subcutaneous	SBP
4396	2.15E-19	3.35E-13	1.74E-09	0.00171	0.99829	CACNB3_Artery_Tibial	SBP
5343	5.56E-51	1.14E-29	1.32E-24	0.00172	0.99828	AGT_Brain_Cerebellum	SBP
5349	8.28E-48	2.24E-29	1.96E-21	0.00433	0.99567	AGT_Brain_Cerebellar_Hemisphere	SBP
6162	1.04E-34	5.93E-29	1.75E-08	0.00895	0.99105	GUCY1A3_Thyroid	SBP
4379	1.25E-13	1.24E-12	0.00101	0.0091	0.98989	CACNB3_Adipose_Visceral_Omentum	SBP
4372	5.22E-13	9.28E-13	0.00422	0.00651	0.98926	CACNB3_Heart_Atrial_Appendage	SBP
2814	2.08E-17	2.06E-16	0.00277	0.02653	0.97070	ACE_Brain_Frontal_Cortex_BA9	SBP
4390	3.04E-12	1.63E-12	0.02461	0.01226	0.96313	CACNB3_Artery_Aorta	SBP
2794	4.36E-18	4.37E-16	0.00058	0.05733	0.94209	ACE_Brain_Cerebellum	SBP
2785	2.33E-16	3.07E-16	0.03106	0.04004	0.92890	ACE_Kidney_Cortex	SBP
4401	8.14E-13	1.62E-11	0.00658	0.13047	0.86295	CACNB3_Thyroid	SBP
2839	2.20E-18	1.16E-15	0.00029	0.15394	0.84576	ACE_Nerve_Tibial	SBP

2836	1.58E-19	2.16E-15	2.00E-05	0.28689	0.71309	ACE_Adipose_Subcutaneous	SBP
6162	1.04E-29	2.31E-27	0.00175	0.38655	0.61170	GUCY1A3_Artery_Tibial	SBP
5325	4.07E-29	1.72E-27	0.00966	0.40886	0.58148	AGT_Brain_Frontal_Cortex_BA9	SBP
6158	1.23E-28	3.87E-27	0.02062	0.64803	0.33135	GUCY1A3_Heart_Atrial_Appendage	SBP
5328	6.35E-29	2.88E-27	0.01507	0.68362	0.30131	AGT_Brain_Putamen_basal_ganglia	SBP
5366	4.25E-29	3.48E-27	0.01011	0.82756	0.16233	AGT_Pancreas	SBP
4506	0.01593	0.80897	0.00147	0.07442	0.09922	PTGER1_Thyroid	SBP
1570	6.99E-21	1.58E-21	0.75885	0.17196	0.06919	SCNN1D_Brain_Nucleus_accumbens_basal_ganglia	SBP
5532	1.99E-06	0.0371	5.00E-05	0.89843	0.06442	AGTR1_Adipose_Subcutaneous	SBP
5528	4.19E-09	0.03782	1.01E-07	0.91578	0.04641	AGTR1_Nerve_Tibial	SBP
1568	6.41E-21	2.44E-21	0.69579	0.26471	0.03951	SCNN1D_Brain_Cerebellum	SBP
1560	3.18E-21	5.82E-21	0.34507	0.63196	0.02298	SCNN1D_Brain_Cortex	SBP
1565	3.31E-21	5.73E-21	0.35934	0.62233	0.01833	SCNN1D_Heart_Left_Ventricle	SBP
6542	4.95E-72	1.90E-71	0.20427	0.78357	0.01216	CACNB2_Brain_Nucleus_accumbens_basal_ganglia	SBP
6551	6.16E-73	2.33E-71	0.02541	0.96371	0.01087	CACNB2_Brain_Cerebellum	SBP
5397	7.31E-43	4.17E-27	1.73E-16	0.99065	0.00935	AGT_Muscle_Skeletal	SBP
1566	7.60E-22	8.39E-21	0.08249	0.91085	0.00666	SCNN1D_Heart_Atrial_Appendage	SBP
5088	6.40E-09	0.36929	1.08E-08	0.6247	0.00601	ADRA2A_Artery_Tibial	SBP
5371	4.49E-29	4.15E-27	0.01068	0.98616	0.00316	AGT_Brain_Anterior_cingulate_cortex_BA24	SBP
5653	2.87E-08	6.12E-05	0.00047	0.9965	0.00298	PDE5A_Thyroid	SBP
5364	2.05E-21	8.95E-20	0.0224	0.97568	0.00193	PPARG_Heart_Atrial_Appendage	SBP
5629	2.41E-10	6.13E-05	3.92E-06	0.99811	0.00183	PDE5A_Artery_Aorta	SBP
5648	4.33E-11	1.58E-09	0.02659	0.97165	0.00176	CACNA1H_Artery_Tibial	SBP
1570	1.93E-22	9.00E-21	0.02105	0.97729	0.00166	SCNN1D_Artery_Aorta	SBP
2838	9.04E-17	7.41E-15	0.01203	0.98656	0.00141	ACE_Thyroid	SBP
4888	5.75E-26	5.10E-24	0.01113	0.98752	0.00135	KCNH2_Adrenal_Gland	SBP
5384	6.70E-29	4.14E-27	0.01591	0.9828	0.00129	AGT_Adipose_Visceral_Omentum	SBP
1573	1.20E-22	9.08E-21	0.01308	0.9862	0.00073	SCNN1D_Adipose_Subcutaneous	SBP
5339	1.28E-35	4.21E-27	3.05E-09	0.99945	0.00055	AGT_Brain_Cortex	SBP

4772	5.85E-37	2.79E-34	0.00209	0.99746	0.00045	ADRB1_Muscle_Skeletal	SBP
1573	1.89E-23	9.19E-21	0.00205	0.9976	0.00035	SCNN1D_Muscle_Skeletal	SBP
4748	3.80E-37	2.80E-34	0.00136	0.99833	0.00032	ADRB1_Pancreas	SBP
5331	1.05E-31	4.21E-27	2.49E-05	0.9998	0.00018	AGT_Brain_Nucleus_accumbens_basal_ganglia	SBP
4875	2.83E-27	5.16E-24	0.00055	0.99936	0.00010	KCNH2_Brain_Cerebellar_Hemisphere	SBP
1569	1.97E-23	9.19E-21	0.00215	0.99776	0.00009	SCNN1D_Nerve_Tibial	SBP
1356	1.32E-12	0.0076	1.72E-10	0.99231	0.00009	CACNA1B_Brain_Cerebellar_Hemisphere	SBP
1356	6.27E-12	0.0076	8.18E-10	0.99235	0.00005	CACNA1B_Brain_Cerebellum	SBP
4773	3.23E-38	2.80E-34	0.00012	0.99986	0.00003	ADRB1_Whole_Blood	SBP
5133	1.36E-26	9.62E-23	0.00014	0.99984	0.00002	SLC12A2_Brain_Cerebellum	SBP
1572	2.56E-24	9.21E-21	0.00028	0.99971	0.00001	SCNN1D_Thyroid	SBP
6141	3.70E-37	5.97E-27	6.20E-11	1	2.43E-06	GUCY1A3_Artery_Aorta	SBP
5229	4.27E-27	9.62E-23	0.00004	0.99995	2.03E-06	SLC12A2_Muscle_Skeletal	SBP
4945	1.93E-27	9.62E-23	0.00002	0.99998	1.41E-06	SLC12A2_Brain_Cerebellar_Hemisphere	SBP
5353	5.75E-35	4.21E-27	1.36E-08	1	9.37E-07	AGT_Brain_Caudate_basal_ganglia	SBP
7096	2.50E-33	1.76E-27	1.41E-06	1	6.55E-08	ADRA1A_Muscle_Skeletal	SBP
4901	5.31E-31	5.16E-24	1.02E-07	1	1.24E-08	KCNH2_Artery_Tibial	SBP
4902	1.36E-32	5.16E-24	2.65E-09	1	3.09E-09	KCNH2_Nerve_Tibial	SBP
1573	2.11E-28	9.21E-21	2.29E-08	1	1.61E-09	SCNN1D_Artery_Tibial	SBP
4865	4.38E-34	5.16E-24	8.48E-11	1	4.64E-11	KCNH2_Brain_Cerebellum	SBP
5374	1.73E-40	4.21E-27	4.12E-14	1	1.16E-11	AGT_Adrenal_Gland	SBP
5335	1.39E-29	9.17E-20	1.52E-10	1	1.13E-11	PPARG_Brain_Cerebellar_Hemisphere	SBP
3808	2.59E-116	3.83E-10	6.78E-107	1	8.12E-12	SLC12A1_Whole_Blood	SBP
5407	2.68E-40	4.21E-27	6.38E-14	1	5.40E-14	AGT_Artery_Tibial	SBP
4901	1.02E-36	5.16E-24	1.974E-13	1	2.89E-14	KCNH2_Artery_Aorta	SBP
5305	5.59E-33	9.17E-20	6.09E-14	1	3.75E-15	PPARG_Brain_Cerebellum	SBP
5458	1.50E-39	1.38E-23	1.08E-16	1	8.28E-18	GUCY1A2_Pancreas	SBP
5394	3.52E-72	4.21E-27	8.35E-46	1	4.74E-28	AGT_Adipose_Subcutaneous	SBP

4.3.2 Coronary artery disease

Gene expression levels of *GUCY1A3* and *PDE5A* showed significant associations with CAD risk. Increased expression level of *GUCY1A3* in the tibial artery was associated with both a decrease in SBP and a decrease in CAD risk. Whereas, increased expression level of *PDE5A* in aorta was associated with both an increase in SBP and an increase in CAD risk.

I observed 26 associations of the expression of 7 genes with the risk of CAD in 13 tissues, of which 3 associations reached significance after adjusting for multiple testing. 18 of the 26 associations passed the HEIDI test (**Figure 4.3, Table 4.3)**.

Using a significant threshold of p-value < 0.0001 (p-value after Bonferroni correction for multiple testing) and a $p_{HEIDI} \geq 0.01$ to eliminate the possibility of association due to linkage, I found significant associations of expression levels of *GUCY1A3* (in tibial artery) and *PDE5A* (in aorta) with CAD risk.

1 unit increase in the expression of *GUCY1A3* in tibial artery was associated with a decrease in SBP (β_{SMR} [95% CI], -2.341 [(-3.307) – (-1.374)], $p_{SMR} = 2.02\times10^{-06}$, $p_{HEIDI} = 0.664$) and a decrease in CAD risk (β_{SMR} [95% CI], -0.404 [(-0.569) – (-0.239)], $p_{SMR} = 1.74\times10^{-06}$, $p_{HEIDI} = 0.345$). I also detected associations for *GUCY1A3* expression in thyroid (β_{SMR}, -0.173 [(-0.247) – (-0.099)], $p_{SMR} = 6.13\times10^{-06}$), aorta ($\beta_{SMR}$, -0.056 [(-0.087) – (-0.025)], $p_{SMR} = 0.0003$), and heart atrial appendage (β_{SMR}, 0.06 [0.009 – 0.111], $p_{SMR} = 0.02$), although these did not pass the HEIDI test (**Figure 4.3, Table 4.3)**.

1 unit increase in the expression of *PDE5A* in aorta was significantly associated with an increase in SBP (β_{SMR} [95% CI], 0.619 [0.223 – 1.015], $p_{SMR} = 0.002$, $p_{HEIDI} = 0.1$) and an increase in CAD risk (β_{SMR} [95% CI], 0.171 [0.095 – 0.247], $p_{SMR} = 9.12\times10^{-06}$, $p_{HEIDI} = 0.338$) (**Figure 4.3, Table 4.3)**.

Moreover, I found suggestive evidence for *KCNH2* expression in tibial artery (β_{SMR}, 0.078 [0.035 – 0.121], $p_{SMR} = 0.0005$; $p_{HEIDI} = 0.02$), and adrenal gland (β_{SMR}, 0.056 [0.023 – 0.089], $p_{SMR} = 0.0008$; $p_{HEIDI} = 0.03$). I also observed the nominal associations (p_{SMR} 0.013 – 0.037; p_{HEIDI} 0.02 – 0.98) for *AGT* expression in brain cerebellar hemisphere; *KCNH2* in nerve tibial; *CACNB3* expression in tibial artery, aorta, adipose visceral omentum, adipose subcutaneous, nerve tibial, heart atrial appendage, and thyroid; *ADRB1* expression in whole blood; *ACE* expression in nerve tibial, adipose subcutaneous, brain cerebellum, and brain frontal cortex BA9. (**Figure 4.3, Table 4.3)**.

For *GUCY1A3* expression in tibial artery and CAD risk, colocalisation analysis gave a high posterior probability for the model with a shared causal variant: $H_4 = 0.99$. For *PDE5A* expression in aorta and CAD risk, the posterior probability for the model with a shared

causal variant is 0.478, and the posterior probability for the distinct causal variants model is 0.521 (**Table 4.4, Figure 4.4).**

Figure 4.3 SMR associations of antihypertensive target gene-expression levels with CAD.
The x-axis represents the β estimates from SMR analysis, and the y-axis represents the -log10(p_{SMR}). The circle, which signifies the corresponding gene, was coloured according to the drug class that targets its protein. β_{SMR} were estimates from SMR analysis per 1 unit increase in the expression of the genes in a particular tissue. $\beta_{SMR} < 0$ suggests that an increase in gene expression levels in a particular tissue was associated with a decrease in CAD risk, and vice versa. The blue and orange dashed lines indicate a p-value of 0.05 and a Bonferroni-corrected p-value of 0.0001, respectively. The asterisk indicates HEIDI p-value greater than 0.01; The identifiers for each SMR association circle are the relevant gene name along with the following abbreviations for tissue types: AVO, Adipose Visceral Omentum; ASc, Adipose Subcutaneous; AdrG, Adrenal Gland; AAo, Artery aorta; Atib, Artery tibial, BCaud, Brain Caudate basal ganglia; BCH, Brain Cerebellar Hemisphere; BC, Brain Cerebellum; BCor, Brain Cortex; BFCB9, Brain Frontal Cortex BA9; BNacc, Brain Nucleus accumbens basal ganglia; BPut, Brain Putamen basal ganglia; HAA, Heart Atrial Appendage; HLV, Heart Left Ventricle, KC, Kidney Cortex, MSk, Muscle Skeletal, NTib, Nerve Tibial; Panc, Pancreas; Thy, Thyroid; WB, Whole Blood.

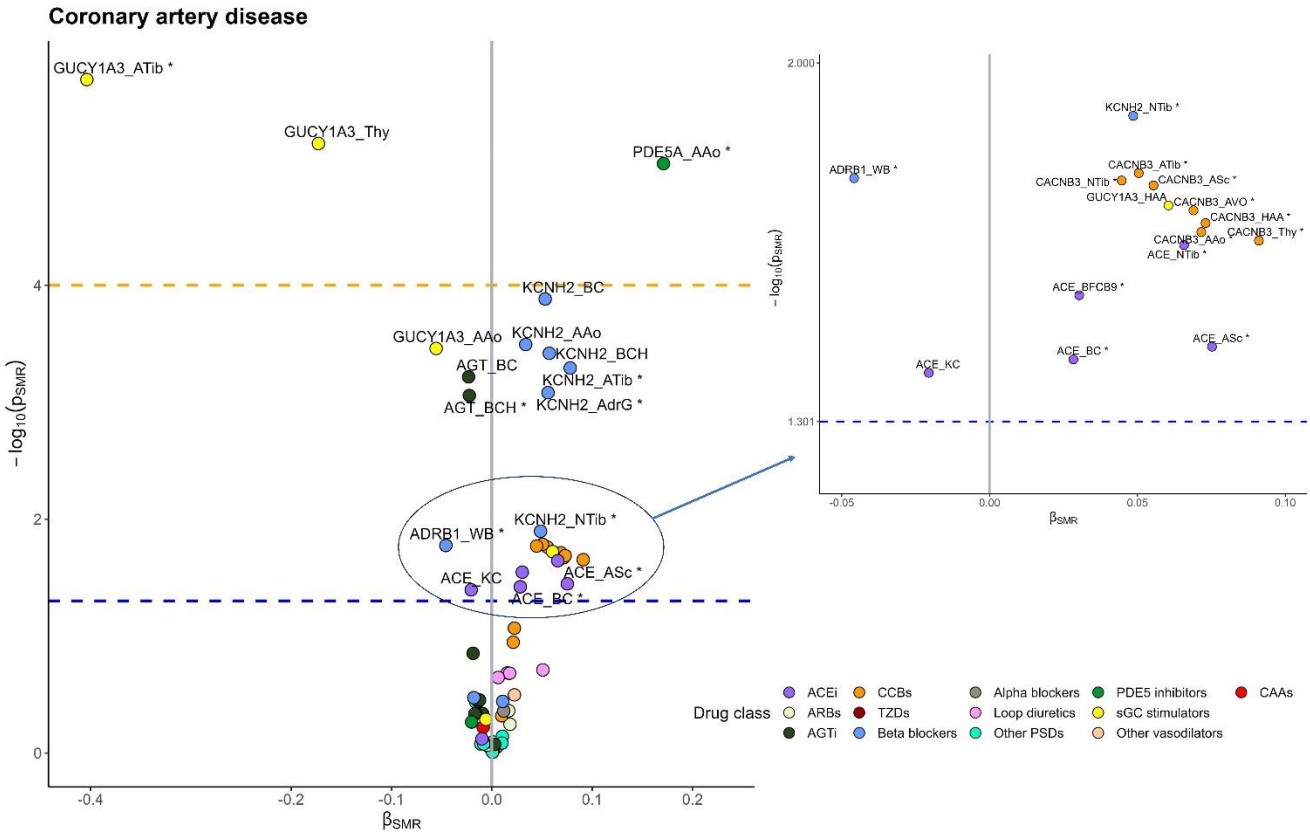


Table 4.3 SMR associations of antihypertensive target gene-expression levels with CAD.

b_{SMR}, SMR effect size per 1 SD increase in gene expression; se_{SMR}, standard error; A1, effect allele; A2, non-effect allele, Freq, frequency of A1.

Drug class	Gene	Outcome	Tissue	TopSNP	Chr	A1	A2	Freq	b_{SMR}	se_{SMR}	P_{SMR}	p_{HEIDI}	$nsnp_{HEIDI}$
ACEi	*ACE*	CAD	ASc	rs4308	17	G	A	0.633	0.075	0.036	0.036	0.267	16
ACEi	*ACE*	SBP	ASc	rs4308	17	G	A	0.633	1.768	0.334	1.22E-07	0.071	16
ACEi	*ACE*	CAD	BC	rs4308	17	G	A	0.633	0.028	0.014	0.038	0.711	12
ACEi	*ACE*	SBP	BC	rs4308	17	G	A	0.633	0.666	0.135	7.50E-07	0.2	11
ACEi	*ACE*	CAD	BFCB9	rs4291	17	A	T	0.628	0.03	0.014	0.028	0.793	4
ACEi	*ACE*	SBP	BFCB9	rs4291	17	A	T	0.628	0.715	0.148	1.46E-06	0.213	4
ACEi	*ACE*	CAD	KC	rs4308	17	G	A	0.633	-0.021	0.01	0.04		
ACEi	*ACE*	SBP	KC	rs4308	17	G	A	0.633	-0.484	0.104	3.50E-06		
ACEi	*ACE*	CAD	NTib	rs4292	17	T	C	0.632	0.066	0.029	0.023	0.555	12
ACEi	*ACE*	SBP	NTib	rs4292	17	T	C	0.632	1.459	0.293	6.61E-07	0.21	12
β-blockers	*ADRB1*	CAD	WB	rs4918889	10	G	C	0.166	-0.046	0.019	0.017	0.019	16
β-blockers	*ADRB1*	SBP	WB	rs4918889	10	G	C	0.166	-0.903	0.173	1.76E-07	6.14E-08	16
AGTi	*AGT*	CAD	BC	rs2493134	1	C	T	0.411	-0.023	0.007	0.001	0.009	20
AGTi	*AGT*	SBP	BC	rs2493134	1	C	T	0.411	-0.472	0.054	4.47E-18	2.10E-06	20
AGTi	*AGT*	CAD	BCH	rs5051	1	T	C	0.413	-0.022	0.007	0.001	0.057	20
AGTi	*AGT*	SBP	BCH	rs5051	1	T	C	0.413	-0.467	0.056	1.15E-16	0.0001	20
CCBs	*CACNB3*	CAD	AAo	rs150857355	12	C	G	0.018	0.072	0.031	0.021	0.987	6
CCBs	*CACNB3*	SBP	AAo	rs150857355	12	C	G	0.018	1.461	0.281	2.06E-07	2.35E-05	5
CCBs	*CACNB3*	CAD	ASc	rs150857355	12	C	G	0.018	0.055	0.023	0.017	0.969	20
CCBs	*CACNB3*	SBP	ASc	rs150857355	12	C	G	0.018	1.131	0.178	2.31E-10	0.006	20
CCBs	*CACNB3*	CAD	ATib	rs150857355	12	C	G	0.018	0.05	0.021	0.016	0.917	20
CCBs	*CACNB3*	SBP	ATib	rs150857355	12	C	G	0.018	1.029	0.152	1.43E-11	0.0001	19
CCBs	*CACNB3*	CAD	AVO	rs150857355	12	C	G	0.018	0.069	0.029	0.019	0.82	9
CCBs	*CACNB3*	SBP	AVO	rs150857355	12	C	G	0.018	1.407	0.249	1.51E-08	8.46E-05	9
CCBs	*CACNB3*	CAD	HAA	rs150857355	12	C	G	0.018	0.073	0.032	0.021	0.945	8

CCBs	*CACNB3*	SBP	HAA	rs150857355	12	C	G	0.018	1.491	0.277	7.63E-08	0.0003	7
CCBs	*CACNB3*	CAD	NTib	rs150857355	12	C	G	0.018	0.045	0.019	0.017	0.978	20
CCBs	*CACNB3*	SBP	NTib	rs150857355	12	C	G	0.018	0.911	0.14	8.20E-11	0.005	20
CCBs	*CACNB3*	CAD	Thy	rs150857355	12	C	G	0.018	0.091	0.04	0.022	0.57	20
CCBs	*CACNB3*	SBP	Thy	rs150857355	12	C	G	0.018	1.858	0.369	4.85E-07	0.001	19
sGC stimulators	*GUCY1A3*	CAD	AAo	rs17033041	4	G	A	0.186	-0.056	0.016	0.0003	0.001	20
sGC stimulators	*GUCY1A3*	SBP	AAo	rs17033041	4	G	A	0.186	-0.882	0.128	6.26E-12	0.003	20
sGC stimulators	*GUCY1A3*	CAD	ATib	rs72687508	4	G	A	0.234	-0.404	0.084	1.74E-06	0.664	20
sGC stimulators	*GUCY1A3*	SBP	ATib	rs72687508	4	G	A	0.234	-2.341	0.493	2.02E-06	0.345	20
sGC stimulators	*GUCY1A3*	CAD	HAA	rs72679986	4	A	T	0.158	0.06	0.026	0.019	2.94E-05	11
sGC stimulators	*GUCY1A3*	SBP	HAA	rs72679986	4	A	T	0.158	0.679	0.187	0.0003	7.87E-07	11
sGC stimulators	*GUCY1A3*	CAD	Thy	rs4690974	4	C	T	0.508	-0.173	0.038	6.13E-06	0.0001	15
sGC stimulators	*GUCY1A3*	SBP	Thy	rs4690974	4	C	T	0.508	-2.174	0.354	7.92E-10	2.16E-07	15
β-blockers	*KCNH2*	CAD	AAo	rs887586	7	G	C	0.605	0.034	0.009	0.0003	0.004	20
β-blockers	*KCNH2*	SBP	AAo	rs887586	7	G	C	0.605	0.32	0.063	3.33E-07	0.022	20
β-blockers	*KCNH2*	CAD	ATib	rs882156	7	C	T	0.605	0.078	0.022	0.001	0.023	20
β-blockers	*KCNH2*	SBP	ATib	rs882156	7	C	T	0.605	0.735	0.156	2.29E-06	0.173	20
β-blockers	*KCNH2*	CAD	AdrG	rs740955	7	C	T	0.452	0.056	0.017	0.001	0.026	20
β-blockers	*KCNH2*	SBP	AdrG	rs740955	7	C	T	0.452	0.547	0.125	1.17E-05	0.106	20
β-blockers	*KCNH2*	CAD	BC	rs11771808	7	G	A	0.577	0.053	0.014	0.0001	1.80E-09	20
β-blockers	*KCNH2*	SBP	BC	rs11771808	7	G	A	0.577	0.291	0.083	0.0004	1.95E-09	20
β-blockers	*KCNH2*	CAD	BCH	rs11771808	7	G	A	0.577	0.057	0.016	0.0004	1.68E-07	20
β-blockers	*KCNH2*	SBP	BCH	rs11771808	7	G	A	0.577	0.313	0.095	0.001	9.59E-08	20
β-blockers	*KCNH2*	CAD	NTib	rs3807375	7	T	C	0.345	0.049	0.019	0.013	0.067	20

β-blockers	*KCNH2*	SBP	NTib	rs3807375	7	T	C	0.345	0.588	0.135	1.22E-05	0.001	20
PDE5 inhibitors	*PDE5A*	CAD	AAo	rs1393966	4	C	T	0.538	0.171	0.039	0.0003	0.338	20
PDE5 inhibitors	*PDE5A*	SBP	AAo	rs1393966	4	C	T	0.538	0.619	0.202	0.002	0.093	20

Table 4.4 Colocalisation analysis of significant gene-disease associations.

nSNP	PP.H0.abf	PP.H1.abf	PP.H2.abf	PP.H3.abf	PP.H4.abf	Trait 1	Trait 2
7986	2.39E-20	5.40E-18	0.0003	0.0604	0.93933	*GUCY1A3* (in artery tibial)	CAD
7241	8.94E-14	2.26E-08	2.07E-06	0.52156	0.47844	*PDE5A* (in artery aorta)	CAD

Table 4.5 SMR results using summary statistics from eQTLGen Consortium.

Gene	topSNP	chr	A_1	A_2	Freq	b_{SMR}	se_{SMR}	p_{SMR}	p_{HEIDI}	$nsnp_{HEIDI}$	outcome	tissue
GUCY1A3	rs72685775	4	C	T	0.013	-0.014	0.047	0.771	0.882	8	CAD	whole blood
GUCY1A3	rs3796585	4	A	G	0.376	-3.129	0.419	8.39E-14	0.09	20	SBP	whole blood
KCNH2	rs2072412	7	G	C	0.294	-0.539	0.212	0.011	1.77E-06	20	SBP	whole blood
KCNH2	rs2072412	7	G	C	0.294	0.249	0.048	2.27E-07	0.109	20	AF	whole blood
PDE5A	rs10518336	4	G	A	0.533	0.191	0.04	1.48E-06	0.005	20	CAD	whole blood
PDE5A	rs10518336	4	G	A	0.533	0.989	0.233	2.26E-05	6.51E-07	20	SBP	whole blood

4.3.3 Myocardial Infarction

8 associations were observed for the expression of 5 genes with the risk of MI in 7 tissues, all of which did not reach significance after multiple testing correction. 5 of the 8 associations passed HEIDI test. Particularly, the association signals were detected for *GUCY1A3* in tibial artery (β_{SMR} [95% CI] per 1 unit increase in gene expression, -0.473 [(-0.714) - (-0.231)]; p_{SMR} = 0.0001, p_{HEIDI} = 0.873) and thyroid (β_{SMR}, -0.2 [(-0.335) - (-0.066)]; p_{SMR} = 0.0035, p_{HEIDI} = 0.03); *PDE5A* in aorta (β_{SMR}, 0.141 [0.016 - 0.266]); p_{SMR} = 0.027, p_{HEIDI} = 0.31); and *SCNN1D* in brain cortex (β_{SMR}, 0.092 [0.006 - 0.178]); p_{SMR} = 0.036, p_{HEIDI} = 0.19) and thyroid (β_{SMR}, 0.163 [0.0007 - 0.326]); p_{SMR} = 0.049, p_{HEIDI} = 0.704) (**Figure 4.4, Table 4.6)**.

Figure 4.4 SMR associations of antihypertensive target gene-expression levels with MI.
The x-axis represents the β estimates from SMR analysis, and the y-axis represents the -log10(p_{SMR}). The circle, which signifies the corresponding gene, was coloured according to the drug class that targets its protein. β_{SMR} were estimates from SMR analysis per 1 unit increase in the expression of the genes in a particular tissue. $\beta_{SMR} < 0$ suggests that an increase in gene expression levels in a particular tissue was associated with a decrease in MI risk, and vice versa. The blue and orange dashed lines indicate a p-value of 0.05 and a Bonferroni-corrected p-value of 0.0001, respectively. The asterisk indicates HEIDI p-value greater than 0.01; Atib, Artery tibial, BCH, Brain Cerebellar Hemisphere; BC, Brain Cerebellum; BCor, Brain Cortex; HAA, Heart Atrial Appendage; Thy, Thyroid.

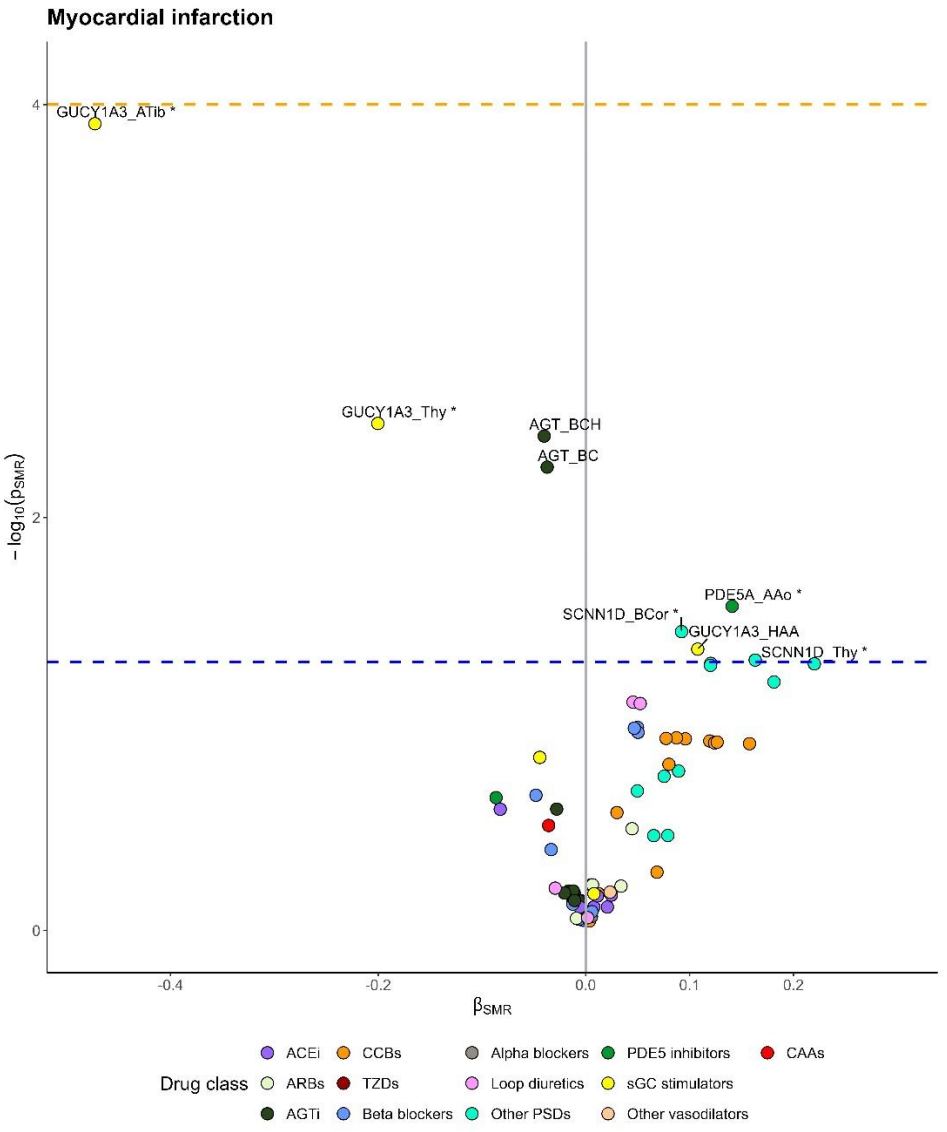


Table 4.6 SMR associations of antihypertensive target gene-expression levels with MI.

b_{SMR}, SMR effect size per 1 SD increase in gene expression; se_{SMR}, standard error; A1, effect allele; A2, non-effect allele, Freq, frequency of A1.

Drug class	Gene	Outcome	Tissue	TopSNP	Chr	A1	A2	Freq	b_{SMR}	se_{SMR}	P_{SMR}	p_{HEIDI}	$nsnp_{HEIDI}$
AGTi	*AGT*	MI	BC	rs2493134	1	C	T	0.411	-0.037	0.013	0.006	0.009	20
AGTi	*AGT*	SBP	BC	rs2493134	1	C	T	0.411	-0.472	0.054	4.47E-18	2.10E-06	20
AGTi	*AGT*	MI	BCH	rs5051	1	T	C	0.413	-0.04	0.014	0.004	0.009	20
AGTi	*AGT*	SBP	BCH	rs5051	1	T	C	0.413	-0.467	0.056	1.15E-16	0.0001	20
sGC stimulators	*GUCY1A3*	MI	ATib	rs72687508	4	G	A	0.234	-0.473	0.123	0.0001	0.873	20
sGC stimulators	*GUCY1A3*	SBP	ATib	rs72687508	4	G	A	0.234	-2.341	0.493	2.02E-06	0.345	20
sGC stimulators	*GUCY1A3*	MI	HAA	rs72679986	4	A	T	0.158	0.108	0.053	0.043	0.001	11
sGC stimulators	*GUCY1A3*	SBP	HAA	rs72679986	4	A	T	0.158	0.679	0.187	0.0003	7.87E-07	11
sGC stimulators	*GUCY1A3*	MI	Thy	rs4690974	4	C	T	0.508	-0.2	0.069	0.004	0.031	15
sGC stimulators	*GUCY1A3*	SBP	Thy	rs4690974	4	C	T	0.508	-2.174	0.354	7.92E-10	2.16E-07	15
PDE5 inhibitors	*PDE5A*	MI	AAo	rs1393966	4	C	T	0.538	0.141	0.064	0.027	0.306	20
PDE5 inhibitors	*PDE5A*	SBP	AAo	rs1393966	4	C	T	0.538	0.619	0.202	0.002	0.093	20
Other PSDs	*SCNN1D*	MI	BCor	rs112631554	1	G	C	0.032	0.092	0.044	0.036	0.19	20
Other PSDs	*SCNN1D*	SBP	BCor	rs112631554	1	G	C	0.032	-0.417	0.134	0.002	0.186	20
Other PSDs	*SCNN1D*	MI	Thy	rs113867217	1	T	C	0.034	0.163	0.083	0.049	0.704	14
Other PSDs	*SCNN1D*	SBP	Thy	rs12068885	1	G	A	0.032	-0.802	0.245	0.001	0.004	13

4.3.4 Atrial Fibrillation

KCNH2 expression in the cerebellum showed a significant association with AF. Increased *KCNH2* expression was associated with a higher SBP and an increase in AF risk. Colocalisation analysis results in distinct causal variants for the gene expression and AF risk.

I observed 8 associations of the expression of 4 genes with the risk of AF in 5 tissues, of which only 1 association reached significance after adjusting for multiple testing. 6 of the 8 associations passed the HEIDI test.

I found significant associations of expression levels of *KCNH2* in the brain cerebellum with AF risk. 1 unit increase in the expression of *KCNH2* in brain cerebellum was associated with an increase in SBP (β_{SMR} [95% CI], 0.290 [0.129 - 0.453], p_{SMR} = 0.0004) and an increase in AF risk (β_{SMR} [95% CI], 0.075 [0.039 - 0.112], $p_{SMR} = 6.02\times10^{-05}$, p_{HEIDI} = 0.195) (**Figure 4.5, Table 4.7)**. Increased *KCNH2* expression in the brain cerebellar hemisphere was also associated with an increase in AF risk (β_{SMR}, 0.081 [0.038 - 0.124], p_{SMR} = 0.0002, p_{HEIDI} = 0.299), although this did not reach significance after adjusting for multiple testing (**Figure 4.5, Table 4.7)**.

Moreover, I also detected the nominal associations (p_{SMR} = 0.005 - 0.037) of the expression of the following genes with AF risk: *AGT* expression in the brain cerebellar hemisphere (-0.02 [(-0.038) - (-0.002)]); *PPARG* expression in the brain cerebellum (-0.024 [(-0.04) - (-0.007)]), brain cerebellar hemisphere (-0.019, [(-0.032) - (-0.006)]), and heart atrial appendage (0.066, [0.004 – 0.127]); and *ADRB1* expression in whole blood (-0.058 [(-0.108) - (-0.008)]) (**Figure 4.5, Table 4.7).**

For *KCNH2* expression in the brain cerebellum and AF risk, colocalisation analysis gave a high posterior probability for the model with distinct causal variants: H_3 = 0.99 (**Table 4.8).** Similarly, colocalisation analysis resulted in a high posterior probability for the model with distinct causal variants (H_3 = 0.99) for *KCNH2* expression in the brain cerebellar hemisphere and AF risk (**Table 4.8).**

Figure 4.5 SMR associations of antihypertensive target gene-expression levels with AF.
The x-axis represents the β estimates from SMR analysis, and the y-axis represents the -log10(p_{SMR}). The circle, which signifies the corresponding gene, was coloured according to the drug class that targets its protein. β_{SMR} were estimates from SMR analysis per 1 unit increase in the expression of the genes in a particular tissue. $\beta_{SMR} < 0$ suggests that an increase in gene expression levels in a particular tissue was associated with a decrease in MI risk, and vice versa. The blue and orange dashed lines indicate a p-value of 0.05 and a Bonferroni-corrected p-value of 0.0001, respectively. The asterisk indicates HEIDI p-value greater than 0.01; The identifiers for each SMR association circle are the relevant gene name along with the following abbreviations for tissue types: BCH, Brain Cerebellar Hemisphere; BC, Brain Cerebellum; HAA, Heart Atrial Appendage; NTib, Nerve Tibial; WB, Whole Blood.

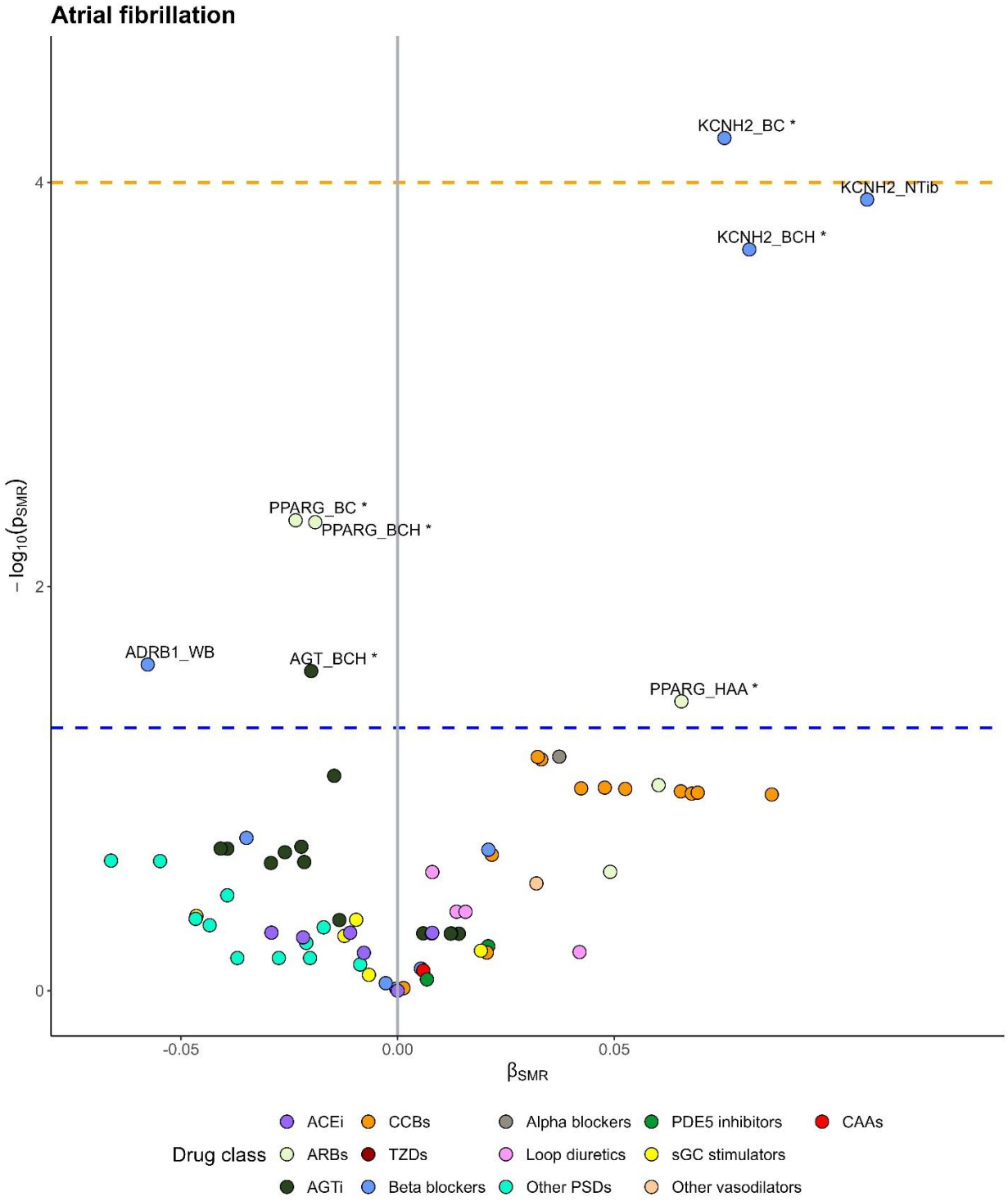


Table 4.7 SMR results of drug target gene expression with the risk of AF.

b_{SMR}, SMR effect size per 1 SD increase in gene expression; se_{SMR}, standard error; A1, effect allele; A2, non-effect allele, Freq, frequency of A1.

Drug class	Gene	Outcome	Tissue	TopSNP	Chr	A1	A2	Freq	b_{SMR}	se_{SMR}	P_{SMR}	p_{HEIDI}	$nsnp_{HEIDI}$
β-blockers	*ADRB1*	AF	WB	rs4918889	10	G	C	0.166	-0.058	0.026	0.024	0.004	16
β-blockers	*ADRB1*	SBP	WB	rs4918889	10	G	C	0.166	-0.903	0.173	1.76E-07	6.14E-08	16
AGTi	*AGT*	AF	BCH	rs5051	1	T	C	0.413	-0.02	0.009	0.026	0.101	20
AGTi	*AGT*	SBP	BCH	rs5051	1	T	C	0.413	-0.467	0.056	1.15E-16	0.0001	20
β-blockers	*KCNH2*	AF	BC	rs11771808	7	G	A	0.577	0.075	0.019	6.01E-05	0.195	20
β-blockers	*KCNH2*	SBP	BC	rs11771808	7	G	A	0.577	0.291	0.083	0.0004	1.95E-09	20
β-blockers	*KCNH2*	AF	BCH	rs11771808	7	G	A	0.577	0.081	0.022	0.0002	0.299	20
β-blockers	*KCNH2*	SBP	BCH	rs11771808	7	G	A	0.577	0.313	0.095	0.001	9.59E-08	20
β-blockers	*KCNH2*	AF	NTib	rs3807375	7	T	C	0.345	0.108	0.028	0.0001	0.004	20
β-blockers	*KCNH2*	SBP	NTib	rs3807375	7	T	C	0.345	0.588	0.135	1.22E-05	0.001	20
ARBs	*PPARG*	AF	BC	rs73025230	3	C	A	0.096	-0.024	0.008	0.005	0.098	20
ARBs	*PPARG*	SBP	BC	rs73025230	3	C	A	0.096	0.089	0.037	0.017	0.01	20
ARBs	*PPARG*	AF	BCH	rs73025230	3	C	A	0.096	-0.019	0.007	0.005	0.15	20
ARBs	*PPARG*	SBP	BCH	rs73025230	3	C	A	0.096	0.072	0.03	0.017	0.006	20
ARBs	*PPARG*	AF	HAA	rs2920502	3	C	G	0.316	0.065	0.031	0.037	0.253	18
ARBs	*PPARG*	SBP	HAA	rs2920502	3	C	G	0.316	-0.534	0.161	0.001	0.083	18

Table 4.8 Colocalisation analysis of significant gene-disease associations.

nsnps	PP.H0.abf	PP.H1.abf	PP.H2.abf	PP.H3.abf	PP.H4.abf	Trait 1	Trait 2
7671	5.04E-16	5.92E-06	8.50E-11	0.9996	0.00039	*KCNH2* (in brain cerebellum)	AF
7793	3.05E-09	5.91E-06	0.00052	0.9981	0.00139	*KCNH2* (in brain cerebellar hemisphere)	AF

4.3.5 Heart Failure

There were 11 nominal associations of the expression of 4 genes with the risk of heart failure in 7 tissues. All of these associations passed the HEIDI test. Suggestive evidence was found for the associations of *KCNH2* expression in nerve tibial (β_{SMR} [95% CI] per 1 unit increase in gene expression, 0.115 [0.051 - 0.179], p_{SMR} = 0.0005, p_{HEIDI} = 0.512), aorta (β_{SMR}, 0.047 [0.018 - 0.075], p_{SMR} = 0.001, p_{HEIDI} = 0.26), and tibial artery (β_{SMR}, 0.106 [0.038 - 0.175], p_{SMR} = 0.002, p_{HEIDI} = 0.23) with the risk of HF. Nominal associations were observed (p_{SMR}: 0.014 - 0.045) for the following genes: *KCNH2* in brain cerebellum (0.051 [0.011 - 0.091]) and brain cerebellar hemisphere (0.055 [0.01 - 0.099]); *GUCY1A3* in aorta (-0.052 [(-0.098) - (-0.006)]); *CACNA1B* in brain cortex (0.043 [0.002 - 0.084]) and brain cerebellar hemisphere (0.054 [0.011 – 0.098]); *ADRB1* in muscle skeletal (-0.062 [(-0.123) - (-0.001)]) and pancreas (0.043 [0.004 - 0.082]) (**Figure 4.6, Table 4.9)**.

4.3.6 Ischemic stroke

There were 8 nominal associations of the expression of 5 genes with ischemic stroke risk in 6 tissues, of which 7 associations passed the HEIDI test (p_{SMR}: 0.004 - 0.022, p_{HEIDI}: 0.03 - 0.986). The association signals were detected for *AGTR1* in nerve tibial (β_{SMR} [95% CI] per 1 unit increase in the gene expression, 0.118 [0.037 - 0.199]); *KCNH2* in brain cerebellum (0.055 [0.017 - 0.093]) and brain cerebellar hemisphere (0.059 [0.016 - 0.102]); *GUCY1A3* in aorta (-0.168 [(-0.298) - (-0.038)]); *AGT* in brain cerebellum (-0.023 [(-0.040) - (-0.005)]); and *ADRB1* in pancreas (0.043 [0.007 - 0.078]) and muscle skeletal (-0.065 [(-0.121) - (-0.009)]) (**Figure 4.7, Table 4.10)**.

Figure 4.6 SMR associations of antihypertensive target gene-expression levels with HF.
The x-axis represents the β estimates from SMR analysis, and the y-axis represents the -log10(p_{SMR}). The circle, which signifies the corresponding gene, was coloured according to the drug class that targets its protein. β_{SMR} were estimates from SMR analysis per 1 unit increase in the expression of the genes in a particular tissue. $\beta_{SMR} < 0$ suggests that an increase in gene expression levels in a particular tissue was associated with a decrease in MI risk, and vice versa. The blue and orange dashed lines indicate a p-value of 0.05 and a Bonferroni-corrected p-value of 0.0001, respectively. The asterisk indicates HEIDI p-value greater than 0.01; The identifiers for each SMR association circle are the relevant gene name along with the following abbreviations for tissue types: AAo, Artery aorta; Atib, Artery tibial; BCH, Brain Cerebellar Hemisphere; BC, Brain Cerebellum; MSk, Muscle Skeletal, NTib, Nerve Tibial.

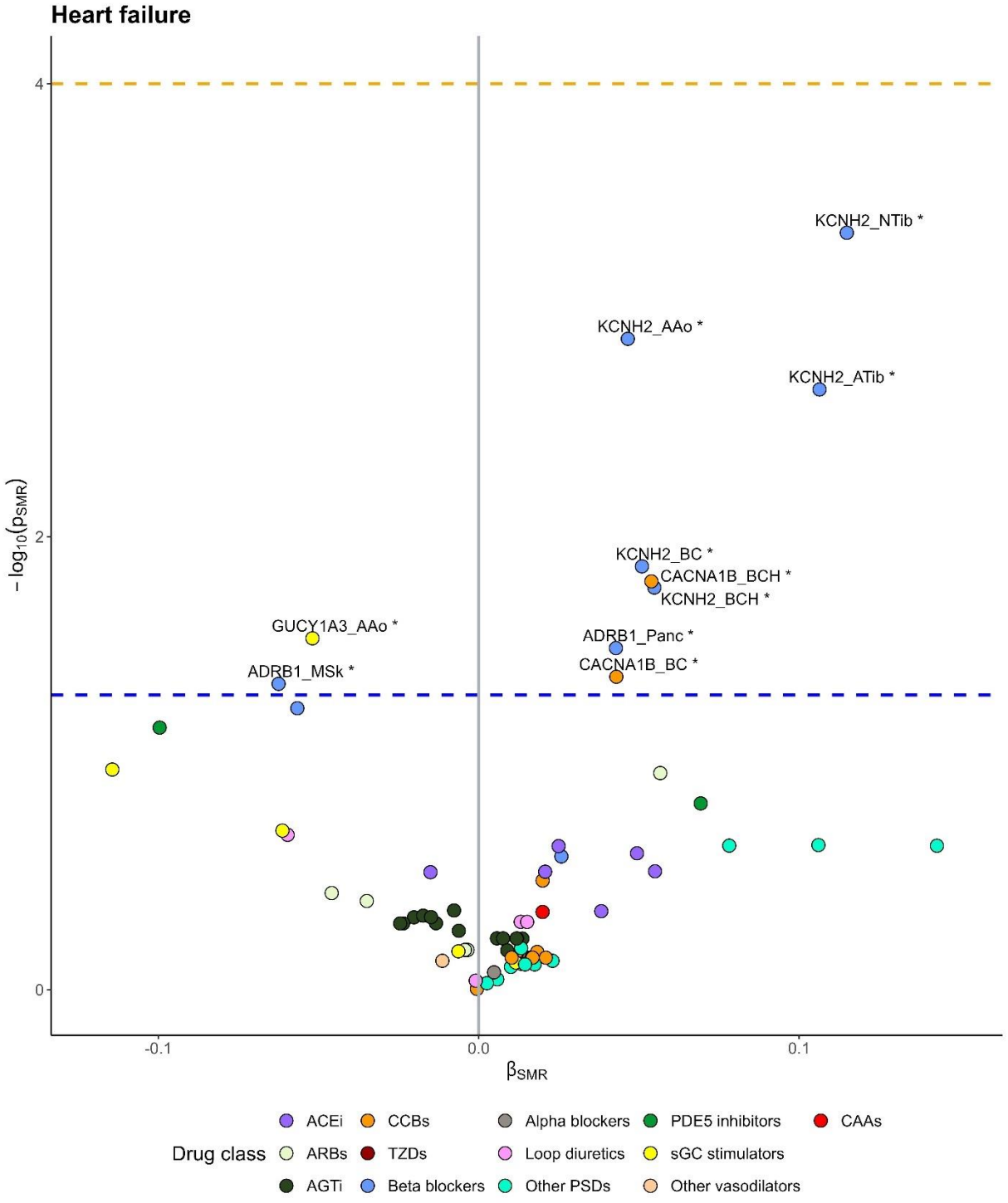


Table 4.9 SMR associations of antihypertensive target gene-expression levels with HF.

b_{SMR}, SMR effect size per 1 SD increase in gene expression; se_{SMR}, standard error; A1, effect allele; A2, non-effect allele, Freq, frequency of A1.

Drug class	Gene	Outcome	Tissue	TopSNP	Chr	A1	A2	Freq	b_{SMR}	se_{SMR}	P_{SMR}	p_{HEIDI}	$nsnp_{HEIDI}$
β-blockers	*ADRB1*	HF	MSk	rs7921036	10	T	C	0.407	-0.062	0.031	0.045	0.162	7
β-blockers	*ADRB1*	SBP	MSk	rs7921036	10	T	C	0.407	0.452	0.14	0.001	0.0002	7
β-blockers	*ADRB1*	HF	Panc	rs2532827	10	C	G	0.408	0.043	0.02	0.031	0.425	13
β-blockers	*ADRB1*	SBP	Panc	rs2532827	10	C	G	0.408	-0.29	0.087	0.001	3.14E-07	13
CCBs	*CACNA1B*	HF	BC	rs3750508	9	T	C	0.353	0.043	0.021	0.042	0.72	20
CCBs	*CACNA1B*	SBP	BC	rs3750508	9	T	C	0.353	-0.169	0.081	0.036	0.817	20
CCBs	*CACNA1B*	HF	BCH	rs11137372	9	A	G	0.445	0.054	0.022	0.016	0.917	20
CCBs	*CACNA1B*	SBP	BCH	rs11137372	9	A	G	0.445	-0.175	0.084	0.037	0.868	20
sGC stimulators	*GUCY1A3*	HF	AAo	rs17033041	4	G	A	0.186	-0.052	0.024	0.028	0.555	20
sGC stimulators	*GUCY1A3*	SBP	AAo	rs17033041	4	G	A	0.186	-0.882	0.128	6.26E-12	0.003	20
β-blockers	*KCNH2*	HF	AAo	rs887586	7	G	C	0.605	0.047	0.015	0.001	0.261	20
β-blockers	*KCNH2*	SBP	AAo	rs887586	7	G	C	0.605	0.32	0.063	3.33E-07	0.022	20
β-blockers	*KCNH2*	HF	ATib	rs882156	7	C	T	0.605	0.106	0.035	0.002	0.23	20
β-blockers	*KCNH2*	SBP	ATib	rs882156	7	C	T	0.605	0.735	0.156	2.29E-06	0.173	20
β-blockers	*KCNH2*	HF	BC	rs11771808	7	G	A	0.577	0.051	0.021	0.014	0.025	20
β-blockers	*KCNH2*	SBP	BC	rs11771808	7	G	A	0.577	0.291	0.083	0.0004	1.95E-09	20
β-blockers	*KCNH2*	HF	BCH	rs11771808	7	G	A	0.577	0.055	0.023	0.017	0.026	20
β-blockers	*KCNH2*	SBP	BCH	rs11771808	7	G	A	0.577	0.313	0.095	0.001	9.59E-08	20
β-blockers	*KCNH2*	HF	NTib	rs3807375	7	T	C	0.345	0.115	0.033	0.0004	0.512	20
β-blockers	*KCNH2*	SBP	NTib	rs3807375	7	T	C	0.345	0.588	0.135	1.22E-05	0.001	20

Figure 4.7 SMR associations of antihypertensive target gene-expression levels with ischemic stroke.

The x-axis represents the β estimates from SMR analysis, and the y-axis represents the -log10(p_{SMR}). The circle, which signifies the corresponding gene, was coloured according to the drug class that targets its protein. β_{SMR} were estimates from SMR analysis per 1 unit increase in the expression of the genes in a particular tissue. $\beta_{SMR} < 0$ suggests that an increase in gene expression levels in a particular tissue was associated with a decrease in MI risk, and vice versa. The blue and orange dashed lines indicate a p-value of 0.05 and a Bonferroni-corrected p-value of 0.0001, respectively. The asterisk indicates HEIDI p-value greater than 0.01; The identifiers for each SMR association circle are the relevant gene name along with the following abbreviations for tissue types: Atib, Artery tibial, BCH, Brain Cerebellar Hemisphere; BC, Brain Cerebellum; MSk, Muscle Skeletal, NTib, Nerve Tibial; Panc, Pancreas; Thy, Thyroid.

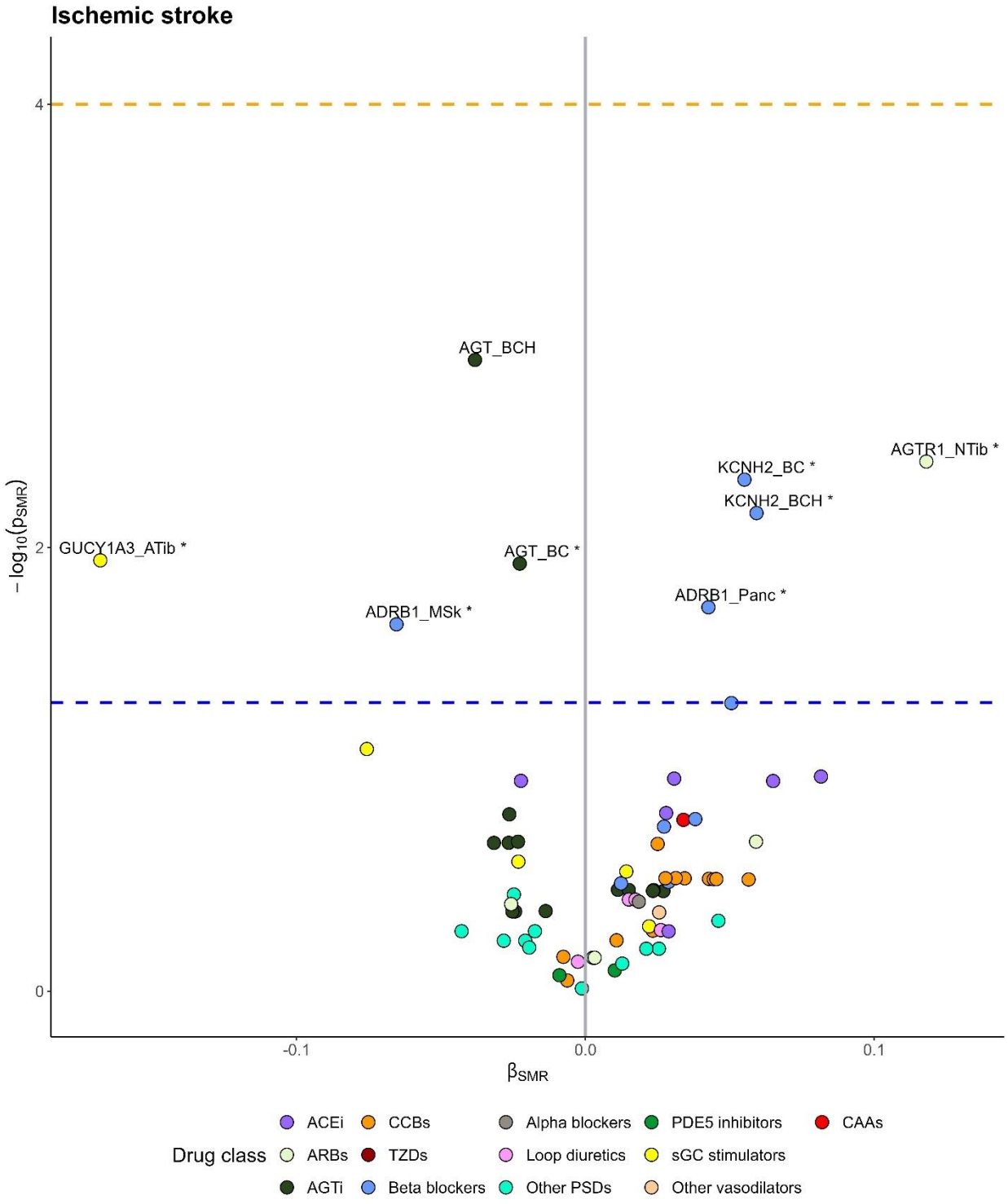


Table 4.10 SMR associations of antihypertensive target gene-expression levels with ischemic stroke.

b_{SMR}, SMR effect size per 1 SD increase in gene expression; se_{SMR}, standard error; A1, effect allele; A2, non-effect allele, Freq, frequency of A1.

Drug class	Gene	Outcome	Tissue	TopSNP	Chr	A1	A2	Freq	b_{SMR}	se_{SMR}	P_{SMR}	p_{HEIDI}	$nsnp_{HEIDI}$
β-blockers	*ADRB1*	Ischemic stroke	MSk	rs7921036	10	T	C	0.407	-0.065	0.029	0.022	0.986	7
β-blockers	*ADRB1*	SBP	MSk	rs7921036	10	T	C	0.407	0.452	0.14	0.001	0.0002	7
β-blockers	*ADRB1*	Ischemic stroke	Panc	rs2532827	10	C	G	0.408	0.043	0.018	0.019	0.31	13
β-blockers	*ADRB1*	SBP	Panc	rs2532827	10	C	G	0.408	-0.29	0.087	0.001	3.14E-07	13
AGTi	*AGT*	Ischemic stroke	BC	rs2493134	1	C	T	0.411	-0.023	0.009	0.012	0.06	20
AGTi	*AGT*	SBP	BC	rs2493134	1	C	T	0.411	-0.472	0.054	4.47E-18	2.10E-06	20
AGTi	*AGT*	Ischemic stroke	BCH	rs5051	1	T	C	0.413	-0.038	0.012	0.001	0.001	20
AGTi	*AGT*	SBP	BCH	rs5051	1	T	C	0.413	-0.467	0.056	1.15E-16	0.0001	20
ARBs	*AGTR1*	Ischemic stroke	NTib	rs275665	3	A	G	0.159	0.118	0.041	0.004	0.137	20
ARBs	*AGTR1*	SBP	NTib	rs275665	3	A	G	0.159	-0.483	0.17	0.005	0.169	20
sGC stimulators	*GUCY1A3*	Ischemic stroke	ATib	rs72687508	4	G	A	0.234	-0.168	0.066	0.011	0.481	20
sGC stimulators	*GUCY1A3*	SBP	ATib	rs72687508	4	G	A	0.234	-2.341	0.493	2.02E-06	0.345	20
β-blockers	*KCNH2*	Ischemic stroke	BC	rs11771808	7	G	A	0.577	0.055	0.02	0.005	0.026	20
β-blockers	*KCNH2*	SBP	BC	rs11771808	7	G	A	0.577	0.291	0.083	0.0004	1.95E-09	20
β-blockers	*KCNH2*	Ischemic stroke	BCH	rs11771808	7	G	A	0.577	0.059	0.022	0.007	0.026	20
β-blockers	*KCNH2*	SBP	BCH	rs11771808	7	G	A	0.577	0.313	0.095	0.001	9.59E-08	20

4.3.7 Chronic kidney disease

13 nominal associations of the expression of 4 genes with CKD risk in 9 tissues were observed, of which all passed the HEIDI test (p_{SMR} 0.01 - 0.039; p_{HEIDI} 0.122 - 0.652). Particularly, the nominal association signals were detected for *CACNB3* in aorta (β_{SMR} [95% CI], 0.140 [0.022 - 0.257]), tibial artery (0.098 [0.019 - 0.177]), adipose subcutaneous (0.108 [0.020 - 0.196]), adipose visceral omentum (0.134 [0.023 - 0.246]), heart atrial appendage (0.143 [0.024 - 0.262]), nerve tibial (0.087 [0.017 - 0.158]), and thyroid (0.178 [0.027 - 0.328]); *PTGER1* in thyroid (0.125 [0.029 - 0.222]); *PPARG* in brain cerebellum (0.028 [0.003 - 0.053]), brain cerebellar hemisphere (0.023 [0.002 - 0.043]), and heart atrial appendage (-0.105 [(-0.194) - (-0.016)]; and *GUCY1A3* in aorta (-0.061 [(-0.118) - (-0.003)]) and tibial artery (-0.189 [(-0.366) - (-0.013)]) (**Figure 4.8, Table 4.11).**

4.3.8 Type 2 diabetes

For type 2 diabetes, there were 23 nominal associations of 9 genes with the disease in 11 tissues. Of these, 21 associations passed the HEIDI test (p_{SMR} 0.002 - 0.048; p_{HEIDI} 0.012 - 0.907). Particularly, the association signals were detected for *PPARG* in heart atrial appendage (β_{SMR} [95% CI] per 1 unit increase in the gene expression, -0.103 [(-0.168) - (-0.039)]); *CACNB3* in tibial artery (0.076 [0.025 - 0.126]), nerve tibial (0.067 [0.022 - 0.112]), aorta (0.107 [0.031 - 0.184]), adipose subcutaneous (0.083 [0.027 - 0.139]), adipose visceral omentum (0.103 [0.032 - 0.175]), heart atrial appendage (0.109 [0.033 - 0.187]), and thyroid (0.137 [0.039 - 0.234]); *KCNH2* in brain cerebellum (-0.053 [(-0.086) - (-0.019)] and brain cerebellar hemisphere (-0.057 [(-0.095) - (-0.019)]); *ACE* in adipose subcutaneous (0.129 [0.036 – 0.222]), brain cerebellum (0.049 [0.013 - 0.084]), brain frontal cortex BA9 (0.051 [0.014 – 0.089]), nerve tibial (0.105 [0.029 - 0.182]), and thyroid (-0.128 [(-0.220) - (-0.036)]); *ADRA2A* in tibial artery (-0.062 [(-0.106) - (-0.017)]); *PDE5A* in aorta (0.105 [0.025 - 0.185]); *CACNA1H* in tibial artery (0.099 [0.01 - 0.188]); *GUCY1A3* in aorta (-0.039 [(-0.077) - (-0.002)]); and *AGT* in adipose subcutaneous (0.017 [0.0001 - 0.035]) (**Figure 4.9, Table 4.12**).

4.3.9 Replication analyses

SMR analysis using data from the eQTLGen consortium found a significant association of increased *PDE5A* expression in blood with an increased risk of CAD (β_{SMR} per 1 unit increase in gene expression in blood, 0.191 [0.113 – 0.269], $p_{SMR} = 1.48\times10^{-06}$), increased *KCNH2* expression in blood with an increased risk of AF (β_{SMR}, 0.249, $p_{SMR} = 2.28\times10^{-07}$). I observed a concordant direction of the association between *GUCY1A3* expression in blood and CAD risk, although the association is not significant (**Table 4.5)**.

Figure 4.8 SMR associations of antihypertensive target gene-expression levels with CKD.
The x-axis represents the β estimates from SMR analysis, and the y-axis represents the -log10(p_{SMR}). The circle, which signifies the corresponding gene, was coloured according to the drug class that targets its protein. β_{SMR} were estimates from SMR analysis per 1 unit increase in the expression of the genes in a particular tissue. $\beta_{SMR} < 0$ suggests that an increase in gene expression levels in a particular tissue was associated with a decrease in MI risk, and vice versa. The blue and orange dashed lines indicate a p-value of 0.05 and a Bonferroni-corrected p-value of 0.0001, respectively. The asterisk indicates HEIDI p-value greater than 0.01; The identifiers for each SMR association circle are the relevant gene name along with the following abbreviations for tissue types: AVO, Adipose Visceral Omentum; ASc, Adipose Subcutaneous; AAo, Artery aorta; Atib, Artery tibial, BCH, Brain Cerebellar Hemisphere; BC, Brain Cerebellum; HAA, Heart Atrial Appendage; NTib, Nerve Tibial; Thy, Thyroid.

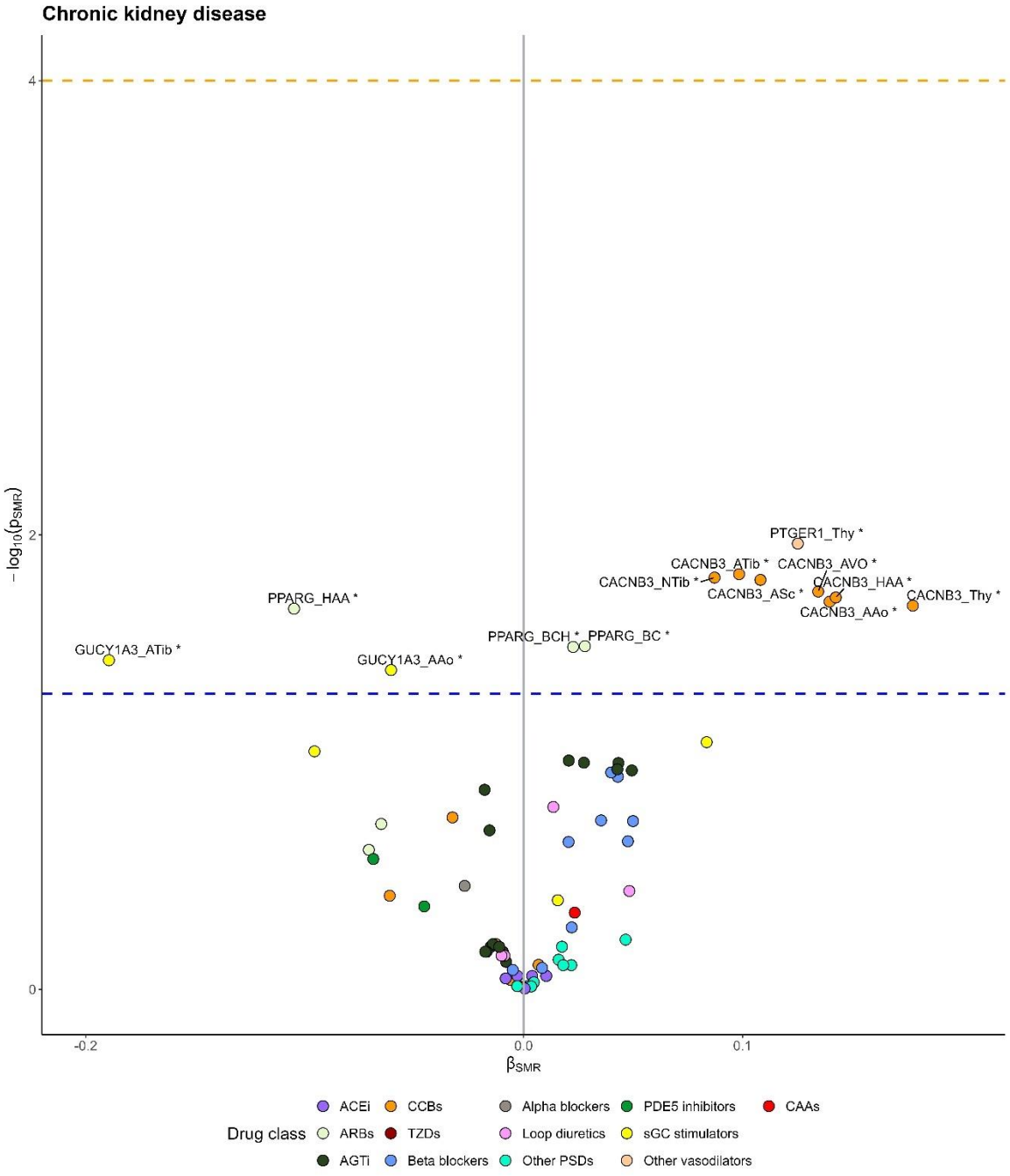


Table 4.11 SMR associations of antihypertensive target gene-expression levels with CKD.

b_{SMR}, SMR effect size per 1 SD increase in gene expression; se_{SMR}, standard error; A1, effect allele; A2, non-effect allele, Freq, frequency of A1.

Drug class	Gene	Outcome	Tissue	TopSNP	Chr	A1	A2	Freq	b_{SMR}	se_{SMR}	P_{SMR}	p_{HEIDI}	$nsnp_{HEIDI}$
CCBs	*CACNB3*	CKD	AAo	rs150857355	12	C	G	0.018	0.14	0.06	0.02	0.468	5
CCBs	*CACNB3*	SBP	AAo	rs150857355	12	C	G	0.018	1.461	0.281	2.06E-07	2.35E-05	5
CCBs	*CACNB3*	CKD	ASc	rs150857355	12	C	G	0.018	0.108	0.045	0.016	0.265	20
CCBs	*CACNB3*	SBP	ASc	rs150857355	12	C	G	0.018	1.131	0.178	2.32E-10	0.005	20
CCBs	*CACNB3*	CKD	ATib	rs150857355	12	C	G	0.018	0.098	0.04	0.015	0.122	20
CCBs	*CACNB3*	SBP	ATib	rs150857355	12	C	G	0.018	1.029	0.152	1.43E-11	0.0001	19
CCBs	*CACNB3*	CKD	AVO	rs150857355	12	C	G	0.018	0.135	0.057	0.018	0.282	9
CCBs	*CACNB3*	SBP	AVO	rs150857355	12	C	G	0.018	1.407	0.249	1.51E-08	8.46E-05	9
CCBs	*CACNB3*	CKD	HAA	rs150857355	12	C	G	0.018	0.143	0.061	0.019	0.33	7
CCBs	*CACNB3*	SBP	HAA	rs150857355	12	C	G	0.018	1.491	0.277	7.63E-08	0.0003	7
CCBs	*CACNB3*	CKD	NTib	rs150857355	12	C	G	0.018	0.087	0.036	0.015	0.262	20
CCBs	*CACNB3*	SBP	NTib	rs150857355	12	C	G	0.018	0.911	0.14	8.20E-11	0.005	20
CCBs	*CACNB3*	CKD	Thy	rs150857355	12	C	G	0.018	0.178	0.077	0.02	0.377	20
CCBs	*CACNB3*	SBP	Thy	rs150857355	12	C	G	0.018	1.858	0.369	4.85E-07	0.001	19
sGC stimulators	*GUCY1A3*	CKD	AAo	rs17033041	4	G	A	0.186	-0.061	0.029	0.039	0.237	20
sGC stimulators	*GUCY1A3*	SBP	AAo	rs17033041	4	G	A	0.186	-0.882	0.128	6.26E-12	0.003	20
sGC stimulators	*GUCY1A3*	CKD	ATib	rs72687508	4	G	A	0.234	-0.189	0.09	0.036	0.514	20
sGC stimulators	*GUCY1A3*	SBP	ATib	rs72687508	4	G	A	0.234	-2.341	0.493	2.02E-06	0.345	20
ARBs	*PPARG*	CKD	BC	rs73025230	3	C	A	0.096	0.028	0.013	0.031	0.57	20
ARBs	*PPARG*	SBP	BC	rs73025230	3	C	A	0.096	0.089	0.037	0.017	0.01	20
ARBs	*PPARG*	CKD	BCH	rs73025230	3	C	A	0.096	0.023	0.01	0.031	0.652	20
ARBs	*PPARG*	SBP	BCH	rs73025230	3	C	A	0.096	0.072	0.03	0.017	0.006	20
ARBs	*PPARG*	CKD	HAA	rs2920502	3	C	G	0.316	-0.105	0.045	0.021	0.416	18
ARBs	*PPARG*	SBP	HAA	rs2920502	3	C	G	0.316	-0.534	0.161	0.001	0.083	18
Other vasodilators	*PTGER1*	CKD	Thy	rs3760702	19	A	G	0.317	0.125	0.049	0.011	0.418	13
Other vasodilators	*PTGER1*	SBP	Thy	rs3760702	19	A	G	0.317	0.403	0.152	0.008	0.985	13

Figure 4.9 SMR associations of antihypertensive target gene-expression levels with T2D.
The x-axis represents the β estimates from SMR analysis, and the y-axis represents the -log10(p_{SMR}). The circle, which signifies the corresponding gene, was coloured according to the drug class that targets its protein. β_{SMR} were estimates from SMR analysis per 1 unit increase in the expression of the genes in a particular tissue. $\beta_{SMR} < 0$ suggests that an increase in gene expression levels in a particular tissue was associated with a decrease in MI risk, and vice versa. The blue and orange dashed lines indicate a p-value of 0.05 and a Bonferroni-corrected p-value of 0.0001, respectively. The asterisk indicates HEIDI p-value greater than 0.01; The identifiers for each SMR association circle are the relevant gene name along with the following abbreviations for tissue types: AVO, Adipose Visceral Omentum; ASc, Adipose Subcutaneous; AdrG, Adrenal Gland; AAo, Artery aorta; Atib, Artery tibial, BCaud, Brain Caudate basal ganglia; BCH, Brain Cerebellar Hemisphere; BC, Brain Cerebellum; BCor, Brain Cortex; BFCB9, Brain Frontal Cortex BA9; BNacc, Brain Nucleus accumbens basal ganglia; BPut, Brain Putamen basal ganglia; HAA, Heart Atrial Appendage; HLV, Heart Left Ventricle, KC, Kidney Cortex, MSk, Muscle Skeletal, NTib, Nerve Tibial; Panc, Pancreas; Thy, Thyroid; WB, Whole Blood.

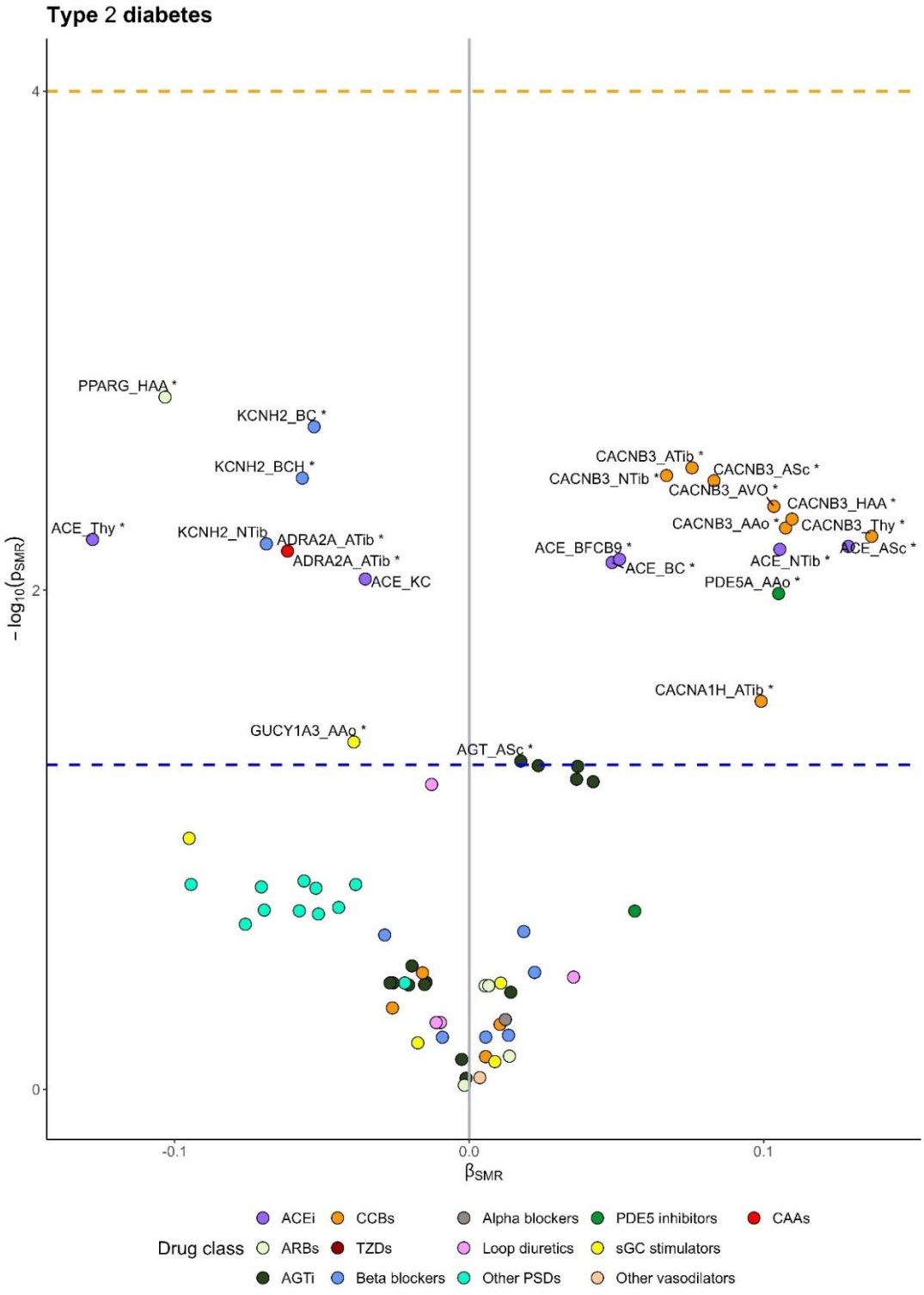


Table 4.12 SMR associations of antihypertensive target gene-expression levels with T2D.

b_{SMR}, SMR effect size per 1 SD increase in gene expression; se_{SMR}, standard error; A1, effect allele; A2, non-effect allele, Freq, frequency of A1.

Drug class	Gene	Outcome	Tissue	TopSNP	Chr	A1	A2	Freq	b_{SMR}	se_{SMR}	P_{SMR}	p_{HEIDI}	$nsnp_{HEIDI}$
ACEi	*ACE*	SBP	ASc	rs4308	17	G	A	0.633	1.768	0.334	1.22E-07	0.071	16
ACEi	*ACE*	T2D	ASc	rs4308	17	G	A	0.633	0.129	0.047	0.007	0.167	16
ACEi	*ACE*	SBP	BC	rs4308	17	G	A	0.633	0.666	0.135	7.50E-07	0.2	11
ACEi	*ACE*	T2D	BC	rs4308	17	G	A	0.633	0.049	0.018	0.008	0.097	11
ACEi	*ACE*	SBP	BFCB9	rs4291	17	A	T	0.628	0.715	0.148	1.46E-06	0.213	4
ACEi	*ACE*	T2D	BFCB9	rs4291	17	A	T	0.628	0.051	0.019	0.008	0.102	4
ACEi	*ACE*	SBP	KC	rs4308	17	G	A	0.633	-0.484	0.104	3.50E-06		
ACEi	*ACE*	T2D	KC	rs4308	17	G	A	0.633	-0.035	0.013	0.009		
ACEi	*ACE*	SBP	NTib	rs4292	17	T	C	0.632	1.459	0.293	6.61E-07	0.21	12
ACEi	*ACE*	T2D	NTib	rs4292	17	T	C	0.632	0.106	0.039	0.007	0.111	12
ACEi	*ACE*	SBP	Thy	rs12449782	17	A	G	0.419	-0.681	0.229	0.003	0.011	13
ACEi	*ACE*	T2D	Thy	rs12449782	17	A	G	0.419	-0.128	0.047	0.006	0.024	13
Alpha blockers	*ADRA2A*	SBP	ATib	rs1343449	10	G	A	0.303	0.222	0.107	0.038	0.908	20
Alpha blockers	*ADRA2A*	T2D	ATib	rs1343449	10	G	A	0.303	-0.062	0.023	0.007	0.012	20
CAAs	*ADRA2A*	SBP	ATib	rs1343449	10	G	A	0.303	0.222	0.107	0.038	0.908	20
CAAs	*ADRA2A*	T2D	ATib	rs1343449	10	G	A	0.303	-0.062	0.023	0.007	0.012	20
AGTi	*AGT*	SBP	ASc	rs5050	1	G	T	0.178	0.126	0.043	0.004	9.36E-08	20
AGTi	*AGT*	T2D	ASc	rs5050	1	G	T	0.178	0.018	0.009	0.048	0.646	20
CCBs	*CACNA1H*	SBP	ATib	rs35198836	16	T	C	0.257	0.593	0.235	0.012	0.659	12
CCBs	*CACNA1H*	T2D	ATib	rs117177120	16	G	T	0.066	0.099	0.045	0.028	0.877	16
CCBs	*CACNB3*	SBP	AAo	rs150857355	12	C	G	0.018	1.461	0.281	2.06E-07	2.35E-05	5
CCBs	*CACNB3*	T2D	AAo	rs150857355	12	C	G	0.018	0.107	0.039	0.006	0.042	5
CCBs	*CACNB3*	SBP	ASc	rs150857355	12	C	G	0.018	1.131	0.178	2.31E-10	0.006	20
CCBs	*CACNB3*	T2D	ASc	rs150857355	12	C	G	0.018	0.083	0.029	0.004	0.107	20
CCBs	*CACNB3*	SBP	ATib	rs150857355	12	C	G	0.018	1.029	0.152	1.43E-11	0.0001	19

CCBs	*CACNB3*	T2D	ATib	rs150857355	12	C	G	0.018	0.076	0.026	0.003	0.177	20
CCBs	*CACNB3*	SBP	AVO	rs150857355	12	C	G	0.018	1.407	0.249	1.51E-08	8.46E-05	9
CCBs	*CACNB3*	T2D	AVO	rs150857355	12	C	G	0.018	0.103	0.037	0.005	0.145	9
CCBs	*CACNB3*	SBP	HAA	rs150857355	12	C	G	0.018	1.491	0.277	7.63E-08	0.0003	7
CCBs	*CACNB3*	T2D	HAA	rs150857355	12	C	G	0.018	0.11	0.039	0.005	0.098	7
CCBs	*CACNB3*	SBP	NTib	rs150857355	12	C	G	0.018	0.911	0.14	8.20E-11	0.005	20
CCBs	*CACNB3*	T2D	NTib	rs150857355	12	C	G	0.018	0.067	0.023	0.003	0.225	20
CCBs	*CACNB3*	SBP	Thy	rs150857355	12	C	G	0.018	1.858	0.369	4.85E-07	0.001	19
CCBs	*CACNB3*	T2D	Thy	rs150857355	12	C	G	0.018	0.137	0.05	0.006	0.027	20
sGC stimulators	*GUCY1A3*	SBP	AAo	rs17033041	4	G	A	0.186	-0.882	0.128	6.26E-12	0.003	20
sGC stimulators	*GUCY1A3*	T2D	AAo	rs17033041	4	G	A	0.186	-0.039	0.019	0.041	0.796	20
Beta blockers	*KCNH2*	SBP	BC	rs11771808	7	G	A	0.577	0.291	0.083	0.0004	1.95E-09	20
Beta blockers	*KCNH2*	T2D	BC	rs11771808	7	G	A	0.577	-0.053	0.017	0.002	0.267	20
Beta blockers	*KCNH2*	SBP	BCH	rs11771808	7	G	A	0.577	0.313	0.095	0.001	9.59E-08	20
Beta blockers	*KCNH2*	T2D	BCH	rs11771808	7	G	A	0.577	-0.057	0.019	0.004	0.397	20
Beta blockers	*KCNH2*	SBP	NTib	rs3807375	7	T	C	0.345	0.588	0.135	1.22E-05	0.001	20
Beta blockers	*KCNH2*	T2D	NTib	rs3807375	7	T	C	0.345	-0.069	0.025	0.007	0.0002	20
PDE5 inhibitors	*PDE5A*	SBP	AAo	rs1393966	4	C	T	0.538	0.619	0.202	0.002	0.093	20
PDE5 inhibitors	*PDE5A*	T2D	AAo	rs1393966	4	C	T	0.538	0.105	0.041	0.01	0.907	20
ARBs	*PPARG*	SBP	HAA	rs2920502	3	C	G	0.316	-0.534	0.161	0.001	0.083	18
ARBs	*PPARG*	T2D	HAA	rs2920502	3	C	G	0.316	-0.103	0.033	0.002	0.022	18

4.4 Discussions

This study utilised a SMR approach to explore the causal relationships between gene expression levels of proteins targeted by antihypertensive drugs and various cardiovascular, diabetic, and renal outcomes. Significant findings emerged from the analysis of 19 genes associated with SBP, linking the expression of these genes to multiple disease states including CAD, MI, AF, HF, ischemic stroke, CKD, and T2D. Among 19 genes, I observed 26 associations of the expression of 7 genes with the risk of CAD in 13 tissues, 8 associations of 5 genes with MI risk in 7 tissues, 8 associations of 4 genes with AF risk in 5 tissues, 11 associations of 4 genes with HF risk in 7 tissues, 8 associations of 5 genes with ischemic stroke in 6 tissues, 13 associations of 14 genes with CKD risk in 9 tissues, and 23 associations of 9 genes with T2D in 11 tissues. 4 associations remained significant after accounting for multiple tests. Particularly, genetically increased *GUCY1A3* expression in tibial artery and thyroid was associated with a decrease in SBP and a decrease in the risk of CAD. Genetically increased *PDE5A* expression in aorta was associated with an increase in SBP and an increase in the risk of CAD. Genetically increased *KCNH2* expression in the brain cerebellum was associated with an increase in SBP and an increase in AF risk. Colocalisation analysis suggested a shared causal variant for genetically predicted *GUCY1A3* expression in tibial artery and CAD risk, whereas it suggested the distinct causal variants for genetically predicted *KCNH2* expression in brain cerebellum and AF risk. For genetically predicted *PDE5A* expression in aorta and CAD risk, colocalisation analysis gave average posterior probability for the model with a shared causal variant (H_4 ~ 50%).

In Chapter 3, 33 associations were detected between antihypertensive drug classes and outcomes, with 17 of those remaining significant after multiple testing corrections. In this chapter, among those 33 associations, SMR analysis detected gene expression-disease associations for 16 pairs of drug classes and disease outcomes (**Table 4.13**). The significant associations detected between gene expression levels and disease risk (*PDE5A* and CAD, *GUCY1A3* and CAD, *KCNH2* and AF) corroborated the significant associations of genetically proxied sGC stimulators, PDE5 inhibitors with CAD risk, and genetically proxied beta-blockers with AF risk.

Table 4.13 Table illustrating the results from MR analysis (Chapter 3) and SMR analysis in this chapter.

y, indicates an association detected between drug class and the disease; (*), indicates significant association after adjusting for multiple tests; AVO, Adipose Visceral Omentum; ASc, Adipose Subcutaneous; AdrG, Adrenal Gland; AAo, Artery aorta; Atib, Artery tibial, BCaud, Brain Caudate basal ganglia; BCH, Brain Cerebellar Hemisphere; BC, Brain Cerebellum; BCor, Brain Cortex; BFCB9, Brain Frontal Cortex BA9; BNacc, Brain Nucleus accumbens basal ganglia; BPut, Brain Putamen basal ganglia; HAA, Heart Atrial Appendage; HLV, Heart Left Ventricle, KC, Kidney Cortex, MSk, Muscle Skeletal, NTib, Nerve Tibial; Panc, Pancreas; Thy, Thyroid; WB, Whole Blood.

Drug Class	**Outcome**	**MR**	**SMR** (gene_significant tissue)
AGTi	CAD	y *	AGT_BC, AGT_BCH
	MI	y	AGT_BC, AGT_BCH
	AF		AGT_BCH
	HF		
	Stroke	y *	AGT_BC, AGT_BCH
	T2D		
	CKD		
ACEi	CAD	y	ACE_Ntib, ACE_BFCB9, ACE_Asc, ACE_BC, ACE_KC
	MI		
	AF		
	HF		
	Stroke		
	T2D	y	ACE_Asc, ACE_Ntib, ACE_BFCB9, ACE_Thy, ACE_KC, ACE_BC
	CKD		
ARB	CAD		
	MI		
	AF		PPARG_BC, PPARB_BCH, PPARG_HAA
	HF		
	Stroke		AGTR1_NTib
	T2D		PPARG_HAA
	CKD		PPARG_HAA, PPARG_BC, PPARG_BCH
CCB	CAD	y *	CACNB3_Atib, CACNB3_Asc, CACNB3_Ntib, CACNB3_AVO, CACNB3_HAA, CACNB3_Thy, CACNB3_Aao
	MI	y *	
	AF	y *	
	HF	y *	CACNA1B_BCH, CACNA1B_BC
	Stroke	y *	
	T2D		CACNB3_Atib, CACNB3_Asc, CACNB3_Ntib, CACNB3_AVO, CACNB3_HAA, CACNB3_Thy, CACNB3_Aao, CACNA1N_Atib
	CKD		CACNB3_Atib, CACNB3_Asc, CACNB3_Ntib, CACNB3_AVO, CACNB3_HAA, CACNB3_Thy, CACNB3_Aao
β-blocker	CAD	y *	ADRB1_WB, KCNH2_BC, KCNH2_Aao, KCNH2_BCH, KCNH2_Atib, KCNH2_AdrG, KCNH2_Ntib
	MI	y *	
	AF	y *	ADRB1_WB, KCNH2_BC *, KCNH2_Ntib, KCNH2_BCH
	HF	y	ADRB1_MSk, ADRB1_Panc, KCNH2_Ntib, KCNH2_Aao, KCNH2_Atib, KCNH2_BC, KCNH2_BCH
	Stroke		ADRB1_MSk, ADRB1_Panc, KCNH2_BC, KCNH2_BCH

	T2D		KCNH2_BC, KCNH2_BCH, KCNH2_Ntib
	CKD		
α-blocker	CAD	y	
	MI		
	AF		
	HF		
	Stroke		
	T2D		
	CKD	y	
Other PSD	CAD		
	MI		
	AF		
	HF		
	Stroke	y	
	T2D	y	
	CKD		
Loop diuretic	CAD	y *	
	MI		
	AF	y	
	HF		
	Stroke		
	T2D		
	CKD		
ERA	CAD	y *	
	MI	y	
	AF	y	
	HF		
	Stroke	y *	
	T2D		
	CKD		
PDE5i	CAD	y *	PDE5A_Aao *
	MI	y	PDE5A_AAo
	AF	y	
	HF	y	
	Stroke	y *	
	T2D		PDE5A_AAo
	CKD		
sGCs	CAD	y *	GUCY1A3_Atib *, GUCY1A3_Thy *, GUCY1A3_AAo, GUCY1A3_HAA
	MI	y *	GUCY1A3_Atib, GUCY1A3_Thy, GUCY1A3_HAA
	AF		
	HF		GUCY1A3_Aao
	Stroke	y	GUCY1A3_Atib
	T2D		GUCY1A3_AAo
	CKD	y *	GUCY1A3_ATib, GUCY1A3_Aao

4.4.1 *GUCY1A3* and CAD risk

GUCY1A3, also known as guanylate cyclase soluble subunit alpha 3, is located on chromosome 4 at band 4q32.1, starting from 156,587,863 bp to 156,653,501 bp (GRCh37/hg 19) (359). *GUCY1A3* encodes an alpha subunit of soluble guanylate cyclase (sGC) complex, which are heterodimeric proteins that consist of an alpha and a beta subunit. The sGC enzyme, which is activated by NO, catalyses the conversion of GTP to cGMP and pyrophosphate. cGMP plays an important role in cellular and physiological responses including vasodilation, inhibition of platelet aggregation and smooth muscle cell relaxation (360). The critical roles of the NO-sGC-cGMP pathway in mediating CAD and MI risk have been highlighted by large-scale GWAS meta-analysis and post-GWAS studies (274, 361). *GUCY1A3* locus is one of the genomic loci that have been identified to be associated with CAD and MI risk in GWAS. The role of genetic variation in the *GUCY1A3* gene in the progression of human atherosclerosis has been emphasized in previous research, which discovered a digenic loss-of-function mutation affecting *GUCY1A3* that is linked with CAD and MI in an extended family (362). Rare coding variants in *GUCY1A3* were found to be enriched in patients with premature CAD and MI (363). More recently, a functional study, which performed a sequence of experiments using human samples and cell lines, highlights the involvement of the lead SNP at the chromosome 4q32.1 risk locus in regulating the gene expression of *GUCY1A3*. Risk allele carriers exhibit reduced suppression of both vascular smooth muscle cell migration and platelet aggregation following the activation of the sGC enzyme. Increased sGC expression is correlated with lower risk of atherosclerosis in data from human and mouse samples (361).

Riociguat and vericiguat are sGC stimulators that have been approved for the treatment of pulmonary hypertension and chronic heart failure. Several novel sGC stimulators or activators are being developed, showing potential for the treatment of cardiovascular and kidney disease (343, 344). Furthermore, in Chapter 3 I also detected a significant association between genetically proxied sGC stimulation with a decreased risk of CAD and MI. Being consistent with the evidence from previous studies and Chapter 1, my study using a MR approach that integrates summary data from GWAS and eQTL studies also demonstrates the potential causal association of an increased *GUCY1A3* gene expression in tibial artery with lower SBP and a reduced risk of CAD and MI. I also detect a nominal association of *GUCY1A3* expression in the thyroid and aorta with CAD risk. The significant association between *GUCY1A3* gene expression in tibial artery and CAD risk was corroborated by the results from the HEIDI test and colocalisation analysis (H_4 > 0.99). Colocalisation analysis gave a moderate posterior probability for the model with a shared causal variant (H_4 ~ 60%) for the pair of SBP and *GUCY1A3* expression in tibial

artery, which suggests insufficient power to show colocalisation for the 2 traits but no evidence for confounding by linkage disequilibrium.

Moreover, *GUCY1A3* gene expression in the tibial artery was nominally associated with the risk of ischemic stroke. Both CAD and stroke share the same pathophysiological mechanism, namely atherosclerosis. Ischemic stroke is attributed to atherosclerosis, where the development of thrombosis can result in the blockage of the cerebral artery, potentially causing permanent brain damage. As the NO-sGC-cGMP pathway plays a crucial role in maintaining vascular homeostasis, disruptions in this pathway can lead to vascular dysfunction, which in turn contributes to the pathogenesis of stroke. For example, a reduced NO production or dysfunction in sGC activity can lead to reduced cGMP levels. This can cause vasoconstriction, platelet aggregation, and inflammation, all of which can elevate the risk of stroke. The role of genetic variation in the *GUCY1A3* gene in the progression of human atherosclerosis has been demonstrated in previous studies (362). Several common genetic variants located in *GUCY1A3* have been identified from GWAS, these are associated with an increased risk of CAD, and MI and a reduction in platelet inhibition when stimulated by NO (364, 365). Genetic predisposition for enhanced signalling of nitric oxide, as mediated by common variants in the *NOS3* and *GUCY1A3* genes, has been associated with a decrease in blood pressure and a reduced susceptibility to CAD, peripheral arterial disease and stroke (364). My current study finds the nominal association of an increased *GUCY1A3* gene expression in tibial artery with a reduced risk of ischemic stroke. This finding reinforces the role of the NO-sGC-cGMP pathway in vascular disease leading to atherosclerosis diseases like CAD, MI, and stroke.

There was also suggestive evidence for the association of *GUCY1A3* gene expression with the risk of HF, CKD and T2D, although these associations did not reach significance after Bonferroni correction. The association signals are mainly detected in the tibial artery and aorta. However, due to the low sample sizes for some tissues (such as the kidney cortex), the analysis might have low statistical power to detect the associations in those tissues. I should not discount the possible association in other tissues, and eQTL studies with larger sample sizes for relevant tissues will help to overcome this limitation.

The findings align with pharmacological studies on sGC stimulators used in treating pulmonary hypertension and heart failure, suggesting that enhancing GUCY1A3 expression could similarly benefit CAD management.

4.4.2 *PDE5A* and CAD risk

PDE5A gene, also known as phosphodiesterase 5A, is located in chromosome 4, starting from 120,415,550 bp to 120,550,146 bp (GRCh37/hg 19) (359). The expression of the *PDE5A* gene has low tissue specificity, with its cytoplasmic protein being observed in most cell types. The protein encoded by the *PDE5A* gene is a cGMP-binding, cGMP-

specific phosphodiesterase, which is a member of the cyclic nucleotide phosphodiesterase family. This phosphodiesterase degrades the second messenger cGMP by hydrolysing it to 5'-GMP. The phosphodiesterase is involved in regulating the levels of intracellular cyclic nucleotides and plays a crucial role in the relaxation of smooth muscle in the cardiovascular system. PDE5 inhibitors are drugs that block the degradative action of cGMP-specific phosphodiesterase type 5, leading to increased levels of cGMP and prolonged signalling actions of NO. These drugs have been used to treat erectile dysfunction and pulmonary arterial hypertension while emerging evidence from animal experiments and observational studies suggests that PDE5 inhibitors may have potential cardioprotective benefits (324, 327, 331).

One of the variants identified by GWAS of CAD is situated in the proximity of *PDE5A* on chromosome 4q27 and linked to increased levels of *PDE5A* mRNA in aorta (366). A functional analysis has identified a variant at the *PDE5A* locus, which is found to be cosegregated with CAD in an extended family with a high prevalence of early-onset CAD. Data from human samples indicated that the variant affected the expression levels of the PDE5A2 isoform, reinforcing the hypothesis that the *PDE5A* gene influences the risk of CAD (367). My current SMR study also supported this hypothesis. I demonstrated a potential causal association between genetically increased *PDE5A* gene expression in aorta was associated with higher SBP and an increase in CAD and MI risk, although it is a nominal association for MI. I obtained a moderate posterior probability (~ 50%) for the colocalisation of *PDE5A* gene expression in aorta with CAD, suggesting a low power to demonstrate colocalisation for the 2 traits. An eQTL data for aorta with a larger sample size might increase the power of the colocalisation analysis. Using data from the eQTLGen consortium, I also observed the consistent association of the *PDE5A* expression in blood and the risk of CAD. My findings support the positive impact of *PDE5A* gene expression, particularly in aorta, on the risk of CAD. This is also corroborated by the results from Chapter 1, in which I demonstrated a causal association between genetically predicted PDE5 inhibition and a decreased risk of CVDs including CAD and MI.

While PDE5 inhibitors are well-known for treating erectile dysfunction and pulmonary hypertension, this finding suggests consideration of their cardiovascular effects, especially in patients at risk of CAD.

4.4.3 *KCNH2* and *AF* risk

KCNH2 gene, also known as potassium voltage-gated channel subfamily H member 2, is located in chromosome 7, starting from 150,642,049 bp to 150,675,403 bp (GRCh37/hg 19) (359). A protein encoded by *KCNH2* is a component of a voltage-activated potassium channel, which is located in cardiac muscle, nerve cells, and microglia. Several types of

potassium channels (such as Kv7.1, Kv11.1, Kv4.3, Kir2.1, KCa2.3), which are present on the surface of cardiac cells, have the responsibility of preserving the resting membrane potential and also contribute to the various stages of the repolarization process. The formation of a functional potassium channel involves the interaction of four pore-forming α-subunits and other partner proteins, including accessory β-subunits. The KCNH2 gene encodes the α-subunit of the voltage-gated potassium channel Kv11.1, which generates the rapidly repolarizing potassium current.

A large-scale GWAS including 60,620 AF cases and 970,216 controls have identified 111 genomic loci and prioritised 151 candidate genes, with KCNH2 being one the candidate genes (275). Previous studies have suggested the association between mutations in the *KCNH2* gene and a higher incidence of AF (368). The common K897-allele of the *KCNH2*-K897T variant was reported to be associated with AF in a two-stage association study involving 1207 AF cases and 2475 controls (369). A missense mutation, which results in the amino acid change in *KCNH2*, was identified in a family with short QT syndrome and a high incidence of AF. This mutation resulted in a gain of function of the rapid repolarizing potassium current, leading to a shortening of the action potential duration (370). Another study of families with AF has identified genetic variants, of which most are mutations, linked to AF. The detected mutations in gene *KCNH2* include both gain-of-function and loss-of-function mutations (371). KCNH2 protein is one of the protein targets of sotalol, a beta-blocker that has been used to treat AF and arrhythmia. In my current SMR study, genetically increased *KCNH2* expression in the brain cerebellum is associated with an increased risk of AF. I also find suggestive evidence for the association of *KCNH2* expression in the brain cerebellar hemisphere. However, colocalisation analysis gave a high posterior probability for the H_3, suggesting distinct causal variants for the 2 traits.

Suggestive evidence was found for the association of *KCNH2* expression with the risk of CAD (tibial artery, nerve tibial, and adrenal gland), HF (tibial artery, nerve tibial), ischemic stroke (brain cerebellum, brain cerebellar hemisphere), and T2D (brain cerebellum, brain cerebellar hemisphere).

These findings suggest a neuro-cardiac regulatory axis that could be explored for novel AF treatments.

4.4.4 Suggestive evidence for *CACNB3*, *CACNA1B* and *ACE* gene expression

There was suggestive evidence for the association of the following genes: *CACNB3* with the risk of CAD, CKD, and T2D; *CACNA1B* with HF risk; and *ACE* with the risk of CAD and T2D. These findings are complementary to the results in Chapter 1, in which I found

protective effects of genetically proxied calcium channel blockers against CAD and HF, and protective effects of genetically proxied ACEi against CAD and T2D.

ACE, also known as angiotensin-converting enzyme, is a gene located on the long arm of chromosome 17 (17q23), starting from 61,554,422 bp to 61,599,205 bp (GRCh37/hg 19) (359). *ACE* gene encodes an enzyme that catalyzes the conversion of angiotensin I into angiotensin II, which is a potent vasopressor and aldosterone-stimulating peptide (36). Angiotensin II plays a crucial role in regulating BP and fluid-electrolyte balance. The enzyme encoded by *ACE* is a target for ACEi, a first-line drug class in treating hypertension. *ACE* Insertion/Deletion (I/D) is a polymorphism in the *ACE* gene, which was reported as a potent risk factor for coronary heart disease in previous studies (372, 373). The DD genotype, associated with higher levels of circulating ACE than other genotypes, was associated with an increased risk of developing myocardial infarction (372). A MR study reported an association between genetically lower ACE serum concentrations and a decreased risk of T2D (374). Evidence from RCTs gives inconclusive evidence regarding the causal relationship between ACE inhibition and protection against T2D (42, 375, 376). My study provides suggestive evidence supporting the association of ACE gene expression, particularly in adipose subcutaneous, nerve tibial, thyroid, brain cerebellum, and brain frontal cortex BA9, with the risk of T2D.

CACNB3, also known as Calcium Voltage-Gated Channel Auxiliary Subunit Beta 3, is a gene located in chromosome 12, starting from 49,207,577 bp to 49,222,724 bp (GRCh37/hg 19) (359). The CACNB3 gene encodes the beta-3 subunit (Cavβ3) of the voltage-dependent calcium channels, which play a crucial role in the contraction of heart muscles and the conduction of nerve impulses. The role of the *CACNB3* gene in cardiovascular and kidney diseases is not well-established. An experimental study has demonstrated that Cavβ3 regulates Ca^{2+} signalling and insulin expression in pancreatic β-cells in a cell-autonomous manner, through the interaction with the IP_3 receptor in isolated β cells (377). In my current SMR study, I have demonstrated suggestive evidence for the association between genetically increased CACNB3 expression and an increased risk of CAD, CKD, and T2D. The association signals were detected in tibial artery, aorta, nerve tibial, heart atrial appendage, adipose subcutaneous, adipose visceral omentum, and thyroid.

I detect a nominal association of genetically proxied *CACNA1B* expression in brain cerebellum and brain cerebellar hemisphere with the risk of HF. *CACNA1B* is a gene located in chromosome 9 (137,877,782 bp to 138,124,624 bp, GRCh37/hg 19) (359). The *CACNA1B* gene encodes the alpha-1B subunit of the N-type voltage-dependent calcium channel (NCC), which plays a pivotal role in regulating neurotransmitter release at sympathetic nerve terminals. Mice that are deficient in the *CACNA1B* gene exhibit a functional decline in their sympathetic nervous system, which can precipitate the

occurrence of ventricular arrhythmias and sudden death in individuals with HF (378). A study was conducted on a transgenic mice model of dilated cardiomyopathy (dnNRSF-Tg) to compare the effects of cilnidipine, a dual N- and L-type Ca^{2+} channel blocker, with those of nitrendipine, a selective L-type Ca^{2+} channel blocker. Cilnidipine was shown to substantially reduce arrhythmias, enhance the survival rate, and rectify the imbalance between cardiac sympathetic and parasympathetic nervous system activity. Genetic titration of the N-type Ca^{2+} channel, which was accomplished by crossing dnNRSF-Tg mice with mice lacking CACNA1B, enhanced the survival rate (379).

4.4.5 Strengths

This SMR study used a strict p-value threshold (5×10^{-8}) to select cis-eQTL for the genetic instruments. Systolic blood pressure was used for positive control testing as well as genetic instrument validation. HEIDI-test and colocalisation analysis were performed to distinguish between true causal association and linkage disequilibrium. I also used eQTL data from the eQTLGen consortium to replicate my findings from the main analysis. My SMR study leverages data from the largest publicly available GWAS for the phenotypes. GTEx version 8 provides cis-eQTL data in a wide range of tissues, facilitating the investigation of gene expression in multiple tissues. Thus, this SMR analysis is not limited to a disease-specific tissue; furthermore, it is often challenging to identify the primary tissue associated with a specific disease.

In Chapter 3, MR analyses were performed to examine the effect of antihypertensive drug classes on the disease outcomes. In this Chapter, I conduct the analysis focusing on the individual genes that encode the target for the antihypertensive drug classes. By doing this, I provide a more profound understanding of the effects of antihypertensive drugs on disease outcomes. My study is the first study to provide this in-depth view of the effects of a wide range of antihypertensive drug classes - at the drug class level as well as the individual target gene expression level - on cardiovascular, diabetic and renal diseases.

4.4.6 Limitations

The statistical power of tissue-specific analysis is limited by the sample size of available eQTL data sets. The sample sizes of eQTL cohorts are considerably less than those of GWAS, which may result in a low statistical power. To alleviate this issue, I conducted replication in different tissues and different eQTL cohorts to evaluate the robustness of my findings. Moreover, given small sample sizes in some tissues from the GTEx cohort (such as the kidney cortex with only 70 samples), the possible function of genes with P_{SMR} less than the Bonferroni-corrected p-value cannot be ruled out.

Ten out of the total 56 genes encoding targets for antihypertensive drugs, summary statistics for eQTLs at the p-value threshold of $5x10^{-8}$ were unavailable in the GTEx

cohort. The SMR analysis is limited to the availability of eQTL data in some specific tissues. As the genetic determinants of gene expression vary between tissues, the lack of eQTL data in tissues relevant to the disease can lead to the missing tissue-specific effects of the gene on the disease outcome. For example, the *AGT* gene (also known as angiotensinogen) primarily functions in the liver. Angiotensinogen was targeted by zilebesiran, an emerging RNA interference therapeutic candidate developed to reduce the production of angiotensinogen in the liver (60). However, eQTL summary statistics for this gene in the liver are unavailable when using the p-value threshold of $5x10^{-8}$. Thus, eQTL data for *AGT* in hepatic tissues from other cohorts with sufficient sample size will be necessary to investigate the effects of this gene.

Cell types and tissue environments utilised in eQTL may be different to those used in GWAS. GWAS investigating the phenotypes were performed using DNA derived from blood samples, which were collected from living population groups and clinical cohorts. Whereas, eQTL analyses in GTEx cohort utilised RNA estimates from post-mortem tissues. The differences in cell types and tissue environments may result in a potential variation in the regulatory environment of the tissue types.

The effect sizes estimated from MR analysis may not reflect the true effect of the gene-expression modulation. The effect sizes from GTEx were calculated in a normalised space where the magnitude cannot be used for any direct biological interpretation (285).

Although I applied additional sensitivity analysis to evaluate the robustness of my findings, horizontal pleiotropy is still possible in the observed associations. Horizontal pleiotropy could occur through the same eQTLs affecting the expression of multiple genes, or the effects of a gene on the disease independently of gene expression levels (e.g. effects through alterations in subcellular localisation and post-translational modifications) (252).

Colocalisation analysis also has limitations. Coloc method assumes that only one single causal variant exists per locus for any given trait, which could limit the accuracy of the analysis when there are multiple causal variants in the genetic region (268). Some methods, such as SuSiE (268), have been developed to address the issue of multiple variants. However, these methods typically require LD information and are highly sensitive to LD misspecifications. Colocalisation analysis can also be constrained by statistical power. For example, colocalisation analysis of PDE5A expression in aorta and CAD gave a moderate posterior probability, suggesting a low power to demonstrate colocalisation for the 2 traits. eQTL data from a cohort with a larger sample size will be beneficial to address this issue.

4.5 Conclusions

This study provides valuable insights into the genetic basis of disease modulation by antihypertensive drugs through the expression of specific target genes. Findings such as the protective effect of GUCY1A3 expression against CAD and the risk elevation associated with PDE5A expression underscore the intricate relationships between genetic regulation, medication effects, and disease outcomes. These results highlight the importance of integrating genetic and pharmacological data to tailor therapies that are not only effective but also minimize adverse effects based on individual genetic backgrounds.

Expanding this research to include more diverse populations and exploring the interactions between multiple genes and environmental factors will enhance the predictive power and clinical utility of genetic studies in hypertension management. Moreover, incorporating advanced statistical models that can handle multiple causal variants and complex linkage disequilibrium structures will improve the accuracy and relevance of findings. Overall, this study paves the way for more personalised and effective interventions in cardiovascular, diabetic, and renal disease management based on genetic insights.

Chapter 5 Overall Discussion

The evidence presented in this thesis utilises large-scale GWAS studies as source data and robust and established methodologies that are appropriately applied to investigate the causal associations between antihypertensive drugs and health outcomes. The consistency of MR estimates across various methods enhances confidence in the results, contributing to the credibility of genetic approaches in pharmacological research. The SMR method applied in the studies to identify if the effect of a genetic variant on a phenotype is mediated by gene expression provides insights into the biological pathways involved in drug actions and potential targets for therapeutic interventions. Causal variants in *GUCY1A3* and *PDE5A* suggest that these genes play a central role in BP regulation across different tissues, indicating that modulation of these genes' expression could influence cardiovascular outcomes.

Sidenafil is the first PDE5 inhibitor that was introduced in 1998 as a treatment for erectile dysfunction. Initially, this drug was developed to treat angina, however, due to its short half-life and interaction with nitrates, this drug was no longer developed for that treatment. Instead, it was repurposed for the treatment of erectile dysfunction treatment because its' common side effect is the induction of penile erection (327). Findings from this thesis highlight the beneficial effects of PDE5 inhibitors on mitigating the CAD risk, suggesting *PDE5A* remains a potential target for CAD management.

The large sample sizes of the GWAS datasets and the focus on populations with substantial European ancestry ensure sufficient statistical power and relevance to the target population. Leveraging MR to investigate the long-term effects and potential adverse outcomes of antihypertensive drugs offers a novel perspective compared to traditional RCTs and observational studies. This approach can identify potential benefits and adverse effects not detected in RCTs, contributing to a more holistic understanding of these drugs. Furthermore, MR's ability to infer lifelong impacts and detect rare adverse effects complements existing RCT evidence, providing a broader evaluation of antihypertensive drugs. The study uncovers previously unexplored effects of antihypertensive drugs on CKD, AF, HF and stroke. The integration of GWAS and eQTL data in SMR provides a clearer understanding of the biological pathways through which BP-lowering drugs exert their effects, offering a more comprehensive view than traditional GWAS alone. The identification of genes like *GUCY1A3* and *PDE5A* as significantly associated with SBP and CAD risk has direct implications for drug development. GUCY1A3's protective role suggests it could be a target for novel antihypertensive therapies that also reduce CAD risk. Conversely, *PDE5A*'s association with increased risk highlights potential adverse effects, cautioning against its use in patients at high risk for CAD. These insights can guide the design of targeted therapies, personalised medicine

approaches, and the repurposing of existing drugs, potentially improving efficacy and safety profiles for cardiovascular treatments. These findings challenge the traditional focus on blood pressure reduction as the primary outcome of antihypertensive therapy, suggesting broader health impacts that need consideration. The notion of the broader impact of antihypertensive drugs has been explored in several previous studies including large studies from Suchard et al. (380) and Law et al. (381) have explored the broader health impacts of antihypertensive drugs, finding similar potential benefits beyond BP reduction, particularly concerning diabetes and kidney function.

The genetic associations found in the study provide important insights, but their translation into clinical practice requires careful consideration. The efficacy and safety of applying these genetic findings in real-world settings need thorough evaluation. While MR can infer causality better than observational studies, it is not free from limitations such as pleiotropy and population-specific effects. It is also pertinent to point out that while the ability of MR to provide insights into lifelong impacts and rare adverse effects is well-supported, MR studies are observational by nature, and though they can address some limitations of RCTs, they cannot completely replace the need for well-conducted RCTs. Despite sensitivity analyses, there might still be residual pleiotropy affecting the MR estimates as genetic variants used as instruments could influence outcomes through pathways other than BP reduction. It is important to differentiate between the efficacy demonstrated in genetic studies and the effectiveness in real-world settings as their real-world effectiveness might differ due to factors like adherence, comorbidities, and healthcare access. Thus, while the findings from this thesis are consistent with previous studies strengthening the case for drug repurposing or repositioning, these conclusions should be approached with caution until further validated by pragmatic clinical trials specifically designed to assess these effects.

The conclusions drawn from this thesis are specific to populations of European ancestry due to the GWAS datasets used. Genetic architectures and drug response profiles can vary significantly across populations. Thus, the findings may not be generalisable to other ethnic groups. Future studies should include more diverse populations to validate the findings and ensure broader applicability.

Clinical implications

Current guidelines for hypertension treatment typically recommend an empirical approach, which focuses on BP reduction as the primary outcome. This approach may not account for the heterogeneity in the impact of antihypertensive drug classes on different health outcomes and the complexity of hypertension pathophysiology. My thesis not only corroborates the clinical effects of existing drug classes on the outcomes but also discovers new therapeutic targets, such as *GUCY1A3* and *PDE5A*, for further

investigation. Further research should focus on (i) validating these results across diverse populations, (ii) triangulating the findings with evidence from clinical trials, and (iii) conducting functional studies to elucidate the biological mechanisms. These drug targets if validated can help to guide the design of targeted therapies, personalised medicine approaches, and the repurposing of existing drugs. A similar method framework can be applied to study the impact of antihypertensive drugs on broader health outcomes, which can help identify off-target effects of the drugs. If validated, it can inform regulatory agencies about potential adverse effects, thereby providing updated prescribing and monitoring guidelines to physicians.

Future work

There are several opportunities for further work based on this thesis. One option is to update the summary statistics data for systolic blood pressure with a recently published GWAS (382) on blood pressure traits with a larger sample size, which may strengthen the power of the study as well as provide genetic instruments for antihypertensive drug classes that lacked genetic instruments in the current research. Other studies can be categorised into the following themes.

Diverse Population Studies: As my thesis focused primarily on the European population, future research can focus on other ancestries such as Asian and African to validate the findings and ensure their applicability across different ethnic groups. This needs large GWAS consortia on those ethnic groups, and the same MR approach can be applied to the research.

Clinical Trials: There are some aspects that future clinical trials can focus on: (i) Biomarker validation study: to test whether the identified genes (*GUCY1A3*, *PDE5A*) can serve as reliable biomarkers for predicting the risk of CAD; (ii) Genotype-phenotype association study: to investigate the association between specific genetic variants in the identified genes and clinical phenotypes; (iii) Drug repurposing trial: to evaluate the repurposing potential of specific drug classes (PDE5i, sGC stimulators) for the management of disease (CAD).

Functional Studies: Further research directions include (i) gene knockout/knockdown studies (e.g. knock out or knock down the expression of the genes in cell lines or animal models to assess the effects of genes on the disease phenotype); (ii) overexpression studies (e.g. overexpressing genes in relevant cell lines or animal models to examine changes in the phenotype); (iii) in vivo studies (e.g. developing transgenic or knockout mouse models to explore the physiological and pathological functions of these genes within an entire organism).

Extended Phenotype Analysis: The integrative approach of MR, SMR, and coloc can be applied to explore the effects of antihypertensive drugs on additional phenotypes, such as

cognitive function and metabolic health, to further understand their broader impacts and potential repurposing opportunities. One-sample MR with individual data or observational studies can also be performed, enabling the triangulation of evidence.

Development of Genetic Screening Tools: Once the roles of detected genes (GUCY1A3, PDE5A) in mitigating the risk of CVDs have been validated, future research on developing genetic screening tools can be conducted to identify individuals (e.g. those with a specific genotype) who would benefit most from the drugs (e.g. sGC stimulators, PDE5 inhibitors). This will help enhance personalised medicine and improve treatment outcomes.

Overall, this study established a foundation for future developments of personalized and effective therapy for cardiovascular, diabetic, and renal diseases based on genetic insights.

Chapter 6 Final Conclusions

This thesis makes a significant contribution to the field by offering new insights and perspectives through the use of Mendelian randomisation and the integration of GWAS and eQTL data using the SMR method providing a more detailed understanding of the genetic mechanisms underlying BP regulation and cardiovascular outcomes. The findings provide strong evidence for the effectiveness of antihypertensive drugs in reducing the risk of various cardiovascular outcomes. The identification of causal genes and the broader implications for drug development represent substantial advancements in precision medicine. The findings suggest that some antihypertensive drugs may have broader health impacts than previously understood, challenging the traditional focus on BP reduction alone. This necessitates a re-evaluation of current therapeutic strategies and encourages further exploration of the multifunctional benefits and risks of BP-lowering drugs. By employing a novel MR approach combined with extensive genetic data integration, offering robust and comprehensive insights into the long-term and broader health impacts of antihypertensive drugs, it confirms findings from similar studies regarding the benefits of BP reduction on cardiovascular risk and expands on these by exploring potential drug repurposing opportunities and additional health outcomes. It advances existing knowledge by addressing the limitations of RCTs, integrating genetic data, and providing a comprehensive evaluation of the broader health impacts of antihypertensive drugs. Overall, the results from this thesis support the need for further research to validate and functionally characterise these findings followed by clinical effectiveness trials for clinical translation.

Chapter 7 List of References

1. Cannon WB. Organization for Physiological Homeostasis. Physiological Reviews. 1929;9(3):399-431.
2. Benetos A, Petrovic M, Strandberg T. Hypertension Management in Older and Frail Older Patients. Circ Res. 2019;124(7):1045-60.
3. Franklin SS, Larson MG, Khan SA, Wong ND, Leip EP, Kannel WB, et al. Does the relation of blood pressure to coronary heart disease risk change with aging? The Framingham Heart Study. Circulation. 2001;103(9):1245-9.
4. Franklin SS, Gustin Wt, Wong ND, Larson MG, Weber MA, Kannel WB, et al. Hemodynamic patterns of age-related changes in blood pressure. The Framingham Heart Study. Circulation. 1997;96(1):308-15.
5. Lewington S, Clarke R, Qizilbash N, Peto R, Collins R, Prospective Studies C. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. Lancet. 2002;360(9349):1903-13.
6. Evans JG, Rose G. Hypertension. Br Med Bull. 1971;27(1):37-42.
7. Williams B, Mancia G, Spiering W, Agabiti Rosei E, Azizi M, Burnier M, et al. 2018 ESC/ESH Guidelines for the management of arterial hypertension. Eur Heart J. 2018;39(33):3021-104.
8. NICE guideline [NG136]. Hypertension in adults: diagnosis and management 28 August 2019 [updated 21 November 2023. Available from: https://www.nice.org.uk/guidance/ng136.
9. Padmanabhan S, Aman A, Dominiczak AF. Genomics of hypertension. Pharmacol Res. 2017;121:219-29.
10. Padmanabhan S, Dominiczak AF. Genomics of hypertension: the road to precision medicine. Nat Rev Cardiol. 2021;18(4):235-50.
11. Frohlich ED, Dustan HP, Bumpus FM. Irvine H. Page: 1901-1991. The celebration of a leader. Hypertension. 1991;18(4):443-5.
12. Carvalho JJ, Baruzzi RG, Howard PF, Poulter N, Alpers MP, Franco LJ, et al. Blood pressure in four remote populations in the INTERSALT Study. Hypertension. 1989;14(3):238-46.
13. Bursztyn M. Occupational and environmental influences on hypertension. J Hum Hypertens. 2020;34(3):202-6.
14. Forouzanfar MH, Afshin A, Alexander LT, Anderson HR, Bhutta ZA, Biryukov S, et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. The Lancet. 2016;388(10053):1659-724.
15. Forouzanfar MH, Liu P, Roth GA, Ng M, Biryukov S, Marczak L, et al. Global Burden of Hypertension and Systolic Blood Pressure of at Least 110 to 115 mm Hg, 1990-2015. Jama. 2017;317(2).
16. Lawes CM, Rodgers A, Bennett DA, Parag V, Suh I, Ueshima H, et al. Blood pressure and cardiovascular disease in the Asia Pacific region. J Hypertens. 2003;21(4):707-16.
17. Rapsomaniki E, Timmis A, George J, Pujades-Rodriguez M, Shah AD, Denaxas S, et al. Blood pressure and incidence of twelve cardiovascular diseases: lifetime risks, healthy life-years lost, and age-specific associations in 1.25 million people. Lancet. 2014;383(9932):1899-911.
18. Flint AC, Conell C, Ren X, Banki NM, Chan SL, Rao VA, et al. Effect of Systolic and Diastolic Blood Pressure on Cardiovascular Outcomes. N Engl J Med. 2019;381(3):243-51.
19. Mancia G, Kreutz R, Brunstrom M, Burnier M, Grassi G, Januszewicz A, et al. 2023 ESH Guidelines for the management of arterial hypertension The Task Force for the management of arterial hypertension of the European Society of Hypertension: Endorsed by the International Society of Hypertension (ISH) and the European Renal Association (ERA). J Hypertens. 2023;41(12):1874-2071.
20. Muntner P, Shimbo D, Carey RM, Charleston JB, Gaillard T, Misra S, et al. Measurement of Blood Pressure in Humans: A Scientific Statement From the American Heart Association. Hypertension. 2019;73(5):e35-e66.
21. Benmira AM, Perez-Martin A, Coudray S, Schuster I, Aichoun I, Laurent J, et al. Systolic peak foot-to-apex time interval, a novel oscillometric technique for systolic blood pressure measurement. J Hypertens. 2017;35(5):1002-10.

22. Staplin N, de la Sierra A, Ruilope LM, Emberson JR, Vinyoles E, Gorostidi M, et al. Relationship between clinic and ambulatory blood pressure and mortality: an observational cohort study in 59 124 patients. Lancet. 2023;401(10393):2041-50.
23. Mancia G, Verdecchia P. Clinical value of ambulatory blood pressure: evidence and limits. Circ Res. 2015;116(6):1034-45.
24. Whelton PK, Carey RM, Aronow WS, Casey DE, Jr., Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Hypertension. 2018;71(6):e13-e115.
25. Ntineri A, Niiranen TJ, McManus RJ, Lindroos A, Jula A, Schwartz C, et al. Ambulatory versus home blood pressure monitoring: frequency and determinants of blood pressure difference and diagnostic disagreement. J Hypertens. 2019;37(10):1974-81.
26. Parati G, Stergiou GS, Bilo G, Kollias A, Pengo M, Ochoa JE, et al. Home blood pressure monitoring: methodology, clinical relevance and practical application: a 2021 position paper by the Working Group on Blood Pressure Monitoring and Cardiovascular Variability of the European Society of Hypertension. J Hypertens. 2021;39(9):1742-67.
27. Niiranen TJ, Asayama K, Thijs L, Johansson JK, Ohkubo T, Kikuya M, et al. Outcome-driven thresholds for home blood pressure measurement: international database of home blood pressure in relation to cardiovascular outcome. Hypertension. 2013;61(1):27-34.
28. Kario K, Hasebe N, Okumura K, Yamashita T, Akao M, Atarashi H, et al. Home Blood Pressure Can Predict the Risk for Stroke/Bleeding Events in Elderly Patients With Nonvalvular Atrial Fibrillation From the ANAFIE Registry. Hypertension. 2022;79(12):2696-705.
29. Omboni S, McManus RJ, Bosworth HB, Chappell LC, Green BB, Kario K, et al. Evidence and Recommendations on the Use of Telemedicine for the Management of Arterial Hypertension: An International Expert Position Paper. Hypertension. 2020;76(5):1368-83.
30. Ettehad D, Emdin CA, Kiran A, Anderson SG, Callender T, Emberson J, et al. Blood pressure lowering for prevention of cardiovascular disease and death: a systematic review and meta-analysis. Lancet. 2016;387(10022):957-67.
31. Thomopoulos C, Parati G, Zanchetti A. Effects of blood pressure lowering on outcome incidence in hypertension. 1. Overview, meta-analyses, and meta-regression analyses of randomized trials. J Hypertens. 2014;32(12):2285-95.
32. Blood Pressure Lowering Treatment Trialists C. Pharmacological blood pressure lowering for primary and secondary prevention of cardiovascular disease across different levels of blood pressure: an individual participant-level data meta-analysis. Lancet. 2021;397(10285):1625-36.
33. Appel LJ. Lifestyle modification as a means to prevent and treat high blood pressure. J Am Soc Nephrol. 2003;14(7 Suppl 2):S99-S102.
34. Parving HH, Brenner BM, McMurray JJ, de Zeeuw D, Haffner SM, Solomon SD, et al. Cardiorenal end points in a trial of aliskiren for type 2 diabetes. N Engl J Med. 2012;367(23):2204-13.
35. Makani H, Bangalore S, Desouza KA, Shah A, Messerli FH. Efficacy and safety of dual blockade of the renin-angiotensin system: meta-analysis of randomised trials. BMJ. 2013;346:f360.
36. James M. Ritter, Rod J. Flower, Graeme Henderson, Yoon Kong Loke, David MacEwan, Emma Robinson, et al. Rang & Dale's Pharmacology, 10th Edition: Elsevier; 2023.
37. van Vark LC, Bertrand M, Akkerhuis KM, Brugts JJ, Fox K, Mourad JJ, et al. Angiotensin-converting enzyme inhibitors reduce mortality in hypertension: a meta-analysis of randomized clinical trials of renin-angiotensin-aldosterone system inhibitors involving 158,998 patients. Eur Heart J. 2012;33(16):2088-97.
38. Savarese G, Costanzo P, Cleland JG, Vassallo E, Ruggiero D, Rosano G, et al. A meta-analysis reporting effects of angiotensin-converting enzyme inhibitors and angiotensin receptor blockers in patients without heart failure. J Am Coll Cardiol. 2013;61(2):131-42.
39. Burnett H, Earley A, Voors AA, Senni M, McMurray JJ, Deschaseaux C, et al. Thirty Years of Evidence on the Efficacy of Drug Treatments for Chronic Heart Failure With Reduced Ejection Fraction: A Network Meta-Analysis. Circ Heart Fail. 2017;10(1):e003529.
40. Thomopoulos C, Parati G, Zanchetti A. Effects of blood pressure lowering on outcome incidence in hypertension: 4. Effects of various classes of antihypertensive drugs--overview and meta-analyses. J Hypertens. 2015;33(2):195-211.

41. Thomopoulos C, Parati G, Zanchetti A. Effects of blood pressure-lowering on outcome incidence in hypertension: 5. Head-to-head comparisons of various classes of antihypertensive drugs - overview and meta-analyses. J Hypertens. 2015;33(7):1321-41.
42. Investigators DT, Bosch J, Yusuf S, Gerstein HC, Pogue J, Sheridan P, et al. Effect of ramipril on the incidence of diabetes. N Engl J Med. 2006;355(15):1551-62.
43. Reboldi G, Angeli F, Cavallini C, Gentile G, Mancia G, Verdecchia P. Comparison between angiotensin-converting enzyme inhibitors and angiotensin receptor blockers on the risk of myocardial infarction, stroke and death: a meta-analysis. J Hypertens. 2008;26(7):1282-9.
44. Gallo G, Volpe M, Rubattu S. Angiotensin Receptor Blockers in the Management of Hypertension: A Real-World Perspective and Current Recommendations. Vasc Health Risk Manag. 2022;18:507-15.
45. Corrao G, Zambon A, Parodi A, Poluzzi E, Baldi I, Merlino L, et al. Discontinuation of and changes in drug therapy for hypertension among newly-treated patients: a population-based study in Italy. J Hypertens. 2008;26(4):819-24.
46. Thomopoulos C, Parati G, Zanchetti A. Effects of blood-pressure-lowering treatment on outcome incidence in hypertension: 10 - Should blood pressure management differ in hypertensive patients with and without diabetes mellitus? Overview and meta-analyses of randomized trials. J Hypertens. 2017;35(5):922-44.
47. Schneider MP, Hua TA, Bohm M, Wachtell K, Kjeldsen SE, Schmieder RE. Prevention of atrial fibrillation by Renin-Angiotensin system inhibition a meta-analysis. J Am Coll Cardiol. 2010;55(21):2299-307.
48. Ducharme A, Swedberg K, Pfeffer MA, Cohen-Solal A, Granger CB, Maggioni AP, et al. Prevention of atrial fibrillation in patients with symptomatic chronic heart failure by candesartan in the Candesartan in Heart failure: Assessment of Reduction in Mortality and morbidity (CHARM) program. Am Heart J. 2006;152(1):86-92.
49. Investigators AI, Yusuf S, Healey JS, Pogue J, Chrolavicius S, Flather M, et al. Irbesartan in patients with atrial fibrillation. N Engl J Med. 2011;364(10):928-38.
50. Goette A, Schon N, Kirchhof P, Breithardt G, Fetsch T, Hausler KG, et al. Angiotensin II-antagonist in paroxysmal atrial fibrillation (ANTIPAF) trial. Circ Arrhythm Electrophysiol. 2012;5(1):43-51.
51. Jensen C, Herold P, Brunner HR. Aliskiren: the first renin inhibitor for clinical treatment. Nat Rev Drug Discov. 2008;7(5):399-410.
52. Sun Q, Sever P. Amiloride: A review. J Renin Angiotensin Aldosterone Syst. 2020;21(4):1470320320975893.
53. Chen C, Zhu XY, Li D, Lin Q, Zhou K. Clinical efficacy and safety of spironolactone in patients with resistant hypertension: A systematic review and meta-analysis. Medicine (Baltimore). 2020;99(34):e21694.
54. Brown MJ, Williams B, Morant SV, Webb DJ, Caulfield MJ, Cruickshank JK, et al. Effect of amiloride, or amiloride plus hydrochlorothiazide, versus hydrochlorothiazide on glucose tolerance and blood pressure (PATHWAY-3): a parallel-group, double-blind randomised phase 4 trial. Lancet Diabetes Endocrinol. 2016;4(2):136-47.
55. Zillich AJ, Garg J, Basu S, Bakris GL, Carter BL. Thiazide diuretics, potassium, and the development of diabetes: a quantitative review. Hypertension. 2006;48(2):219-24.
56. Nazarzadeh M, Bidel Z, Canoy D, Copland E, Wamil M, Majert J, et al. Blood pressure lowering and risk of new-onset type 2 diabetes: an individual participant data meta-analysis. Lancet. 2021;398(10313):1803-10.
57. Vidt DG. Alpha-blockers and congestive heart failure: early termination of an arm of the ALLHAT trial. Cleve Clin J Med. 2000;67(6):429-33.
58. Williams B, MacDonald TM, Morant S, Webb DJ, Sever P, McInnes G, et al. Spironolactone versus placebo, bisoprolol, and doxazosin to determine the optimal treatment for drug-resistant hypertension (PATHWAY-2): a randomised, double-blind, crossover trial. Lancet. 2015;386(10008):2059-68.
59. Hu B, Zhong L, Weng Y, Peng L, Huang Y, Zhao Y, et al. Therapeutic siRNA: state of the art. Signal Transduct Target Ther. 2020;5(1):101.
60. Desai AS, Webb DJ, Taubel J, Casey S, Cheng Y, Robbie GJ, et al. Zilebesiran, an RNA Interference Therapeutic Agent for Hypertension. New England Journal of Medicine. 2023;389(3):228-38.
61. Cohen JB, Bress AP. Entering a New Era of Antihypertensive Therapy. Am J Kidney Dis. 2024;83(3):411-4.

62. Cruz-Lopez EO, Ye D, Wu C, Lu HS, Uijl E, Mirabito Colafella KM, et al. Angiotensinogen Suppression: A New Tool to Treat Cardiovascular and Renal Disease. Hypertension. 2022;79(10):2115-26.
63. Uijl E, Mirabito Colafella KM, Sun Y, Ren L, van Veghel R, Garrelds IM, et al. Strong and Sustained Antihypertensive Effect of Small Interfering RNA Targeting Liver Angiotensinogen. Hypertension. 2019;73(6):1249-57.
64. Bakris GL, Saxena M, Gupta A, Chalhoub F, Lee J, Stiglitz D, et al. RNA Interference With Zilebesiran for Mild to Moderate Hypertension: The KARDIA-1 Randomized Clinical Trial. JAMA. 2024;331(9):740-9.
65. Funder JW. Mineralocorticoid receptor antagonists: emerging roles in cardiovascular medicine. Integr Blood Press Control. 2013;6:129-38.
66. Nishiyama A. Pathophysiological mechanisms of mineralocorticoid receptor-dependent cardiovascular and chronic kidney disease. Hypertens Res. 2019;42(3):293-300.
67. Ito S, Itoh H, Rakugi H, Okuda Y, Yamakawa S. Efficacy and safety of esaxerenone (CS-3150) for the treatment of essential hypertension: a phase 2 randomized, placebo-controlled, double-blind study. J Hum Hypertens. 2019;33(7):542-51.
68. Takahashi M, Ubukata O, Homma T, Asoh Y, Honzumi M, Hayashi N, et al. Crystal structure of the mineralocorticoid receptor ligand-binding domain in complex with a potent and selective nonsteroidal blocker, esaxerenone (CS-3150). FEBS Lett. 2020;594(10):1615-23.
69. Ito S, Itoh H, Rakugi H, Okuda Y, Yoshimura M, Yamakawa S. Double-Blind Randomized Phase 3 Study Comparing Esaxerenone (CS-3150) and Eplerenone in Patients With Essential Hypertension (ESAX-HTN Study). Hypertension. 2020;75(1):51-8.
70. Wan N, Rahman A, Nishiyama A. Esaxerenone, a novel nonsteroidal mineralocorticoid receptor blocker (MRB) in hypertension and chronic kidney disease. J Hum Hypertens. 2021;35(2):148-56.
71. Okuda Y, Ito S, Kashihara N, Shikata K, Nangaku M, Wada T, et al. The renoprotective effect of esaxerenone independent of blood pressure lowering: a post hoc mediation analysis of the ESAX-DN trial. Hypertens Res. 2023;46(2):437-44.
72. Hoshide S. Is esaxerenone the ultimate mineralocorticoid receptor antagonist? Hypertens Res. 2023;46(2):516-7.
73. Hermidorff MM, de Assis LV, Isoldi MC. Genomic and rapid effects of aldosterone: what we know and do not know thus far. Heart Fail Rev. 2017;22(1):65-89.
74. Laffin LJ, Rodman D, Luther JM, Vaidya A, Weir MR, Rajicic N, et al. Aldosterone Synthase Inhibition With Lorundrostat for Uncontrolled Hypertension: The Target-HTN Randomized Clinical Trial. JAMA. 2023;330(12):1140-50.
75. Calhoun DA, White WB, Krum H, Guo W, Bermann G, Trapani A, et al. Effects of a novel aldosterone synthase inhibitor for treatment of primary hypertension: results of a randomized, double-blind, placebo- and active-controlled phase 2 trial. Circulation. 2011;124(18):1945-55.
76. Bomback AS, Klemmer PJ. The incidence and implications of aldosterone breakthrough. Nat Clin Pract Nephrol. 2007;3(9):486-92.
77. Weldon SM, Brown NF. Inhibitors of Aldosterone Synthase. Vitam Horm. 2019;109:211-39.
78. Gomez-Sanchez CE, Qi X, Velarde-Miranda C, Plonczynski MW, Parker CR, Rainey W, et al. Development of monoclonal antibodies against human CYP11B1 and CYP11B2. Mol Cell Endocrinol. 2014;383(1-2):111-7.
79. Hargovan M, Ferro A. Aldosterone synthase inhibitors in hypertension: current status and future possibilities. JRSM Cardiovasc Dis. 2014;3:2048004014522440.
80. Freeman MW, Bond M, Murphy B, Hui J, Isaacsohn J. Results from a phase 1, randomized, double-blind, multiple ascending dose study characterizing the pharmacokinetics and demonstrating the safety and selectivity of the aldosterone synthase inhibitor baxdrostat in healthy volunteers. Hypertens Res. 2023;46(1):108-18.
81. Freeman MW, Halvorsen YD, Marshall W, Pater M, Isaacsohn J, Pearce C, et al. Phase 2 Trial of Baxdrostat for Treatment-Resistant Hypertension. N Engl J Med. 2023;388(5):395-405.
82. Follmann M, Griebenow N, Hahn MG, Hartung I, Mais FJ, Mittendorf J, et al. The chemistry and biology of soluble guanylate cyclase stimulators and activators. Angew Chem Int Ed Engl. 2013;52(36):9442-62.
83. Vakalopoulos A, Wunder F, Hartung IV, Redlich G, Jautelat R, Buchgraber P, et al. New Generation of sGC Stimulators: Discovery of Imidazo[1,2-a]pyridine Carboxamide BAY

1165747 (BAY-747), a Long-Acting Soluble Guanylate Cyclase Stimulator for the Treatment of Resistant Hypertension. J Med Chem. 2023;66(11):7280-303.
84. Sauzeau V, Le Jeune H, Cario-Toumaniantz C, Smolenski A, Lohmann SM, Bertoglio J, et al. Cyclic GMP-dependent protein kinase signaling pathway inhibits RhoA-induced Ca2+ sensitization of contraction in vascular smooth muscle. J Biol Chem. 2000;275(28):21722-9.
85. Packer CS. Soluble guanylate cyclase (sGC) down-regulation by abnormal extracellular matrix proteins as a novel mechanism in vascular dysfunction: Implications in metabolic syndrome. Cardiovasc Res. 2006;69(2):302-3.
86. Stasch JP, Hobbs AJ. NO-independent, haem-dependent soluble guanylate cyclase stimulators. Handb Exp Pharmacol. 2009(191):277-308.
87. Ghofrani HA, Galie N, Grimminger F, Grunig E, Humbert M, Jing ZC, et al. Riociguat for the treatment of pulmonary arterial hypertension. N Engl J Med. 2013;369(4):330-40.
88. Azizi M, Rossignol P, Hulot JS. Emerging Drug Classes and Their Potential Use in Hypertension. Hypertension. 2019;74(5):1075-83.
89. Armstrong PW, Pieske B, Anstrom KJ, Ezekowitz J, Hernandez AF, Butler J, et al. Vericiguat in Patients with Heart Failure and Reduced Ejection Fraction. N Engl J Med. 2020;382(20):1883-93.
90. Udelson JE, Lewis GD, Shah SJ, Zile MR, Redfield MM, Burnett J, Jr., et al. Effect of Praliciguat on Peak Rate of Oxygen Consumption in Patients With Heart Failure With Preserved Ejection Fraction: The CAPACITY HFpEF Randomized Clinical Trial. JAMA. 2020;324(15):1522-31.
91. Tchernychev B, Li H, Lee SK, Gao X, Ramanarasimhaiah R, Liu G, et al. Olinciguat, a stimulator of soluble guanylyl cyclase, attenuates inflammation, vaso-occlusion and nephropathy in mouse models of sickle cell disease. Br J Pharmacol. 2021;178(17):3463-75.
92. Wunder F, Stasch JP, Knorr A, Mondritzki T, Brockschnieder D, Becker-Pelster EM, et al. Identification and characterization of the new generation soluble guanylate cyclase stimulator BAY-747 designed for the treatment of resistant hypertension. Br J Pharmacol. 2023;180(19):2500-13.
93. Abraham GR, Williams TL, Maguire JJ, Greasley PJ, Ambery P, Davenport AP. Current and future strategies for targeting the endothelin pathway in cardiovascular disease. Nature Cardiovascular Research. 2023;2(11):972-90.
94. Haynes WG, Webb DJ. Contribution of endogenous generation of endothelin-1 to basal vascular tone. Lancet. 1994;344(8926):852-4.
95. Chester AH, Yacoub MH. The role of endothelin-1 in pulmonary arterial hypertension. Glob Cardiol Sci Pract. 2014;2014(2):62-78.
96. Davenport AP, Hyndman KA, Dhaun N, Southan C, Kohan DE, Pollock JS, et al. Endothelin. Pharmacol Rev. 2016;68(2):357-418.
97. Anand I, McMurray J, Cohn JN, Konstam MA, Notter T, Quitzau K, et al. Long-term effects of darusentan on left-ventricular remodelling and clinical outcomes in the EndothelinA Receptor Antagonist Trial in Heart Failure (EARTH): randomised, double-blind, placebo-controlled trial. Lancet. 2004;364(9431):347-54.
98. Veenit V, Heerspink HJL, Ahlstrom C, Greasley PJ, Skritic S, van Zuydam N, et al. The sodium glucose co-transporter 2 inhibitor dapagliflozin ameliorates the fluid-retaining effect of the endothelin A receptor antagonist zibotentan. Nephrol Dial Transplant. 2023;38(10):2289-97.
99. Rubin LJ, Badesch DB, Barst RJ, Galie N, Black CM, Keogh A, et al. Bosentan therapy for pulmonary arterial hypertension. N Engl J Med. 2002;346(12):896-903.
100. Pulido T, Adzerikho I, Channick RN, Delcroix M, Galie N, Ghofrani HA, et al. Macitentan and morbidity and mortality in pulmonary arterial hypertension. N Engl J Med. 2013;369(9):809-18.
101. Galie N, Olschewski H, Oudiz RJ, Torres F, Frost A, Ghofrani HA, et al. Ambrisentan for the treatment of pulmonary arterial hypertension: results of the ambrisentan in pulmonary arterial hypertension, randomized, double-blind, placebo-controlled, multicenter, efficacy (ARIES) study 1 and 2. Circulation. 2008;117(23):3010-9.
102. Schlaich MP, Bellet M, Weber MA, Danaietash P, Bakris GL, Flack JM, et al. Dual endothelin antagonist aprocitentan for resistant hypertension (PRECISION): a multicentre, blinded, randomised, parallel-group, phase 3 trial. Lancet. 2022;400(10367):1927-37.
103. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, et al. 10 Years of GWAS Discovery: Biology, Function, and Translation. Am J Hum Genet. 2017;101(1):5-22.
104. Tam V, Patel N, Turcotte M, Bosse Y, Pare G, Meyre D. Benefits and limitations of genome-wide association studies. Nat Rev Genet. 2019;20(8):467-84.

105. Sollis E, Mosaku A, Abid A, Buniello A, Cerezo M, Gil L, et al. The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. Nucleic Acids Res. 2023;51(D1):D977-D85.
106. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature. 2012;491(7422):119-24.
107. Sekar A, Bialas AR, de Rivera H, Davis A, Hammond TR, Kamitaki N, et al. Schizophrenia risk from complex variation of complement component 4. Nature. 2016;530(7589):177-83.
108. Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores. Nat Rev Genet. 2018;19(9):581-90.
109. Nelson MR, Tipney H, Painter JL, Shen J, Nicoletti P, Shen Y, et al. The support of human genetic evidence for approved drug indications. Nat Genet. 2015;47(8):856-60.
110. Havlik RJ, Garrison RJ, Feinleib M, Kannel WB, Castelli WP, McNamara PM. Blood pressure aggregation in families. Am J Epidemiol. 1979;110(3):304-12.
111. Kupper N, Willemsen G, Riese H, Posthuma D, Boomsma DI, de Geus EJ. Heritability of daytime ambulatory blood pressure in an extended twin design. Hypertension. 2005;45(1):80-5.
112. Niiranen TJ, McCabe EL, Larson MG, Henglin M, Lakdawala NK, Vasan RS, et al. Risk for hypertension crosses generations in the community: a multi-generational cohort study. Eur Heart J. 2017;38(29):2300-8.
113. Timpson NJ, Greenwood CMT, Soranzo N, Lawson DJ, Richards JB. Genetic architecture: the shape of the genetic contribution to human traits and disease. Nat Rev Genet. 2018;19(2):110-24.
114. Padmanabhan S, Joe B. Towards Precision Medicine for Hypertension: A Review of Genomic, Epigenomic, and Microbiomic Effects on Blood Pressure in Experimental Rat Models and Humans. Physiol Rev. 2017;97(4):1469-528.
115. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. Nature Genetics. 2018;50(10):1412-25.
116. Jhamb D, Magid-Slav M, Hurle MR, Agarwal P. Pathway analysis of GWAS loci identifies novel drug targets and repurposing opportunities. Drug Discov Today. 2019;24(6):1232-6.
117. Stasch JP, Pacher P, Evgenov OV. Soluble guanylate cyclase as an emerging therapeutic target in cardiopulmonary disease. Circulation. 2011;123(20):2263-73.
118. O'Connor CM, Starling RC, Hernandez AF, Armstrong PW, Dickstein K, Hasselblad V, et al. Effect of nesiritide in patients with acute decompensated heart failure. N Engl J Med. 2011;365(1):32-43.
119. Cowie MR, Fisher M. SGLT2 inhibitors: mechanisms of cardiovascular benefit beyond glycaemic control. Nat Rev Cardiol. 2020;17(12):761-72.
120. Chauquet S, Zhu Z, O'Donovan MC, Walters JTR, Wray NR, Shah S. Association of Antihypertensive Drug Target Genes With Psychiatric Disorders: A Mendelian Randomization Study. JAMA Psychiatry. 2021;78(6):623-31.
121. Yarmolinsky J, Diez-Obrero V, Richardson TG, Pigeyre M, Sjaarda J, Pare G, et al. Genetically proxied therapeutic inhibition of antihypertensive drug targets and risk of common cancers: A mendelian randomization analysis. PLoS Med. 2022;19(2):e1003897.
122. Gill D, Georgakis MK, Koskeridis F, Jiang L, Feng Q, Wei WQ, et al. Use of Genetic Variants Related to Antihypertensive Drugs to Inform on Efficacy and Side Effects. Circulation. 2019;140(4):270-9.
123. Wittrup A, Lieberman J. Knocking down disease: a progress report on siRNA therapeutics. Nat Rev Genet. 2015;16(9):543-52.
124. Gallagher MD, Chen-Plotkin AS. The Post-GWAS Era: From Association to Function. Am J Hum Genet. 2018;102(5):717-30.
125. Padmanabhan S, Melander O, Johnson T, Di Blasio AM, Lee WK, Gentilini D, et al. Genome-wide association study of blood pressure extremes identifies variant near UMOD associated with hypertension. PLoS Genet. 2010;6(10):e1001177.
126. US National Library of Medicine. ClinicalTrials.gov 2017 [Available from: https://clinicaltrials.gov/ct2/show/NCT03354897?term=NCT03354897.
127. Gupta RM, Hadaya J, Trehan A, Zekavat SM, Roselli C, Klarin D, et al. A Genetic Variant Associated with Five Vascular Diseases Is a Distal Regulator of Endothelin-1 Gene Expression. Cell. 2017;170(3):522-33 e15.

128. Verweij P, Danaietash P, Flamion B, Menard J, Bellet M. Randomized Dose-Response Study of the New Dual Endothelin Receptor Antagonist Aprocitentan in Hypertension. Hypertension. 2020;75(4):956-65.
129. Ren M, Ng FL, Warren HR, Witkowska K, Baron M, Jia Z, et al. The biological impact of blood pressure-associated genetic variants in the natriuretic peptide receptor C gene on human vascular smooth muscle. Hum Mol Genet. 2018;27(1):199-210.
130. Zhang R, Witkowska K, Afonso Guerra-Assuncao J, Ren M, Ng FL, Mauro C, et al. A blood pressure-associated variant of the SLC39A8 gene influences cellular cadmium accumulation and toxicity. Hum Mol Genet. 2016;25(18):4117-26.
131. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. Nature. 2009;461(7265):747-53.
132. Aschard H, Chen J, Cornelis MC, Chibnik LB, Karlson EW, Kraft P. Inclusion of gene-gene and gene-environment interactions unlikely to dramatically improve risk prediction for complex diseases. Am J Hum Genet. 2012;90(6):962-72.
133. Mancuso N, Shi H, Goddard P, Kichaev G, Gusev A, Pasaniuc B. Integrating Gene Expression with Summary Association Statistics to Identify Genes Associated with 30 Complex Traits. Am J Hum Genet. 2017;100(3):473-87.
134. Edwards SL, Beesley J, French JD, Dunning AM. Beyond GWASs: illuminating the dark road from association to function. Am J Hum Genet. 2013;93(5):779-97.
135. Claussnitzer M, Dankel SN, Kim KH, Quon G, Meuleman W, Haugen C, et al. FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. N Engl J Med. 2015;373(10):895-907.
136. Schaub MA, Boyle AP, Kundaje A, Batzoglou S, Snyder M. Linking disease associations with regulatory information in the human genome. Genome Res. 2012;22(9):1748-59.
137. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, et al. The structure of haplotype blocks in the human genome. Science. 2002;296(5576):2225-9.
138. Hemani G, Bowden J, Davey Smith G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. Hum Mol Genet. 2018;27(R2):R195-R208.
139. Wagner GP, Zhang J. The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. Nat Rev Genet. 2011;12(3):204-13.
140. Sibbald B, Roland M. Understanding controlled trials. Why are randomised controlled trials important? BMJ. 1998;316(7126):201.
141. Pantzaris ND, Karanikolas E, Tsiotsios K, Velissaris D. Renin Inhibition with Aliskiren: A Decade of Clinical Experience. J Clin Med. 2017;6(6).
142. Galie N, Hoeper MM, Gibbs JS, Simonneau G. Liver toxicity of sitaxentan in pulmonary arterial hypertension. Eur Respir J. 2011;37(2):475-6.
143. Silas JH, Lennard MS, Tucker GT, Smith AJ, Malcolm SL, Marten TR. Why hypertensive patients vary in their response to oral debrisoquine. Br Med J. 1977;1(6058):422-5.
144. Woosley RL, Nies AS. Guanethidine. N Engl J Med. 1976;295(19):1053-7.
145. Po AL, Zhang WY. What lessons can be learnt from withdrawal of mibefradil from the market? Lancet. 1998;351(9119):1829-30.
146. Strawbridge R, Javed RR, Cave J, Jauhar S, Young AH. The effects of reserpine on depression: A systematic review. J Psychopharmacol. 2023;37(3):248-60.
147. LiverTox: Clinical and Research Information on Drug-Induced Liver Injury [Internet]: Bethesda (MD): National Institute of Diabetes and Digestive and Kidney Diseases; 2012 [updated 2020 Jan 10]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK548173/.
148. Ball SG. Discontinuation of doxazosin arm of ALLHAT. Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack. Lancet. 2000;355(9214):1558.
149. Murphy MB, Murray C, Shorten GD. Fenoldopam: a selective peripheral dopamine-receptor agonist for the treatment of severe hypertension. N Engl J Med. 2001;345(21):1548-57.
150. Pickering TG. Effects of stress and behavioral interventions in hypertension: the rise and fall of omapatrilat. J Clin Hypertens (Greenwich). 2002;4(5):371-3.
151. Tikkanen I, Narko K, Zeller C, Green A, Salsali A, Broedl UC, et al. Empagliflozin reduces blood pressure in patients with type 2 diabetes and hypertension. Diabetes Care. 2015;38(3):420-8.
152. Pfeifer M, Townsend RR, Davies MJ, Vijapurkar U, Ren J. Effects of canagliflozin, a sodium glucose co-transporter 2 inhibitor, on blood pressure and markers of arterial stiffness

in patients with type 2 diabetes mellitus: a post hoc analysis. Cardiovasc Diabetol. 2017;16(1):29.
153. Wiviott SD, Raz I, Bonaca MP, Mosenzon O, Kato ET, Cahn A, et al. Dapagliflozin and Cardiovascular Outcomes in Type 2 Diabetes. N Engl J Med. 2019;380(4):347-57.
154. Omenn GS, Goodman GE, Thornquist MD, Balmes J, Cullen MR, Glass A, et al. Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular disease. N Engl J Med. 1996;334(18):1150-5.
155. Hooper L, Ness AR, Smith GD. Antioxidant strategy for cardiovascular diseases. Lancet. 2001;357(9269):1705-6.
156. Khaw KT, Bingham S, Welch A, Luben R, Wareham N, Oakes S, et al. Relation between plasma ascorbic acid and mortality in men and women in EPIC-Norfolk prospective study: a prospective population study. European Prospective Investigation into Cancer and Nutrition. Lancet. 2001;357(9257):657-63.
157. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol. 2003;32(1):1-22.
158. Molander L, Lovheim H, Norman T, Nordstrom P, Gustafson Y. Lower systolic blood pressure is associated with greater mortality in people aged 85 and older. J Am Geriatr Soc. 2008;56(10):1853-9.
159. Ravindrarajah R, Hazra NC, Hamada S, Charlton J, Jackson SHD, Dregan A, et al. Systolic Blood Pressure Trajectory, Frailty, and All-Cause Mortality >80 Years of Age: Cohort Study Using Electronic Health Records. Circulation. 2017;135(24):2357-68.
160. Feinstein AR. Multivariable analysis: an introduction: Yale University Press; 1996.
161. Robins J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. Mathematical Modelling. 1986;7(9-12):1393-512.
162. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behav Res. 2011;46(3):399-424.
163. Angrist J, Pischke J-S. Mostly Harmless Econometrics: An Empiricist's Companion. 2009.
164. Greenland S. An introduction To instrumental variables for epidemiologists. Int J Epidemiol. 2000;29(6):1102.
165. Sanderson E, Glymour MM, Holmes MV, Kang H, Morrison J, Munafo MR, et al. Mendelian randomization. Nat Rev Methods Primers. 2022;2.
166. Smith GD, Ebrahim S. Mendelian randomization: prospects, potentials, and limitations. Int J Epidemiol. 2004;33(1):30-42.
167. Labrecque JA, Swanson SA. Interpretation and Potential Biases of Mendelian Randomization Estimates With Time-Varying Exposures. Am J Epidemiol. 2019;188(1):231-8.
168. Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, et al. Guidelines for performing Mendelian randomization investigations: update for summer 2023. Wellcome Open Res. 2019;4:186.
169. Burgess S, Thompson SG, Collaboration CCG. Avoiding bias from weak instruments in Mendelian randomization studies. Int J Epidemiol. 2011;40(3):755-64.
170. Taylor AE, Davies NM, Ware JJ, VanderWeele T, Smith GD, Munafo MR. Mendelian randomization in health research: using appropriate genetic variants and avoiding biased estimates. Econ Hum Biol. 2014;13(100):99-106.
171. Burgess S, Scott RA, Timpson NJ, Davey Smith G, Thompson SG, Consortium E-I. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. Eur J Epidemiol. 2015;30(7):543-52.
172. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. Genet Epidemiol. 2016;40(7):597-608.
173. Burgess S, Thompson SG. Mendelian Randomization2021.
174. Smith GD, Lawlor DA, Harbord R, Timpson N, Day I, Ebrahim S. Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. PLoS Med. 2007;4(12):e352.
175. Lawlor DA. Commentary: Two-sample Mendelian randomization: opportunities and challenges. Int J Epidemiol. 2016;45(3):908-15.
176. Schmidt AF, Finan C, Gordillo-Maranon M, Asselbergs FW, Freitag DF, Patel RS, et al. Genetic drug target validation using Mendelian randomisation. Nat Commun. 2020;11(1):3255.

177. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med. 2008;27(8):1133-63.
178. Munafo MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. Int J Epidemiol. 2018;47(1):226-35.
179. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol. 2013;37(7):658-65.
180. Thompson JR, Minelli C, Del Greco MF. Mendelian Randomization using Public Data from Genetic Consortia. Int J Biostat. 2016;12(2).
181. Burgess S, Dudbridge F, Thompson SG. Combining information on multiple instrumental variables in Mendelian randomization: comparison of allele score and summarized data methods. Stat Med. 2016;35(11):1880-906.
182. Hartwig FP, Davies NM, Hemani G, Davey Smith G. Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. Int J Epidemiol. 2016;45(6):1717-26.
183. Bowden J, Del Greco MF, Minelli C, Davey Smith G, Sheehan N, Thompson J. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. Stat Med. 2017;36(11):1783-802.
184. Burgess S, Thompson SG. Bias in causal estimates from Mendelian randomization studies with weak instruments. Stat Med. 2011;30(11):1312-23.
185. Lawson DJ, Davies NM, Haworth S, Ashraf B, Howe L, Crawford A, et al. Is population structure in the genetic biobank era irrelevant, a challenge, or an opportunity? Hum Genet. 2020;139(1):23-41.
186. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet. 2018;50(5):693-8.
187. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. BMJ. 2018;362:k601.
188. Wardle J, Carnell S, Haworth CM, Farooqi IS, O'Rahilly S, Plomin R. Obesity associated genetic variation in FTO is associated with diminished satiety. J Clin Endocrinol Metab. 2008;93(9):3640-3.
189. Slatkin M. Linkage disequilibrium — understanding the evolutionary past and mapping the medical future. Nature Reviews Genetics. 2008;9(6):477-85.
190. Kang H, Zhang A, Cai TT, Small DS. Instrumental Variables Estimation With Some Invalid Instruments and its Application to Mendelian Randomization. Journal of the American Statistical Association. 2016;111(513):132-44.
191. Slob EAW, Burgess S. A comparison of robust Mendelian randomization methods using summary data. Genetic Epidemiology. 2020;44(4):313-29.
192. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol. 2016;40(4):304-14.
193. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. Int J Epidemiol. 2017;46(6):1985-98.
194. Rees JMB, Wood AM, Dudbridge F, Burgess S. Robust methods in Mendelian randomization via penalization of heterogeneous causal estimates. PLoS One. 2019;14(9):e0222362.
195. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol. 2015;44(2):512-25.
196. Zhao Q, Wang J, Hemani G, Bowden J, Small DS. Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score. The Annals of Statistics. 2020;48(3).
197. Qi G, Chatterjee N. Mendelian randomization analysis using mixture models for robust and efficient estimation of causal effects. Nat Commun. 2019;10(1):1941.
198. Burgess S, Foley CN, Allara E, Staley JR, Howson JMM. A robust and efficient method for Mendelian randomization with hundreds of genetic variants. Nat Commun. 2020;11(1):376.
199. Skrivankova VW, Richmond RC, Woolf BAR, Yarmolinsky J, Davies NM, Swanson SA, et al. Strengthening the Reporting of Observational Studies in Epidemiology Using Mendelian Randomization: The STROBE-MR Statement. JAMA. 2021;326(16):1614-21.

200. Haycock PC, Burgess S, Wade KH, Bowden J, Relton C, Davey Smith G. Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. Am J Clin Nutr. 2016;103(4):965-78.
201. Zacho J, Tybjaerg-Hansen A, Jensen JS, Grande P, Sillesen H, Nordestgaard BG. Genetically elevated C-reactive protein and ischemic vascular disease. N Engl J Med. 2008;359(18):1897-908.
202. Collaboration CRPCHDG, Wensley F, Gao P, Burgess S, Kaptoge S, Di Angelantonio E, et al. Association between C reactive protein and coronary heart disease: mendelian randomisation analysis based on individual participant data. BMJ. 2011;342:d548.
203. Relton CL, Davey Smith G. Is epidemiology ready for epigenetics? Int J Epidemiol. 2012;41(1):5-9.
204. Zuber V, Colijn JM, Klaver C, Burgess S. Selecting likely causal risk factors from high-throughput experiments using multivariable Mendelian randomization. Nat Commun. 2020;11(1):29.
205. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. eLife. 2018;7.
206. Burgess S, Davey Smith G. How humans can contribute to Mendelian randomization analyses. Int J Epidemiol. 2019;48(3):661-4.
207. Loewa A, Feng JJ, Hedtrich S. Human disease models in drug development. Nat Rev Bioeng. 2023:1-15.
208. Hay M, Thomas DW, Craighead JL, Economides C, Rosenthal J. Clinical development success rates for investigational drugs. Nat Biotechnol. 2014;32(1):40-51.
209. Scannell JW, Blanckley A, Boldon H, Warrington B. Diagnosing the decline in pharmaceutical R&D efficiency. Nat Rev Drug Discov. 2012;11(3):191-200.
210. Ochoa D, Karim M, Ghoussaini M, Hulcoop DG, McDonagh EM, Dunham I. Human genetics evidence supports two-thirds of the 2021 FDA-approved drugs. Nat Rev Drug Discov. 2022;21(8):551.
211. Finan C, Gaulton A, Kruger FA, Lumbers RT, Shah T, Engmann J, et al. The druggable genome and support for target identification and validation in drug development. Sci Transl Med. 2017;9(383).
212. Walker VM, Davey Smith G, Davies NM, Martin RM. Mendelian randomization: a novel approach for the prediction of adverse drug events and drug repurposing opportunities. Int J Epidemiol. 2017;46(6):2078-89.
213. Swerdlow DI, Kuchenbaecker KB, Shah S, Sofat R, Holmes MV, White J, et al. Selecting instruments for Mendelian randomization in the wake of genome-wide association studies. Int J Epidemiol. 2016;45(5):1600-16.
214. Gill D, Walker VM, Martin RM, Davies NM, Tzoulaki I. Comparison with randomized controlled trials as a strategy for evaluating instruments in Mendelian randomization. Int J Epidemiol. 2020;49(4):1404-6.
215. Abifadel M, Rabes JP, Devillers M, Munnich A, Erlich D, Junien C, et al. Mutations and polymorphisms in the proprotein convertase subtilisin kexin 9 (PCSK9) gene in cholesterol metabolism and disease. Hum Mutat. 2009;30(4):520-9.
216. Benn M, Nordestgaard BG, Grande P, Schnohr P, Tybjaerg-Hansen A. PCSK9 R46L, low-density lipoprotein cholesterol levels, and risk of ischemic heart disease: 3 independent studies and meta-analyses. J Am Coll Cardiol. 2010;55(25):2833-42.
217. Cohen JC, Boerwinkle E, Mosley TH, Jr., Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. N Engl J Med. 2006;354(12):1264-72.
218. Raal FJ, Stein EA, Dufour R, Turner T, Civeira F, Burgess L, et al. PCSK9 inhibition with evolocumab (AMG 145) in heterozygous familial hypercholesterolaemia (RUTHERFORD-2): a randomised, double-blind, placebo-controlled trial. Lancet. 2015;385(9965):331-40.
219. Robinson JG, Farnier M, Krempf M, Bergeron J, Luc G, Averna M, et al. Efficacy and safety of alirocumab in reducing lipids and cardiovascular events. N Engl J Med. 2015;372(16):1489-99.
220. Daghlas I, Gill D. Mendelian randomization as a tool to inform drug development using human genetics. Camb Prism Precis Med. 2023;1:e16.
221. Gill D, Georgakis MK, Walker VM, Schmidt AF, Gkatzionis A, Freitag DF, et al. Mendelian randomization for studying the effects of perturbing drug targets. Wellcome Open Res. 2021;6:16.
222. Santos R, Ursu O, Gaulton A, Bento AP, Donadi RS, Bologa CG, et al. A comprehensive map of molecular drug targets. Nat Rev Drug Discov. 2017;16(1):19-34.

223. Ference BA, Ray KK, Catapano AL, Ference TB, Burgess S, Neff DR, et al. Mendelian Randomization Study of ACLY and Cardiovascular Disease. N Engl J Med. 2019;380(11):1033-42.
224. Force USPST, Mangione CM, Barry MJ, Nicholson WK, Cabana M, Chelmow D, et al. Statin Use for the Primary Prevention of Cardiovascular Disease in Adults: US Preventive Services Task Force Recommendation Statement. JAMA. 2022;328(8):746-53.
225. Swerdlow DI, Preiss D, Kuchenbaecker KB, Holmes MV, Engmann JE, Shah T, et al. HMG-coenzyme A reductase inhibition, type 2 diabetes, and bodyweight: evidence from genetic analysis and randomised trials. Lancet. 2015;385(9965):351-61.
226. Denny JC, Ritchie MD, Basford MA, Pulley JM, Bastarache L, Brown-Gentry K, et al. PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. Bioinformatics. 2010;26(9):1205-10.
227. Cai T, Zhang Y, Ho YL, Link N, Sun J, Huang J, et al. Association of Interleukin 6 Receptor Variant With Cardiovascular Disease Effects of Interleukin 6 Receptor Blocking Therapy: A Phenome-Wide Association Study. JAMA Cardiol. 2018;3(9):849-57.
228. Rosa M, Chignon A, Li Z, Boulanger MC, Arsenault BJ, Bosse Y, et al. A Mendelian randomization study of IL6 signaling in cardiovascular diseases, immune-related disorders and longevity. NPJ Genom Med. 2019;4:23.
229. Lp PLASC, Thompson A, Gao P, Orfei L, Watson S, Di Angelantonio E, et al. Lipoprotein-associated phospholipase A(2) and risk of coronary disease, stroke, and mortality: collaborative analysis of 32 prospective studies. Lancet. 2010;375(9725):1536-44.
230. Millwood IY, Bennett DA, Walters RG, Clarke R, Waterworth D, Johnson T, et al. Lipoprotein-Associated Phospholipase A2 Loss-of-Function Variant and Risk of Vascular Diseases in 90,000 Chinese Adults. J Am Coll Cardiol. 2016;67(2):230-1.
231. Investigators S, White HD, Held C, Stewart R, Tarka E, Brown R, et al. Darapladib for preventing ischemic events in stable coronary heart disease. N Engl J Med. 2014;370(18):1702-11.
232. Collaboration IRGCERF, Sarwar N, Butterworth AS, Freitag DF, Gregson J, Willeit P, et al. Interleukin-6 receptor pathways in coronary heart disease: a collaborative meta-analysis of 82 studies. Lancet. 2012;379(9822):1205-13.
233. Georgakis MK, Malik R, Gill D, Franceschini N, Sudlow CLM, Dichgans M, et al. Interleukin-6 Signaling Effects on Ischemic Stroke and Other Cardiovascular Outcomes: A Mendelian Randomization Study. Circ Genom Precis Med. 2020;13(3):e002872.
234. Toshner M, Church C, Harbaum L, Rhodes C, Villar Moreschi SS, Liley J, et al. Mendelian randomisation and experimental medicine approaches to interleukin-6 as a drug target in pulmonary arterial hypertension. Eur Respir J. 2022;59(3).
235. Georgakis MK, Malik R, Richardson TG, Howson JMM, Anderson CD, Burgess S, et al. Associations of genetically predicted IL-6 signaling with cardiovascular disease risk across population subgroups. BMC Med. 2022;20(1):245.
236. Rees JMB, Foley CN, Burgess S. Factorial Mendelian randomization: using genetic variants to assess interactions. Int J Epidemiol. 2020;49(4):1147-58.
237. Krishnamurthy N, Grimshaw AA, Axson SA, Choe SH, Miller JE. Drug repurposing: a systematic review on root causes, barriers and facilitators. BMC Health Serv Res. 2022;22(1):970.
238. Hauser AS, Attwood MM, Rask-Andersen M, Schioth HB, Gloriam DE. Trends in GPCR drug discovery: new agents, targets and indications. Nat Rev Drug Discov. 2017;16(12):829-42.
239. Tang B, Wang Y, Jiang X, Thambisetty M, Ferrucci L, Johnell K, et al. Genetic Variation in Targets of Antidiabetic Drugs and Alzheimer Disease Risk: A Mendelian Randomization Study. Neurology. 2022;99(7):e650-e9.
240. Gill D, Vujkovic M. The Potential of Genetic Data for Prioritizing Drug Repurposing Efforts. Neurology. 2022;99(7):267-8.
241. Novo Nordisk. Novo Nordisk to enter phase 3 development in Alzheimer's disease with oral semaglutide 2020 [Available from: novonordisk.com/news-and-media/news-and-ir-materials/news-details.html?id=39097.
242. Higgins H, Mason AM, Larsson SC, Gill D, Langenberg C, Burgess S. Estimating the Population Benefits of Blood Pressure Lowering: A Wide-Angled Mendelian Randomization Study in UK Biobank. J Am Heart Assoc. 2021;10(17):e021098.
243. Le NN, Tran TQB, Lip S, McCallum L, McClure J, Dominiczak AF, et al. Unravelling the Distinct Effects of Systolic and Diastolic Blood Pressure Using Mendelian Randomisation. Genes (Basel). 2022;13(7).

244. Lian J, Shi X, Jia X, Fan J, Wang Y, Zhao Y, et al. Genetically predicted blood pressure, antihypertensive drugs and risk of heart failure: a Mendelian randomization study. J Hypertens. 2023;41(1):44-50.
245. Zhao JV, Liu F, Schooling CM, Li J, Gu D, Lu X. Using genetics to assess the association of commonly used antihypertensive drugs with diabetes, glycaemic traits and lipids: a trans-ancestry Mendelian randomisation study. Diabetologia. 2022;65(4):695-704.
246. Zhao JV, Schooling CM. Using Mendelian randomization study to assess the renal effects of antihypertensive drugs. BMC Med. 2021;19(1):79.
247. Walker VM, Kehoe PG, Martin RM, Davies NM. Repurposing antihypertensive drugs for the prevention of Alzheimer's disease: a Mendelian randomization study. Int J Epidemiol. 2020;49(4):1132-40.
248. Burgess S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and a binary outcome. Int J Epidemiol. 2014;43(3):922-9.
249. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet. 2016;48(5):481-7.
250. Wu Y, Zeng J, Zhang F, Zhu Z, Qi T, Zheng Z, et al. Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. Nat Commun. 2018;9(1):918.
251. Pavlides JM, Zhu Z, Gratten J, McRae AF, Wray NR, Yang J. Predicting gene targets from integrative analyses of summary data from GWAS and eQTL studies for 28 human complex traits. Genome Med. 2016;8(1):84.
252. Jacobs BM, Taylor T, Awad A, Baker D, Giovanonni G, Noyce AJ, et al. Summary-data-based Mendelian randomization prioritizes potential druggable targets for multiple sclerosis. Brain Commun. 2020;2(2):fcaa119.
253. Li Y, Sundquist K, Zhang N, Wang X, Sundquist J, Memon AA. Mitochondrial related genome-wide Mendelian randomization identifies putatively causal genes for multiple cancer types. EBioMedicine. 2023;88:104432.
254. Li Z, Zhang B, Liu Q, Tao Z, Ding L, Guo B, et al. Genetic association of lipids and lipid-lowering drug target genes with non-alcoholic fatty liver disease. EBioMedicine. 2023;90:104543.
255. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet. 2014;10(5):e1004383.
256. Guo H, Fortune MD, Burren OS, Schofield E, Todd JA, Wallace C. Integration of disease association and eQTL data using a Bayesian colocalisation approach highlights six candidate causal genes in immune-mediated diseases. Hum Mol Genet. 2015;24(12):3305-13.
257. Wen X, Pique-Regi R, Luca F. Integrating molecular QTL data into genome-wide genetic association analysis: Probabilistic assessment of enrichment and colocalization. PLoS Genet. 2017;13(3):e1006646.
258. Jaffe AE, Gao Y, Deep-Soboslay A, Tao R, Hyde TM, Weinberger DR, et al. Mapping DNA methylation across development, genotype and schizophrenia in the human frontal cortex. Nat Neurosci. 2016;19(1):40-7.
259. Foley CN, Staley JR, Breen PG, Sun BB, Kirk PDW, Burgess S, et al. A fast and efficient colocalization algorithm for identifying shared genetic risk factors across multiple traits. Nat Commun. 2021;12(1):764.
260. Zuber V, Grinberg NF, Gill D, Manipur I, Slob EAW, Patel A, et al. Combining evidence from Mendelian randomization and colocalization: Review and comparison of approaches. Am J Hum Genet. 2022;109(5):767-82.
261. Plagnol V, Smyth DJ, Todd JA, Clayton DG. Statistical independence of the colocalized association signals for type 1 diabetes and RPS26 gene expression on chromosome 12q13. Biostatistics. 2009;10(2):327-34.
262. Cotsapas C, Voight BF, Rossin E, Lage K, Neale BM, Wallace C, et al. Pervasive sharing of genetic effects in autoimmune disease. PLoS Genet. 2011;7(8):e1002254.
263. Smyth DJ, Plagnol V, Walker NM, Cooper JD, Downes K, Yang JH, et al. Shared and distinct genetic variants in type 1 diabetes and celiac disease. N Engl J Med. 2008;359(26):2767-77.
264. Zheng J, Haberland V, Baird D, Walker V, Haycock PC, Hurle MR, et al. Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. Nat Genet. 2020;52(10):1122-31.

265. Holmes MV, Richardson TG, Ference BA, Davies NM, Davey Smith G. Integrating genomics with biomarkers and therapeutic targets to invigorate cardiovascular drug development. Nat Rev Cardiol. 2021;18(6):435-53.
266. Wallace C. Statistical testing of shared genetic control for potentially related traits. Genet Epidemiol. 2013;37(8):802-13.
267. Wang G, Sarkar A, Carbonetto P, Stephens M. A simple new approach to variable selection in regression, with application to genetic fine mapping. J R Stat Soc Series B Stat Methodol. 2020;82(5):1273-300.
268. Wallace C. A more accurate method for colocalisation analysis allowing for multiple causal variants. PLoS Genet. 2021;17(9):e1009440.
269. Joint Formulary Committee. British National Formulary (online) London: BMJ and Pharmaceutical Press [Available from: http://www.medicinescomplete.com.
270. Gormley M, Yarmolinsky J, Dudding T, Burrows K, Martin RM, Thomas S, et al. Using genetic variants to evaluate the causal effect of cholesterol lowering on head and neck cancer risk: A Mendelian randomization study. PLoS Genet. 2021;17(4):e1009525.
271. Ference BA, Robinson JG, Brook RD, Catapano AL, Chapman MJ, Neff DR, et al. Variation in PCSK9 and HMGCR and Risk of Cardiovascular Disease and Diabetes. N Engl J Med. 2016;375(22):2144-53.
272. Clarke L, Zheng-Bradley X, Smith R, Kulesha E, Xiao C, Toneva I, et al. The 1000 Genomes Project: data management and community access. Nature Methods. 2012;9(5):459-62.
273. Aragam KG, Jiang T, Goel A, Kanoni S, Wolford BN, Atri DS, et al. Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. Nature Genetics. 2022;54(12):1803-15.
274. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet. 2015;47(10):1121-30.
275. Nielsen JB, Thorolfsdottir RB, Fritsche LG, Zhou W, Skov MW, Graham SE, et al. Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. Nature Genetics. 2018;50(9):1234-9.
276. Shah S, Henry A, Roselli C, Lin H, Sveinbjörnsson G, Fatemifar G, et al. Genome-wide association and Mendelian randomisation analysis provide insights into the pathogenesis of heart failure. Nature Communications. 2020;11(1).
277. Mishra A, Malik R, Hachiya T, Jurgenson T, Namba S, Posner DC, et al. Stroke genetics informs drug discovery and risk prediction across ancestries. Nature. 2022;611(7934):115-23.
278. Stanzick KJ, Li Y, Schlosser P, Gorski M, Wuttke M, Thomas LF, et al. Discovery and prioritization of variants and genes for kidney function in >1.2 million individuals. Nat Commun. 2021;12(1):4350.
279. Mahajan A, Spracklen CN, Zhang W, Ng MCY, Petty LE, Kitajima H, et al. Multi-ancestry genetic study of type 2 diabetes highlights the power of diverse populations for discovery and translation. Nat Genet. 2022;54(5):560-72.
280. Palmer TM, Lawlor DA, Harbord RM, Sheehan NA, Tobias JH, Timpson NJ, et al. Using multiple genetic variants as instrumental variables for modifiable risk factors. Statistical Methods in Medical Research. 2011;21(3):223-42.
281. Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. Int J Epidemiol. 2017;46(6):1734-9.
282. Karhunen V, Daghlas I, Zuber V, Vujkovic M, Olsen AK, Knudsen LB, et al. Leveraging human genetic data to investigate the cardiometabolic effects of glucose-dependent insulinotropic polypeptide signalling. Diabetologia. 2021;64(12):2773-8.
283. Emerging Risk Factors C, Kaptoge S, Di Angelantonio E, Lowe G, Pepys MB, Thompson SG, et al. C-reactive protein concentration and risk of coronary heart disease, stroke, and mortality: an individual participant meta-analysis. Lancet. 2010;375(9709):132-40.
284. Prins BP, Abbasi A, Wong A, Vaez A, Nolte I, Franceschini N, et al. Investigating the Causal Relationship of C-Reactive Protein with 32 Complex Somatic and Psychiatric Outcomes: A Large-Scale Cross-Consortium Mendelian Randomization Study. PLoS Med. 2016;13(6):e1001976.
285. Consortium GT. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science. 2020;369(6509):1318-30.

286. Consortium GT. The Genotype-Tissue Expression (GTEx) project. Nat Genet. 2013;45(6):580-5.
287. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature. 2015;526(7571):68-74.
288. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4:7.
289. Vosa U, Claringbould A, Westra HJ, Bonder MJ, Deelen P, Zeng B, et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. Nat Genet. 2021;53(9):1300-10.
290. Wakefield J. Bayes factors for genome-wide association studies: comparison with P-values. Genet Epidemiol. 2009;33(1):79-86.
291. Yusuf S, Joseph P, Rangarajan S, Islam S, Mente A, Hystad P, et al. Modifiable risk factors, cardiovascular disease, and mortality in 155 722 individuals from 21 high-income, middle-income, and low-income countries (PURE): a prospective cohort study. Lancet. 2020;395(10226):795-808.
292. World Health Organization. Global report on hypertension: the race against a silent killer. 2023. Report No.: ISBN: 978-92-4-008106-2.
293. Abuissa H, Jones PG, Marso SP, O'Keefe JH, Jr. Angiotensin-converting enzyme inhibitors or angiotensin receptor blockers for prevention of type 2 diabetes: a meta-analysis of randomized clinical trials. J Am Coll Cardiol. 2005;46(5):821-6.
294. Group NS, McMurray JJ, Holman RR, Haffner SM, Bethel MA, Holzhauer B, et al. Effect of valsartan on the incidence of diabetes and cardiovascular events. N Engl J Med. 2010;362(16):1477-90.
295. McFarlane SI, Kumar A, Sowers JR. Mechanisms by which angiotensin-converting enzyme inhibitors prevent diabetes and cardiovascular disease. Am J Cardiol. 2003;91(12A):30H-7H.
296. Jones NR, McCormack T, Constanti M, McManus RJ. Diagnosis and management of hypertension in adults: NICE guideline update 2019. Br J Gen Pract. 2020;70(691):90-1.
297. Frieden TR. Evidence for Health Decision Making - Beyond Randomized, Controlled Trials. N Engl J Med. 2017;377(5):465-75.
298. Tsang R, Colley L, Lynd LD. Inadequate statistical power to detect clinically significant differences in adverse event rates in randomized controlled trials. J Clin Epidemiol. 2009;62(6):609-16.
299. Hyman MC, Levin MG, Gill D, Walker VM, Georgakis MK, Davies NM, et al. Genetically Predicted Blood Pressure and Risk of Atrial Fibrillation. Hypertension. 2021;77(2):376-82.
300. Anderson AH, Yang W, Townsend RR, Pan Q, Chertow GM, Kusek JW, et al. Time-updated systolic blood pressure and the progression of chronic kidney disease: a cohort study. Ann Intern Med. 2015;162(4):258-65.
301. Emdin CA, Anderson SG, Woodward M, Rahimi K. Usual Blood Pressure and Risk of New-Onset Diabetes: Evidence From 4.1 Million Adults and a Meta-Analysis of Prospective Studies. J Am Coll Cardiol. 2015;66(14):1552-62.
302. Emdin CA, Anderson SG, Salimi-Khorshidi G, Woodward M, MacMahon S, Dwyer T, et al. Usual blood pressure, atrial fibrillation and vascular risk: evidence from 4.3 million adults. Int J Epidemiol. 2017;46(1):162-72.
303. Wuttke M, Li Y, Li M, Sieber KB, Feitosa MF, Gorski M, et al. A catalog of genetic loci associated with kidney function from analyses of a million individuals. Nat Genet. 2019;51(6):957-72.
304. Chen J, Spracklen CN, Marenne G, Varshney A, Corbin LJ, Luan J, et al. The trans-ancestral genomic architecture of glycemic traits. Nat Genet. 2021;53(6):840-60.
305. . !!! INVALID CITATION !!! (299).
306. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. Stat Methods Med Res. 2017;26(5):2333-55.
307. Sanderson E, Richardson TG, Hemani G, Davey Smith G. The use of negative control outcomes in Mendelian randomization to detect potential population stratification. Int J Epidemiol. 2021;50(4):1350-61.
308. Jones G, Trajanoska K, Santanasto AJ, Stringa N, Kuo CL, Atkins JL, et al. Genome-wide meta-analysis of muscle weakness identifies 15 susceptibility loci in older men and women. Nat Commun. 2021;12(1):654.
309. Jiang L, Zheng Z, Fang H, Yang J. A generalized linear mixed model association tool for biobank-scale data. Nat Genet. 2021;53(11):1616-21.

310. Nalls MA, Blauwendraat C, Vallerga CL, Heilbron K, Bandres-Ciga S, Chang D, et al. Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. Lancet Neurol. 2019;18(12):1091-102.
311. Loh PR, Kichaev G, Gazal S, Schoech AP, Price AL. Mixed-model association for biobank-scale datasets. Nat Genet. 2018;50(7):906-8.
312. Neale B. Rapid GWAS of Thousands of Phenotypes in the UK Biobank 2020 [
313. Georgakis MK, Gill D, Webb AJS, Evangelou E, Elliott P, Sudlow CLM, et al. Genetically determined blood pressure, antihypertensive drug classes, and risk of stroke subtypes. Neurology. 2020;95(4):e353-e61.
314. Li Y, Xiao W, Huang N, Zhuang Z, Zhang L, Wang W, et al. The effects of blood pressure and antihypertensive drugs on heart failure: A Mendelian randomization study. Nutr Metab Cardiovasc Dis. 2023;33(7):1420-8.
315. Pathak RK, Middeldorp ME, Lau DH, Mehta AB, Mahajan R, Twomey D, et al. Aggressive risk factor reduction study for atrial fibrillation and implications for the outcome of ablation: the ARREST-AF cohort study. J Am Coll Cardiol. 2014;64(21):2222-31.
316. Rahman F, Yin X, Larson MG, Ellinor PT, Lubitz SA, Vasan RS, et al. Trajectories of Risk Factors and Risk of New-Onset Atrial Fibrillation in the Framingham Heart Study. Hypertension. 2016;68(3):597-605.
317. International Consortium for Blood Pressure Genome-Wide Association S, Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature. 2011;478(7367):103-9.
318. Aikens RC, Zhao W, Saleheen D, Reilly MP, Epstein SE, Tikkanen E, et al. Systolic Blood Pressure and Risk of Type 2 Diabetes: A Mendelian Randomization Study. Diabetes. 2017;66(2):543-50.
319. Sun D, Zhou T, Heianza Y, Li X, Fan M, Fonseca VA, et al. Type 2 Diabetes and Hypertension. Circ Res. 2019;124(6):930-7.
320. Xie X, Atkins E, Lv J, Bennett A, Neal B, Ninomiya T, et al. Effects of intensive blood pressure lowering on cardiovascular and renal outcomes: updated systematic review and meta-analysis. Lancet. 2016;387(10017):435-43.
321. Beddhu S, Rocco MV, Toto R, Craven TE, Greene T, Bhatt U, et al. Effects of Intensive Systolic Blood Pressure Control on Kidney and Cardiovascular Outcomes in Persons Without Kidney Disease: A Secondary Analysis of a Randomized Trial. Ann Intern Med. 2017;167(6):375-83.
322. Yu Z, Coresh J, Qi G, Grams M, Boerwinkle E, Snieder H, et al. A bidirectional Mendelian randomization study supports causal effects of kidney function on blood pressure. Kidney Int. 2020;98(3):708-16.
323. Officers A, Coordinators for the ACRGTA, Lipid-Lowering Treatment to Prevent Heart Attack T. 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. 2002;288(23):2981-97.
324. Kloner RA, Stanek E, Crowe CL, Singhal M, Pepe RS, Bradsher J, et al. Effect of phosphodiesterase type 5 inhibitors on major adverse cardiovascular events and overall mortality in a large nationwide cohort of men with erectile dysfunction and cardiovascular risk factors: A retrospective, observational study based on healthcare claims and national death index data. J Sex Med. 2023;20(1):38-48.
325. Zilebesiran as Add-on Therapy in Patients With Hypertension Not Adequately Controlled by a Standard of Care Antihypertensive Medication (KARDIA-2) [Available from: https://classic.clinicaltrials.gov/ct2/show/NCT05103332.
326. McCallum L, Lip S, McConnachie A, Brooksbank K, MacIntyre IM, Doney A, et al. UMOD Genotype-Blinded Trial of Ambulatory Blood Pressure Response to Torasemide. Hypertension. 2024;81(10):2049-59.
327. Hutchings DC, Anderson SG, Caldwell JL, Trafford AW. Phosphodiesterase-5 inhibitors and the heart: compound cardioprotection? Heart. 2018;104(15):1244-50.
328. Corinaldesi C, Di Luigi L, Lenzi A, Crescioli C. Phosphodiesterase type 5 inhibitors: back and forward from cardiac indications. J Endocrinol Invest. 2016;39(2):143-51.
329. Nguyen H, Amanullah AM. Therapeutic potentials of phosphodiesterase-5 inhibitors in cardiovascular disease. Rev Cardiovasc Med. 2014;15(2):158-67.
330. Gebska MA, Stevenson BK, Hemnes AR, Bivalacqua TJ, Haile A, Hesketh GG, et al. Phosphodiesterase-5A (PDE5A) is localized to the endothelial caveolae and modulates NOS3 activity. Cardiovasc Res. 2011;90(2):353-63.

331. Andersson DP, Landucci L, Lagerros YT, Grotta A, Bellocco R, Lehtihet M, et al. Association of Phosphodiesterase-5 Inhibitors Versus Alprostadil With Survival in Men With Coronary Artery Disease. J Am Coll Cardiol. 2021;77(12):1535-50.
332. Kukreja RC, Salloum FN, Das A. Cyclic guanosine monophosphate signaling and phosphodiesterase-5 inhibitors in cardioprotection. J Am Coll Cardiol. 2012;59(22):1921-7.
333. Dousa TP. Cyclic-3',5'-nucleotide phosphodiesterase isozymes in cell biology and pathophysiology of the kidney. Kidney Int. 1999;55(1):29-62.
334. Brown KE, Dhaun N, Goddard J, Webb DJ. Potential therapeutic role of phosphodiesterase type 5 inhibition in hypertension and chronic kidney disease. Hypertension. 2014;63(1):5-11.
335. Rodriguez-Iturbe B, Ferrebuz A, Vanegas V, Quiroz Y, Espinoza F, Pons H, et al. Early treatment with cGMP phosphodiesterase inhibitor ameliorates progression of renal damage. Kidney Int. 2005;68(5):2131-42.
336. Kuno Y, Iyoda M, Shibata T, Hirai Y, Akizawa T. Sildenafil, a phosphodiesterase type 5 inhibitor, attenuates diabetic nephropathy in non-insulin-dependent Otsuka Long-Evans Tokushima Fatty rats. Br J Pharmacol. 2011;162(6):1389-400.
337. Lee KW, Jeong JY, Lim BJ, Chang YK, Lee SJ, Na KR, et al. Sildenafil attenuates renal injury in an experimental model of rat cisplatin-induced nephrotoxicity. Toxicology. 2009;257(3):137-43.
338. Jeong KH, Lee TW, Ihm CG, Lee SH, Moon JY, Lim SJ. Effects of sildenafil on oxidative and inflammatory injuries of the kidney in streptozotocin-induced diabetic rats. Am J Nephrol. 2009;29(3):274-82.
339. Oruc O, Inci K, Aki FT, Zeybek D, Muftuoglu SF, Kilinc K, et al. Sildenafil attenuates renal ischemia reperfusion injury by decreasing leukocyte infiltration. Acta Histochem. 2010;112(4):337-44.
340. Scheele W, Diamond S, Gale J, Clerin V, Tamimi N, Le V, et al. Phosphodiesterase Type 5 Inhibition Reduces Albuminuria in Subjects with Overt Diabetic Nephropathy. J Am Soc Nephrol. 2016;27(11):3459-68.
341. Morrow AJ, Ford TJ, Mangion K, Kotecha T, Rakhit R, Galasko G, et al. Rationale and design of the Medical Research Council's Precision Medicine with Zibotentan in Microvascular Angina (PRIZE) trial. Am Heart J. 2020;229:70-80.
342. Civieri G, Iop L, Tona F. Antibodies against Angiotensin II Type 1 and Endothelin 1 Type A Receptors in Cardiovascular Pathologies. Int J Mol Sci. 2022;23(2).
343. Buys ES, Zimmer DP, Chickering J, Graul R, Chien YT, Profy A, et al. Discovery and development of next generation sGC stimulators with diverse multidimensional pharmacology and broad therapeutic potential. Nitric Oxide. 2018;78:72-80.
344. Carlstrom M. Nitric oxide signalling in kidney regulation and cardiometabolic health. Nat Rev Nephrol. 2021;17(9):575-90.
345. Stasch JP, Schlossmann J, Hocher B. Renal effects of soluble guanylate cyclase stimulators and activators: a review of the preclinical evidence. Curr Opin Pharmacol. 2015;21:95-104.
346. Boyle EA, Li YI, Pritchard JK. An Expanded View of Complex Traits: From Polygenic to Omnigenic. Cell. 2017;169(7):1177-86.
347. Wu Y, Zheng Z, Visscher PM, Yang J. Quantifying the mapping precision of genome-wide association studies using whole-genome sequencing data. Genome Biol. 2017;18(1):86.
348. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, et al. Systematic localization of common disease-associated variation in regulatory DNA. Science. 2012;337(6099):1190-5.
349. Porcu E, Rueger S, Lepik K, e QC, Consortium B, Santoni FA, et al. Mendelian randomization integrating GWAS and eQTL data reveals genetic determinants of complex and clinical traits. Nat Commun. 2019;10(1):3300.
350. Fortune MD, Guo H, Burren O, Schofield E, Walker NM, Ban M, et al. Statistical colocalization of genetic risk variants for related autoimmune diseases in the context of common controls. Nat Genet. 2015;47(7):839-46.
351. Wallace C, Rotival M, Cooper JD, Rice CM, Yang JH, McNeill M, et al. Statistical colocalization of monocyte gene expression and genetic risk variants for type 1 diabetes. Hum Mol Genet. 2012;21(12):2815-24.
352. Ryan DK, Karhunen V, Su B, Traylor M, Richardson TG, Burgess S, et al. Genetic Evidence for Protective Effects of Angiotensin-Converting Enzyme Against Alzheimer Disease But Not Other Neurodegenerative Diseases in European Populations. Neurol Genet. 2022;8(5):e200014.

353. Zhang L, Xiong Y, Zhang J, Feng Y, Xu A. Systematic proteome-wide Mendelian randomization using the human plasma proteome to identify therapeutic targets for lung adenocarcinoma. J Transl Med. 2024;22(1):330.
354. Nica AC, Dermitzakis ET. Expression quantitative trait loci: present and future. Philosophical Transactions of the Royal Society B: Biological Sciences. 2013;368(1620).
355. Chauquet S, Zhu Z, O'Donovan MC, Walters JTR, Wray NR, Shah S. Association of Antihypertensive Drug Target Genes With Psychiatric Disorders. JAMA Psychiatry. 2021;78(6).
356. Richardson TG, Hemani G, Gaunt TR, Relton CL, Davey Smith G. A transcriptome-wide Mendelian randomization study to uncover tissue-dependent regulatory mechanisms across the human phenome. Nature Communications. 2020;11(1).
357. Xiao J, Huang W, Ji J, Chen L. Association of lipid-lowering drugs with COVID-19 outcomes from a Mendelian randomization study. eLife. 2021;10.
358. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nature Genetics. 2016;48(5):481-7.
359. Martin FJ, Amode MR, Aneja A, Austine-Orimoloye O, Azov AG, Barnes I, et al. Ensembl 2023. Nucleic Acids Res. 2023;51(D1):D933-D41.
360. Wobst J, Kessler T, Dang TA, Erdmann J, Schunkert H. Role of sGC-dependent NO signalling and myocardial infarction risk. J Mol Med (Berl). 2015;93(4):383-94.
361. Kessler T, Wobst J, Wolf B, Eckhold J, Vilne B, Hollstein R, et al. Functional Characterization of the GUCY1A3 Coronary Artery Disease Risk Locus. Circulation. 2017;136(5):476-89.
362. Erdmann J, Stark K, Esslinger UB, Rumpf PM, Koesling D, de Wit C, et al. Dysfunctional nitric oxide signalling increases risk of myocardial infarction. Nature. 2013;504(7480):432-6.
363. Wobst J, von Ameln S, Wolf B, Wierer M, Dang TA, Sager HB, et al. Stimulators of the soluble guanylyl cyclase: promising functional insights from rare coding atherosclerosis-related GUCY1A3 variants. Basic Res Cardiol. 2016;111(4):51.
364. Emdin CA, Khera AV, Klarin D, Natarajan P, Zekavat SM, Nomura A, et al. Phenotypic Consequences of a Genetic Predisposition to Enhanced Nitric Oxide Signaling. Circulation. 2018;137(3):222-32.
365. Dichgans M, Malik R, Konig IR, Rosand J, Clarke R, Gretarsdottir S, et al. Shared genetic susceptibility to ischemic stroke and coronary artery disease: a genome-wide analysis of common variants. Stroke. 2014;45(1):24-36.
366. Nelson CP, Goel A, Butterworth AS, Kanoni S, Webb TR, Marouli E, et al. Association analyses based on false discovery rate implicate new loci for coronary artery disease. Nat Genet. 2017;49(9):1385-91.
367. Dang TA, Kessler T, Wobst J, Wierer M, Braenne I, Strom TM, et al. Identification of a Functional PDE5A Variant at the Chromosome 4q27 Coronary Artery Disease Locus in an Extended Myocardial Infarction Family. Circulation. 2021;144(8):662-5.
368. Feghaly J, Zakka P, London B, MacRae CA, Refaat MM. Genetics of Atrial Fibrillation. J Am Heart Assoc. 2018;7(20):e009884.
369. Sinner MF, Pfeufer A, Akyol M, Beckmann BM, Hinterseer M, Wacker A, et al. The non-synonymous coding IKr-channel variant KCNH2-K897T is associated with atrial fibrillation: results from a systematic candidate gene-based analysis of KCNH2 (HERG). Eur Heart J. 2008;29(7):907-14.
370. Hong K, Bjerregaard P, Gussak I, Brugada R. Short QT syndrome and atrial fibrillation caused by mutation in KCNH2. J Cardiovasc Electrophysiol. 2005;16(4):394-6.
371. Mann SA, Otway R, Guo G, Soka M, Karlsdotter L, Trivedi G, et al. Epistatic effects of potassium channel variation on cardiac repolarization and atrial fibrillation risk. J Am Coll Cardiol. 2012;59(11):1017-25.
372. Cambien F, Poirier O, Lecerf L, Evans A, Cambou JP, Arveiler D, et al. Deletion polymorphism in the gene for angiotensin-converting enzyme is a potent risk factor for myocardial infarction. Nature. 1992;359(6396):641-4.
373. Chen YH, Liu JM, Hsu RJ, Hu SC, Harn HJ, Chen SP, et al. Angiotensin converting enzyme DD genotype is associated with acute coronary syndrome severity and sudden cardiac death in Taiwan: a case-control emergency room study. BMC Cardiovasc Disord. 2012;12:6.
374. Pigeyre M, Sjaarda J, Chong M, Hess S, Bosch J, Yusuf S, et al. ACE and Type 2 Diabetes Risk: A Mendelian Randomization Study. Diabetes Care. 2020;43(4):835-42.

375. Yusuf S, Gerstein H, Hoogwerf B, Pogue J, Bosch J, Wolffenbuttel BH, et al. Ramipril and the development of diabetes. JAMA. 2001;286(15):1882-5.
376. Vermes E, Ducharme A, Bourassa MG, Lessard M, White M, Tardif JC, et al. Enalapril reduces the incidence of diabetes in patients with chronic heart failure: insight from the Studies Of Left Ventricular Dysfunction (SOLVD). Circulation. 2003;107(9):1291-6.
377. Becker A, Wardas B, Salah H, Amini M, Fecher-Trost C, Sen Q, et al. Cavbeta3 Regulates Ca(2+) Signaling and Insulin Expression in Pancreatic beta-Cells in a Cell-Autonomous Manner. Diabetes. 2021;70(11):2532-44.
378. Triposkiadis F, Karayannis G, Giamouzis G, Skoularigis J, Louridas G, Butler J. The sympathetic nervous system in heart failure physiology, pathophysiology, and clinical implications. J Am Coll Cardiol. 2009;54(19):1747-62.
379. Yamada Y, Kinoshita H, Kuwahara K, Nakagawa Y, Kuwabara Y, Minami T, et al. Inhibition of N-type Ca2+ channels ameliorates an imbalance in cardiac autonomic nerve activity and prevents lethal arrhythmias in mice with heart failure. Cardiovasc Res. 2014;104(1):183-93.
380. Suchard MA, Schuemie MJ, Krumholz HM, You SC, Chen R, Pratt N, et al. Comprehensive comparative effectiveness and safety of first-line antihypertensive drug classes: a systematic, multinational, large-scale analysis. Lancet. 2019;394(10211):1816-26.
381. Law MR, Morris JK, Wald NJ. Use of blood pressure lowering drugs in the prevention of cardiovascular disease: meta-analysis of 147 randomised trials in the context of expectations from prospective epidemiological studies. BMJ. 2009;338:b1665.
382. Keaton JM, Kamali Z, Xie T, Vaez A, Williams A, Goleva SB, et al. Genome-wide analysis in over 1 million individuals of European ancestry yields improved polygenic risk scores for blood pressure traits. Nat Genet. 2024;56(5):778-91.