



TÉCNICO
LISBOA

Analysis of a genetic test in the context of arterial hypertension

João Luís Lopes Amorim

Thesis to obtain the Master of Science Degree in

Biomedical Engineering

Supervisor(s): Prof. Ana Teresa Correia de Freitas
Dr. José Miguel Ranhadas Vellez Caldas

Examination Committee

Chairperson: Prof. João Pedro Estrela Rodrigues Conde
Supervisor: Dr. José Miguel Ranhadas Vellez Caldas
Member of the Committee: Prof. Susana de Almeida Mendes Vinga Martins

November 2018

Acknowledgments

I would like to express my most sincere thanks to everyone that helped and supported me during the elaboration of this thesis.

Firstly, I would like to thank my advisers Professor Ana Teresa Freitas and Doctor José Caldas for all the guidance provided and important feedback given that improved the quality of this work. Their knowledge and insight were crucial to the development of this thesis. I would also like to express my gratitude to Professor João André Nogueira Custódio Carriço for being a member of the supervisor committee.

I would also like to thank all the HeartGenetics staff for welcoming me and making me feel as member of the team.

To my family, friends, and colleagues I express my many thanks for accompanying me during this long journey.

Resumo

A hipertensão essencial, caracterizada por elevação patológica da pressão arterial, afeta globalmente mil milhões de indivíduos e constitui um importante fator de risco para doenças cardiovasculares. Na maioria dos casos, esta condição pode ser controlada por terapia farmacológica. A evolução das tecnologias de análise de DNA e metodologias de análise de dados permitem explorar a genética da hipertensão.

O objetivo desta tese é analisar um painel genético da HeartGenetics relacionado com a hipertensão essencial. Para tal, foi criado um modelo que calcula o risco genético para a hipertensão e um modelo discriminativo de doentes com hipertensão resistente com base neste painel. Este estudo foi suportado por uma revisão bibliográfica relativa à associação entre variantes genéticas e hipertensão.

Num conjunto de 417 indivíduos foi utilizado um classificador de regressão logística para o risco de hipertensão, cuja performance foi avaliada através da área sob a curva (AUC = 0.568). Foi também efetuada uma análise exploratória dos dados na qual foram identificadas seis variantes associadas à hipertensão. Usando um coorte com 322 indivíduos, foi possível discriminar com significância estatística pacientes com hipertensão resistente (valor-p = 0.003). A revisão bibliográfica levou à sugestão da inclusão de quatro variantes ao painel genético considerado.

Os resultados sugerem que a previsão do risco genético para a hipertensão essencial é complexa e necessita de um painel genético alargado para obter resultados com aplicabilidade clínica. No entanto, no contexto da hipertensão resistente existe um potencial significativo e com aplicabilidade clínica para desenvolver um teste genético para esta condição.

Palavras-chave: hipertensão, genética da hipertensão, farmacogenética, testes genéticos, variantes genéticas

Abstract

Essential hypertension, characterized by pathological blood pressure elevation, affects globally one billion individuals and constitutes one of the most important risk factors for cardiovascular diseases. In most cases, this condition can be controlled through pharmacological therapy. Due to advances in DNA analysis technologies and data analysis methodologies the genetics of hypertension can be explored.

The objective of this thesis is the analysis of a HeartGenetics' genetic panel associated with essential hypertension. For this, a genetic risk score model for hypertension and a discriminative model for patients with treatment-resistant hypertension were developed using this panel. This study was supported by a bibliographic revision relative to the association between genetic variants and hypertension.

In a cohort with 417 individuals a logistic regression classifier was used to calculate the risk for hypertension whose performance was evaluated by the area under the curve (AUC = 0.568). Additionally, a univariate analysis was performed from which six genetic variants associated with hypertension were identified. Additionally, using a cohort with 322 individuals it was possible to discriminate patients with treatment-resistant hypertension (p-value = 0.003). The bibliographic revision led to the suggestion of inclusion of four genetic variants to the HeartGenetics' genetic panel.

The results of this thesis suggest that genetic risk prediction of essential hypertension is complex and requires an extensive genetic panel to obtain results with a clear clinical applicability. Nonetheless, in the context of treatment-resistant hypertension, the results suggest that there is potential for the development of a genetic test for this condition with clinical applicability.

Keywords: hypertension, hypertension genetics, pharmacogenetics, genetic testing, genetic variants

Contents

Acknowledgments	iii
Resumo	v
Abstract	vii
List of Tables	xi
List of Figures	xiii
Nomenclature	xv
1 Introduction	1
1.1 Motivation and Objectives	1
1.2 Thesis Outline	2
2 Genetics, Pharmacogenetics and Bioinformatics	3
2.1 Principles of Genetics	3
2.2 Pharmacogenomics and Pharmacogenetics	6
2.2.1 Drug-Protein Interactions	6
2.3 Genetic Markers and Association Studies	8
2.4 Overview of Statistical Methods for Genetic Association Studies	10
2.4.1 Genetic Models and Genotype Data Quality Control	10
2.4.2 Univariate Association Tests	12
2.4.3 Logistic Regression	13
3 Hypertension Background	15
3.1 Definition and Impact of Hypertension	15
3.2 Pathophysiology of Hypertension	16
3.2.1 Sodium Homeostasis Regulation	17
3.2.2 Renin-Angiotensin-Aldosterone System (RAAS)	18
3.2.3 Natriuretic Peptides	18
3.2.4 Endothelium	19
3.2.5 Sympathetic Nervous System (SNS)	19
3.2.6 Inflammation and the immune system	20
3.3 Risk Factors for Hypertension	21
3.4 Hypertension as a Risk Factor for Several Diseases	24

3.5	Hypertension Treatment	24
3.5.1	Non-Pharmacological Approach	25
3.5.2	Pharmacological Approach	27
3.5.3	Monotherapy and Combination Therapy	30
3.5.4	Treatment-Resistant Hypertension	31
4	Materials and Methods	33
4.1	Data Selection and Characterization	33
4.2	Genotype Data Encoding and Genetic Models	36
4.3	Genotype Data Quality Control	36
4.4	Univariate Association Analysis for the Identification of Relevant Genetic Variants in Hypertension	37
4.5	Construction of a Genetic Risk Score for Hypertension	37
4.6	Criteria for the Selection of Relevant Genetic Variants with Pharmacogenetic Associations	37
4.7	Analysis of the Genetic Influence on Treatment-Resistant Hypertension	39
5	Results	41
5.1	Genotype Data Quality Control	41
5.2	Genetic Variants Associated with Hypertension	42
5.3	Genetic Risk Score Model Evaluation for Essential Hypertension	42
5.4	Genetic Variants with Relevant Pharmacogenetic Associations	42
5.5	Genetic Risk Score Model Evaluation for Treatment-Resistant Hypertension	43
6	Discussion and Conclusions	47
	Bibliography	53
A	HeartGenetics DNArterial Genetic Panel	69
B	Association Analysis Results for the Hypertension Trait	73

List of Tables

2.1	Contingency tables for case control analyses, by genetic model.	12
3.1	Blood pressure categories in adults	16
3.2	Recommended treatment strategies for each stratum of total cardiovascular risk.	25
3.3	Drugs for the management of hypertension preferred in specific health conditions.	28
4.1	Demographic information of the two cohorts under study.	34
A.1	General genetic information for the genetic variants on the DNArterial test genetic panel.	70
A.2	General genetic information for the genetic variants on the DNArterial test genetic panel.	71
B.1	Association results for hypertension trait and quality control results in the HTN samples. .	74

List of Figures

2.1	Molecular structure of the DNA molecule.	4
3.1	The major neuroendocrine systems involved in blood pressure regulation.	17
3.2	Genome wide association for essential hypertension using the UK biobank dataset.	22
3.3	Recommended therapeutic strategies used to achieve blood pressure target.	31
4.1	Flow diagram of pharmacogenetic associations selection.	35
4.2	Decision diagram with the different criteria considered for the selection of genetic variants with relevant pharmacogenetic associations.	38
5.1	Population stratification analysis results using a PCA followed by a k-means.	41
5.2	Receiver operator characteristic (ROC) curves of the logistic regressions used for hypertension discrimination.	43
5.3	Boxplot of the genetic risk scores calculated using a total of 24 genetic variants for the cases and controls groups by genetic model used.	45
5.4	Boxplot of the genetic risk scores calculated using a total of six genetic variants for the cases and controls groups by genetic model used.	46

Nomenclature

A Adenine.

ABC ATP binding cassette.

ACE Angiotensin I converting enzyme.

ADD1 Adducin 1 (alpha).

ADRA1A Adrenoceptor alpha 1A .

ADRB1 Adrenoceptor beta 1.

ADRB2 Adrenoceptor beta 2.

AGT Angiotensinogen.

AGTR1 Angiotensin II receptor, type 1.

AGTR2 Angiotensin II receptor, type 2.

ANOVA Analysis of variance.

ANP Atrial natriuretic peptide.

AUC Area under the curve.

BDKRB2 Bradykinin receptor B2.

BIC Bayesian information criterion.

BMI Body mass index.

BNP Brain natriuretic peptide.

BP Blood Pressure.

CACNA1C Calcium voltage-gated channel subunit alpha 1 C.

CACNB2 Calcium voltage-gated channel auxiliary subunit beta 2.

CAD Coronary artery disease.

CALCA Calcitonin-related polypeptide alpha.

C Cytosine.

CCB Calcium-channel blockers.

CHUC Coimbra Hospital and University Centre.

CKD chronic kidney disease.

CLCNKA Chloride voltage-gated channel Ka.

CLCNKB Chloride voltage-gated channel Kb.

CO Cardiac Output.

Corin Corin, serine peptidase.

Corin Atrial natriuretic peptide-converting enzyme.

CV Cardiovascular.

CVD Cardiovascular disease.

CVP Central venous pressure.

CYBA Cytochrome B-245, alpha polypeptide.

CYP17A1 Cytochrome P450 family 17 subfamily A member 1.

CYP2C9 Cytochrome P450 2C9.

CYP4A11 Cytochrome P450 family 4 subfamily A member 11.

CYP Cytochrome.

DASH Dietary approaches to stop hypertension.

DBP Diastolic blood pressure.

DNA Deoxyribonucleic acid.

DRD3 Dopamine receptor D3.

ECE1 Endothelin converting enzyme 1.

EDN1 Endothelin.

EDNRA Endothelin receptor type A.

ESC European society of cardiology.

ESH European society of hypertension.

ESRD End-stage renal disease.

FDA Food and drug administration.

FGF5 Fibroblast growth factor 5.

FPR False positive rate.

GCH1 GTP cyclohydrolase 1.

G Guanine.

GRK4 G protein-coupled receptor kinase 4.

GWAS Genome-wide association study.

HF Heart Failure.

HWE Hardy-Weinberg equilibrium.

ISH Isolated systolic hypertension.

KCNMB1 Potassium calcium-activated channel subfamily M regulatory beta subunit 1.

LD Linkage disequilibrium.

LVH Left ventricular hypertrophy.

MAF Minor allele frequency.

MAP Mean arterial pressure.

MDR1 Multidrug resistance protein 1.

miRNA MicroRNA.

mRNA Messenger ribonucleic acid.

NO Nitric oxide.

NOS2 Nitric oxide synthase 2.

NOS3 Nitric oxide synthase 3.

NOS Nitric oxide synthase.

NPPA Natriuretic peptide A.

NPPC Natriuretic peptide C.

NR3C2 Nuclear receptor subfamily 3 group C member 2.

OD Organ damage.

OR Odds ratio.

PCA Principal component analysis.

PD Pharmacokinetics.

PharmGKB Pharmacogenomics Knowledgebase.

PK Pharmacodynamics.

PP Pulse pressure.

ProANP Precursor of the atrial natriuretic peptide.

ProBNP Precursor of the brain natriuretic peptide.

RAAS Renin-Angiotensin-Aldosterone System.

REN Renin.

RETN Resistin.

RF Risk factor.

RNA Ribonucleic acid.

ROC Receiver operating characteristic curve.

rRNA Ribosomal ribonucleic acid.

SBP Systolic blood pressure.

SCNN1A Sodium channel epithelial 1 alpha subunit.

siRNA Short interfering ribonucleic acid.

SLC12A3 Solute carrier family 12 member 3.

SNP Single nucleotide polymorphism.

SNS Sympathetic nervous system.

STK39 Serine threonine kinase 39.

SVR Systemic vascular resistance.

TPR True positive rate.

tRNA Transcribed ribonucleic acid.

T Thymine.

VKORC1 Vitamin K epoxide reductase complex subunit 1.

WHO World Health Organization.

WNK1 WNK lysine deficient protein kinase 1.

WNK4 Serine/threonine-protein kinase.

Chapter 1

Introduction

1.1 Motivation and Objectives

Cardiovascular diseases (CVDs) are the global leading cause of death. It is estimated that 17 million people die from CVDs each year, corresponding to 30% of all global deaths [1]. Most CVDs can be prevented by addressing risk factors that contribute to their development.

Arterial hypertension is a medical condition that affects almost 1 billion individuals worldwide and constitutes one of the most important risk factors for the development of CVDs such as coronary artery disease, myocardial infarction and heart failure [2].

The development of hypertension involves the complex interplay of both non-genetic and genetic factors. The underlying genetic factors involved in the development of hypertension and in the therapeutic response to antihypertensive drugs have been highlighted in several clinical studies [3–6]. The most common strategy used for the discovery and validation of these genetic factors is based on association studies. These studies test the association between genetic markers, usually involved in metabolic pathways related with blood pressure control, and the phenotype under study.

Although a plethora of antihypertensive drugs such as diuretics, β -blockers, angiotensin converting enzyme inhibitors, calcium channel blockers, and angiotensin receptor blockers are used for blood pressure control, approximately 40% of all hypertensive patients under treatment do not have hypertension under control [7].

Considering the genetic factors underlying hypertension, the development of a genetic test for this condition with a well-defined clinical utility, using a personalized medicine approach, may contribute to the management and treatment of hypertension which can lead ultimately to a decrease on cardiovascular events and to the sustainability of healthcare systems.

The HeartGenetics project, funded by Portugal2020, in cooperation with six leading European institutions in the cardiovascular area (Monzino Cardiology Centre (CCM) and Italian Diagnostic Center (CDI) in Italy, Paris-Cardiovascular research Center (PARCC) in France, Hospital da Luz (HL) and Portuguese Society of Hypertension (SPH) in Portugal, and Hannover Medical School (HMS) in Germany), and the work developed in this thesis had as main objective the validation of the clinical utility of the

HeartGenetics DNArterial genetic test for arterial hypertension.

In more detail, the aim of this thesis was to perform a clinical validation study of the DNArterial genetic panel on two components of the test: risk prediction and pharmacogenetic-guided therapy for hypertension. To evaluate the risk prediction component of the test a genetic association study of hypertension was performed. The objectives of this analysis consisted on the identification of genetic variants from the DNArterial genetic panel with statistical significant associations with hypertension and on the construction of a genetic risk score for this condition. To evaluate the pharmacogenetic component of the test an analysis of the genetic influence of pharmacogenetic variants on treatment-resistant hypertension was performed. Additionally, a literature search with the objective of suggesting the addition of additional pharmacogenetic variants to be included in the DNArterial genetic panel was performed.

1.2 Thesis Outline

In this section, a brief description of this thesis structure is provided to the reader. The present chapter introduces the motivations and objectives of this dissertation.

Chapter 2 reviews the basic concepts of genetics and pharmacogenetics, as well as the main bioinformatic topics related with genetic association studies.

Chapter 3 focus on the biological background of hypertension, specifically regarding the pathophysiology of this condition, as well as the most relevant risk factors contributing to elevated blood pressure and the impact of hypertension on other diseases. In addition, are also described the treatment approaches used to manage blood pressure.

Chapter 4 describes the materials and methods used in this thesis, starting with a characterization of the genotype data, followed by a description of how the genotype data was prepared and tested for quality control. Sections 4.4 and 4.5 present the methodology used to identify genetic variants associated with hypertension and quantify the genetic risk of developing hypertension. Section 4.6 describes the methods used for selecting genetic variants with relevant pharmacogenetic associations. Finally, in section 4.7 are described the methods used to quantify the genetic risk of developing treatment-resistant hypertension and identify the genetic variants associated with this condition.

In Chapter 5 the obtained results concerning the previously mentioned analyses performed are presented. The Dissertation ends in Chapter 6 with the discussion of the obtained results and an overall conclusion of the work developed.

Chapter 2

Genetics, Pharmacogenetics and Bioinformatics

2.1 Principles of Genetics

The human genome is composed of long molecules of deoxyribonucleic acid (DNA) and can be divided in two parts, the nuclear and the mitochondrial genome. The nuclear genome has approximately 3 billion base pairs which are divided into 23 chromosomes. Of the 23 pairs of chromosomes, 22 are autosomes and one is a sex chromosome (X or Y). In opposition, the mitochondrial genome is composed of circular molecules located in the cell mitochondria [8].

A DNA molecule is a linear and double-helical structure composed of two long intertwined strands that have as subunits four types of nucleotides (Figure 2.1). Each nucleotide has a deoxyribose molecule, a phosphate group, and a nitrogen-containing base. The deoxyribose molecule is a monosaccharide (simple sugar) composed of five carbon atoms that are numbered from 1' to 5'. The phosphate group is composed of up to three phosphate units, termed α , β , and γ , with the α being the one directly connected to the 5'-carbon of the deoxyribose. Lastly, the nitrogen-containing bases are attached to the 1'-carbon of the deoxyribose and they may be either an adenine (A), a cytosine (C), a guanine (G), or a thymine (T). In the same strand, consecutive nucleotides are connected by strong phosphodiester bonds between the sugar and phosphate portions of each nucleotide. Weaker hydrogen bonds between bases of opposite strands hold together the intertwined strands (adenine pairs with thymine and cytosine pairs with guanine). Each strand has two different ends, one having a terminal phosphate group attached to a 5'-carbon (5' end) and the other having a terminal hydroxyl group connected to a 3'-carbon (3' end). Being the two strands anti-parallel *i.e.* oriented in opposite directions, one can be expressed as $5' \rightarrow 3'$ (positive strand) and the other as $3' \rightarrow 5'$ (negative strand) [8].

During replication, the two DNA strands are cleaved and act as templates for the construction of new complementary strands, resulting in two identical copies of the original molecule. During the transcription process, a single strand of DNA is also used as a template for the formation of a complementary strand of ribonucleic acid (RNA). RNA is almost identical to the DNA molecule, with ribose and uracil base (U)

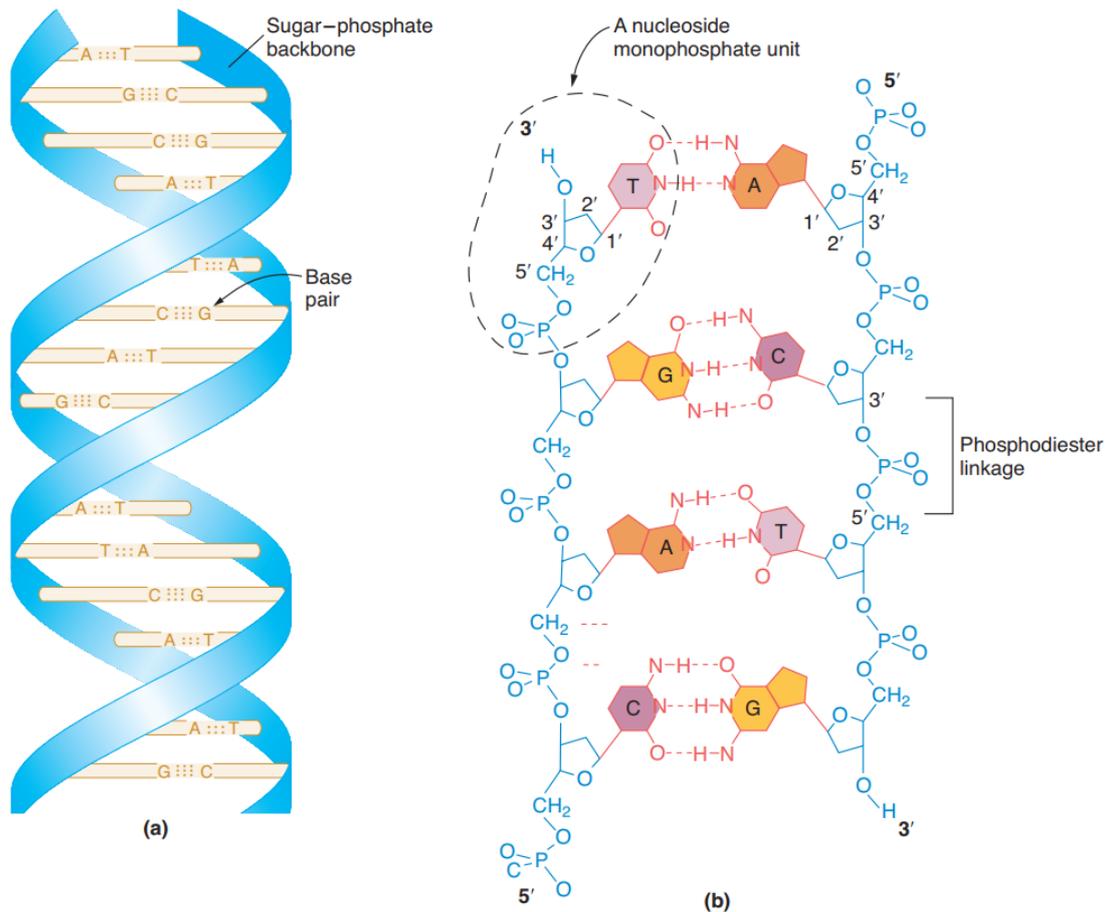


Figure 2.1: Molecular structure of the DNA molecule. (a) Simplified model showing the helical structure of DNA (b) Detailed chemical diagram of the DNA double helix. From Suzuki et al. [9].

replacing the DNA deoxyribose and thymine bases, respectively.

The transcribed RNA is an identical copy of the entire gene which includes introns and exons. Introns are non-coding nucleotide sequences of the gene that may have regulatory functions. By contrast, exons are coding sequences. The transcribed RNA undergoes post-transcriptional processing termed alternative splicing in which the introns are removed and the exons are spliced to form messenger RNA (mRNA) [8, 10].

Proteins are formed by translating mRNA to amino acids. During translation, the mRNA molecule is read in the positive direction ($5' \rightarrow 3'$) according to a genetic code. This code relates each group of three bases (codon) with a specific amino acid, the basic molecular structure of a protein. For translation to occur, mRNA molecules leave the nucleus and travel to the ribosome, characterized by having two subunits termed the large (50S) and small (30S) subunits. In the ribosome, the mRNA is placed between the two ribosome subunits where ribosomal RNA (rRNA) and transfer RNA (tRNA) molecules help the translation process. The tRNA molecules serve as adaptors that facilitate the base-pairing process, by having in one end a complementary codon to a mRNA codon and in the other a specific amino acid that is incorporated into the translated protein. In its turn, the rRNAs molecules are responsible for catalysing this process [11].

In addition, other RNA molecule types exist such as short interfering RNAs (siRNAs) and microRNAs (miRNAs). These RNA molecules are usually termed small RNAs since they are characterized by having approximately 20-30 nucleotides. Despite being non-coding RNA sequences, they are involved in gene expression regulation [12].

As previously mentioned, nuclear DNA is divided into 23 different pairs of chromosomes. Human cell nucleus are diploid which means that they contain two almost identical versions of each chromosomes, having 22 homologous pairs of autosomes and one pair of sexual chromosomes. For each chromosome pair, one is inherited from the father and the other from the mother. By contrast, haploid cells such as gametes contain half the chromosomes present in a diploid cell. During meiosis, through cell division and differentiation of diploid precursor cells, mature haploid gametes are formed. It would be reasonable to think that one chromosome from each homologous pair is chosen. However, the gamete receives a mixture of the two chromosomes due to crossover events. This random recombination process during meiosis is a source of genetic variability [13].

Although homologous chromosomes have the same sequence of genes, there are small genetic variations caused by mutations at several specific chromosome locations *i.e.* loci (singular: locus). If a locus has different DNA variations in the population, each version is termed allele. An individual with the same allele in the same locus of two homologous chromosomes is called homozygous at that locus. In opposition, if the alleles are different then the individual is called heterozygous at that locus [13].

Different types of genetic variation found in the genome range from structural changes involving whole chromosomes to single base-pair alterations. Those variations can be classified according to the frequency of the minor allele (*i.e.* the less common allele) as common or rare. Common variants or polymorphisms have a minor allele frequency (MAF) of at least 1% in the population, whereas rare variants have a MAF lower than 1% [14].

When a genetic variant is in correlation with nearby variants, it is said to be in linkage disequilibrium (LD). LD is formally defined as the non-random association of alleles in adjacent loci. This occurs because specific loci tend to be inherited together because of their physical proximity. Over successive generations this association tends to break up (LD decay) due to several reasons such as mutation rates, genetic drift, population growth, variable recombination, and natural selection [15].

The vast majority of genetic variations among individuals are single nucleotide polymorphisms (SNPs). It is estimated that the human genome has at least 11 million SNPs, of which 7 million SNPs have a MAF higher than 5% and the remaining between 1 and 5% [14]. Despite the large number of SNPs, most of them tend to be in LD with each other in different regions forming genomic regions known as LD blocks. Hence, a small number of SNPs (tag SNPs) can be used to identify each block and consequently can be used to capture approximately 90% of the genetic variation in a population [14]. Other types of genetic variations exist such as insertion–deletions (indels), block substitutions, and DNA inversions. These are commonly defined as structural variants [14].

Two critical concepts in genetics are genotype and phenotype. In a broad sense, the term genotype refers to the genetic constitution of an individual. However, in a narrower sense it can be used to refer to a particular locus that determines one of its characteristics. In turn, phenotype is an individual's

observable characteristic such as its morphology, biochemical, or physiological properties that result from the expression of the genotype. Due to the importance of these concepts for this thesis both genotype and phenotype will be described in more detail in section 4.1.

2.2 Pharmacogenomics and Pharmacogenetics

The terms pharmacogenomics and pharmacogenetics are usually used interchangeably, despite the first being a newer and broader concept. In general, pharmacogenetics is concerned with how an unexpected drug response may be influenced by a single gene. In its turn, pharmacogenomics is concerned about the influence of the genome on drug response [16].

The passage of a drug through the body can be divided in four stages. First, it needs to be absorbed by the body which can be accomplished through different methods, the most common being the non-injection methods. Then, the bloodstream helps with the distribution of the drug to its target where it is metabolized. This metabolization stage is responsible by converting the drug to an active form or to an inactive form, after being used by the organism. In the end, drug metabolites are excreted by the body. These four processes describe the pharmacokinetics (PK) of a drug which are facilitated by genes usually referred to as pharmacokinetic genes. By contrast, the pharmacodynamics (PD) of a drug is the study of the action or effect a drug has on the organism [17]. In this case, the drug acts directly on the target genes or indirectly on downstream pathways. These genes are usually referred as pharmacodynamic genes [18]. Given this, by understanding what genes are involved in the pharmacokinetic and pharmacodynamic pathways it is possible to optimize drug dosage and consequently improve its efficacy and avoid adverse effects [18]. Genetic association studies are performed with the propose of identifying the genetic factors involved in the pharmacological response of a specific drug. In this case the studied phenotype is related with the response to the drug or with possible adverse effects [18].

2.2.1 Drug-Protein Interactions

Proteins are macromolecules involved in most of the cellular processes. A protein's characteristics and shape depend on its amino acid sequence that can range in size from tens to thousands of amino acids. Most proteins have complex structures that contain active sites which affect its activity. These active sites are protein regions that form temporary bonds with substrate molecules [8].

Drugs usually act by binding to the active site of proteins which can either activate the protein function (agonist activity) or prevent its function by blocking the active site to agonist molecules (antagonist activity) [19].

These drug-protein interactions are expected whether they result by the organism action on the drug (pharmacokinetics) or by the drug action on the body (pharmacodynamics). Nevertheless, sometimes these interactions can be unexpected leading to adverse events. Genetic variation can lead to structural changes of the proteins involved in these interactions which will ultimately affect the pharmacological action of the drug [18].

Pharmacokinetic (PK) Interactions

The drug pathway, from ingestion to reaching its target and until leaving the organism, involves the interaction with several proteins that facilitate or delay the drug's progress. As seen previously, these interactions are explained by the pharmacokinetics of the drug that determines how quickly the effects of the drug are felt and their duration.

After being administered, a drug needs to be absorbed and distributed to its target. Bioavailability of a drug refers to the fraction or rate of the drug that is absorbed and reaches the systemic circulation, which is dependent on how the drug is administered. For instance, intravenous administration leads to complete bioavailability of the drug, whereas non-injection methods such as oral ingestion may not lead to complete bioavailability since, in this case, drug absorption is usually dependent on enzymatic action [20]. Once in systemic circulation, the drug is distributed to its target depending on the activity of proteins. Therefore, genetic variations that can lead to modification of the proteins involved in these processes can affect the pharmacokinetics of a drug [18]. As an example, ATP binding cassette (ABC) transporters are involved in the transport of drugs and genetic variants of the genes that encode these proteins are associated with bioavailability of several drugs. This is the case of the polymorphism C3435T on the gene multidrug resistance protein 1 (*MDR1*). Individuals with this polymorphism present elevated plasma levels of the cardiac drug digoxin as compared with individuals without this genetic variant [21].

The metabolism of a drug can either lead to the activation of the drug by converting it to its active metabolite or can lead to its inactivation by breaking down its active form into an inactive metabolite. As seen in other processes involved in the pharmacokinetics of a drug, the metabolism stage can also be influenced by existing genetic variations [18]. For instance, several proteins of the cytochrome P450 family, also known as CYP genes, are involved in the metabolism of most drugs. Polymorphisms in the genes that encode these proteins have been associated with individual drug response and adverse effects [22]. A clear example is the gene *CYP2C9* that is involved in the metabolism of approximately 10% of all drugs including warfarin, an anticoagulant [22]. It has been shown that genetic variability in this gene has an important influence on warfarin metabolism and consequently on the initial dose of the drug required to obtain the desired outcome [23].

To conclude, unwanted metabolites, consequence of the drug metabolism or activity, need be to be excreted of the body. As it would be expected, the excretion of drug metabolites may also be influenced by inter-individual genetic variation. For instance, individuals that are carriers of the genetic variants *CYP2C9*2* and *CYP2C9*3* (gene *CYP2C9*) have been associated with a 30% and 90% reduction in warfarin clearance, respectively [24].

Pharmacodynamic (PD) Interactions

The pharmacodynamics of a drug refers to the effects of the drug on its targets and the consequence of these effects on downstream pathways. The effect of the drug can be "on-target" which is the desired effect with a therapeutic action. However, if the drug acts "off-target", due to interaction with other proteins during its journey, it can lead to adverse effects. Pharmacodynamics is also concerned with the

relationship between drug concentration and its effect [18].

In many cases, drugs can mimic the natural ligands of proteins by being structurally identical to those molecules. This usually results in a pharmacological effect. As an example, VKORC1 (vitamin K epoxide reductase complex subunit 1), an enzyme responsible for converting the vitamin K 2,3-epoxide to its reduced and active form, can bind to vitamin K and to the drug warfarin [25]. By binding to VKORC1, warfarin inhibits the reduction reaction which leads to a reduced amount of available active vitamin K. Consequently, since vitamin K is involved in blood coagulation, the warfarin action leads to an anticoagulant effect by inhibiting VKORC1 action [26]. The efficacy of warfarin is compromised by *VKORC1* polymorphisms that affect the binding of warfarin to the VKORC1 enzyme. It was been shown that almost 50% of warfarin dosing variability was due to *VKORC1* polymorphisms [27]. Therefore, it is crucial to consider these polymorphisms in the dosing equation.

Many polymorphisms have been shown to influence the effect of drugs on its targets which could be helpful in the prediction of patient responsiveness to a drug. For instance, β -adrenergic receptors, present in many cells, which are targeted by catecholamines such as norepinephrine and epinephrine and by medications such as β -Blockers, have several polymorphisms that influence responsiveness to these drugs. Therefore, for some drugs, it is crucial to understand if genetic variation between individuals is responsible for differences in drug efficacy [18].

2.3 Genetic Markers and Association Studies

A genetic marker is a DNA sequence with a known location in the genome that can be used to describe polymorphisms and identify individuals. The length of a genetic marker is variable, it can range from a single base-pair change (SNPs) to hundred or thousand nucleotide base-pairs such as minisatellites and microsatellites which are characterized by having several repetitions of one DNA sequence of 10-60 and 1-6 base pairs length, respectively [28].

Genetic markers can have functional consequences that confer a higher probability of developing certain diseases. If no functional consequence is found, the marker may be located near a functional variant associated with a particular disease. In this case, both markers may be inherited if they are in linkage disequilibrium. In this scenario, genetic markers can be used indirectly to study the relationship between a disease and its genetic cause [29].

An ideal genetic marker needs to be easily identifiable and polymorphic. In addition, the genotyping process to detect that marker should have a cost-effective and easy implementation.

A series of genetic markers exist. Among all, SNPs remain the most popular because of their higher density and their wide distribution across the genome, despite being less polymorphic than other markers such as microsatellites. Recent advancements in high-throughput genotyping technologies have also contributed to the increased interest in using SNPs as genetic markers, mostly due to its amenability for high-throughput genotyping [30]. First-generation genotyping technologies amplified and evaluated DNA fragments one by one. By contrast, microarray technologies allow the sequencing of thousands or millions DNA sequences in parallel [31].

As mentioned above, genetic markers can be used as a primary source for identifying the genetic roots of complex diseases. Genetic association analysis, frequently based on SNPs, test the correlation between a disease status and genetic variants [32]. These studies can be divided into four categories: candidate polymorphism, candidate gene, fine mapping, and genome-wide scans. [33]

Candidate polymorphism studies test the association between a set of polymorphisms and a disease trait. The selection of the polymorphisms under investigation is based on prior scientific evidence, being selected the ones suggested as relevant to the development of the disease under study. Therefore, these types of studies have as goal to validate if a given polymorphism or a set of polymorphisms have direct influence in the disease [33].

Similar to candidate polymorphism studies, candidate gene studies focus on the selection of genes with known biological or functional relevance to the disease in study. This is followed by the selection of SNPs within each candidate gene that are used as identifiers of specific linkage disequilibrium blocks and therefore may not be necessarily functional. The basic premise of these studies is that the genetic variability of the genes in study is captured by the selected SNPs [34].

Fine mapping studies differ from candidate polymorphism and candidate gene studies. Their aim is to identify precisely where in the genome is located the disease-causing variant. For this, it is necessary to know with high confidence the exact location of the genotyped SNPs and it is also required to distinguish between SNPs in linkage disequilibrium. Additionally, rigorous quality control is necessary [35].

Genome-wide association studies (GWAS) aim at identifying associations between SNPs and a phenotype. Contrary to candidate polymorphism and candidate gene studies, GWAS involve the characterization of a larger number of SNPs, usually ranging from 100Kb and 500Kb DNA segments in partial scans to between 500Kb and 1000Kb in whole-genome scans. This leads to a more extensive and computationally exhaustive data analysis [33]. GWAS have gained popularity in recent years mostly due to its success in identifying SNPs in loci associated with various diseases [36].

Two primary association study designs exist: case-control and quantitative designs. In the first, individuals are classified by a binary categorical variable as either affected, *i.e.* cases, or unaffected, *i.e.* controls. The use of a binary variable to classify individuals is most useful when the disease trait cannot be quantitatively measured. From a researcher point of view, a quantitative design has a more easily interpretable outcome when quantitative traits are considered [37].

Independently of the study design, standardized phenotype rules must be used to ensure a consistent phenotype definition. By decreasing the variability among clinicians in the classification of individuals as case or control, these protocols are particularly important in case-control studies. Additionally, these rules minimize measurement errors associated with quantitative traits [37].

2.4 Overview of Statistical Methods for Genetic Association Studies

2.4.1 Genetic Models and Genotype Data Quality Control

Regardless of the analysis method and the type of the disease trait, the genotype data can be encoded in a variety of ways for association tests. The type of encoding chosen for the data influences the number of genotype-based groups and consequently the degrees of freedom of the test. As a result, its statistical power may be influenced [37]. For a single SNP three different genotypes exist, the dominant and recessive homozygous genotypes and the heterozygous genotype, that can be grouped according to different models such as the dominant, recessive, and additive models [38].

Each genetic model has different underlying assumptions about the genetic effects of the alleles in the disease. Consider that a given SNP has two alleles, A and B, and that allele B is the effect allele *i.e.* the allele responsible for a specific phenotype. Under the dominant model which considers that having one or more copies of the effect allele leads to the manifestation of the phenotype, the genotypes AB and BB are grouped together and encoded as 1 while the genotype AA is encoded as 0. In opposition, under a recessive model in which it is hypothesized that two copies of allele B are necessary for the manifestation of the phenotype, genotypes AA and AB are encoded as 0 while the genotype BB is encoded as 1. Lastly, the additive model considers a linear increase in risk for each copy of the effect allele. In this case, genotypes AA, AB and BB are encoded as 0, 1, and 2, respectively.

After encoding the genotype data a quality control analysis should be performed to guarantee the validity of the genetic association study. For this, for each genetic variant the departure from the Hardy-Weinberg equilibrium (HWE) needs to be evaluated and the minor allele frequency computed. Additionally, is necessary to evaluate if the population under study has different ancestries *i.e.* if there is population stratification.

The Hardy-Weinberg equilibrium (HWE) states that, in the absence of migration, mutation, natural selection, and assortative mating, the alleles segregate randomly in the population and as result the genotype frequencies can be calculated from the allele frequencies. A perfect HWE is considered an idealized state since it rarely occurs in reality, mostly due to the several causes responsible for its disturbance [39]. Several factors may lead to departures from HWE in controls such as genotyping problems, population heterogeneity, assortative mating, and random chance, since 5% of the tested markers will have a p-value smaller than 0.05 by chance [38, 40].

The validity of the analysis conducted in genetic association studies depends on the use of appropriate controls. Theoretically, the control group should follow the Hardy-Weinberg equilibrium [39]. Considering the population relative frequency of alleles A and B, p and q , respectively, under HWE the frequency of the genotypes AA, AB, and BB, should be p^2 , $2pq$, and q^2 , respectively. By comparing the observed and expected frequencies of genotypes, using a Fisher exact test, it is possible to test HWE in the control group [38].

The p-value of the test is computed using equation 2.1 which sums the probabilities smaller than or

equal to the probability of having the observed number of heterozygotes, [41]

$$\text{p-value} = \sum_{x \in S} P(x), \quad S = \{x : P(x) \leq P(N_{AB \text{ observed}})\}. \quad (2.1)$$

In practice, since there are no standard guidelines for rejecting SNPs that are not in HWE, SNPs with a p-value smaller than 0.05 should be checked manually to understand if there are any genotyping errors [40].

After this step, it is important to filter genetic variants based on its minor allele frequency. Given the fact that the power for detecting statistical significant associations is low for rare genetic variants *i.e.* with $\text{MAF} < 5\%$, it is necessary to compute the relative frequency of the less frequent allele so that rare genetic variants can be identified and removed [42].

If the population under study has different genetic backgrounds and simultaneously different phenotypes, apparent associations between the genotype and phenotype may appear. However, these associations are usually due to differences in ancestry which leads to a higher false-positive association rate. Therefore, it should be ensured that the population under study has the same ancestry. For this, a successful analysis can be achieved by detecting population stratification through a principal component analysis (PCA) followed by a k-means clustering [43].

PCA can be done by eigenvalue decomposition of the covariance matrix of the data which allows the data to be represented by its eigenvalues and eigenvectors also referred as principal components. These principal components explain variations, among the used samples, in the alleles distribution of the tested genetic variants and can be interpreted as continuous axes of genetic variation which usually have a geographic interpretation. After performing the PCA, it is possible to reduce the number of dimensions of the data by selecting the most significant principal components that best describe the data. For this a Tracy-Widow test can be used to determine the most significant eigenvalues which allows the selection of its respective eigenvectors *i.e.* principal components. After selecting the top principal components, a k-means clustering algorithm can be applied to the data which assigns each individual to a cluster which, in this case, represents a given population. To define the optimal number of clusters to be used in the k-means methods such as the Bayesian information criteria (BIC) are used. For this, an iterative approach can be used in which for each j iteration ($j \in \{1, n\}$) the k-means algorithm with j clusters is applied to the data and the respective BIC value is computed (equation 2.2) Then, of the n models the one with lowest BIC is preferred. A more detailed description of this technique was presented by Liu et al. [44].

$$\text{BIC} = n \cdot \ln(\text{RSS}/n) + k \cdot \ln(n) \quad (2.2)$$

where:

n = number of observations

RSS = residual sum of squares

k = number of parameters estimated by the model

Table 2.1: Contingency tables for case control analyses, by genetic model.

	AA	AB	BB		AA	AB + BB		AA + AB	BB
Cases	a	b	c	Cases	a	b + c	Cases	a + b	c
Controls	d	e	f	Controls	d	e + f	Controls	d + e	f
(a) Additive Model			(b) Dominant Model			(c) Recessive Model			

Consider a given SNP with two alleles A and B and that B is the effect allele. In (a), an additive model is used and each genotype AA, AB and BB is analysed separately. In (b), it is considered that allele B is dominant. Therefore, the phenotype will be present alongside with genotypes AB and BB, for this reason they are grouped together. In (c), it is considered that allele B is recessive and, following the same line of thought as in (b), the genotypes AA and AB are grouped together.

2.4.2 Univariate Association Tests

After performing a genotype data quality control, the statistical analysis of the data can begin. In practice, the analysis consists of a series of single-locus statistical tests in which the association between each genetic marker and the phenotype under study is examined. According to the type of phenotype, categorical or quantitative, there are different statistical tests that can be used [37].

Regression analysis are generally used in the analysis of quantitative traits. More specifically, the Analysis of Variance (ANOVA) approach is commonly used. It is similar to linear regression however it uses a categorical predictor variable, the genotype classes [37].

The analysis of categorical traits uses contingency table methods. One of the most used form of the contingency table tests is the chi-square test. [37]. However, the chi-squared test is not the most appropriate when the number of samples is small because this test relies on an approximation. In opposition, the Fisher test is exact and despite being more computationally expensive is widely used [45].

The Fisher exact test tests the null hypothesis of independence between the rows and columns of a contingency table or, in this case, tests if the cases and controls are independent of the genotype class. For each table, the rows represent the cases and controls while the columns represent the genotype class. The entries of the tables represent the number of individuals defined by these two variables. Table 2.1 presents, as an example, the contingency tables used for the case/control analysis of a given genetic variant, by genetic model.

Considering a contingency table with m rows and n columns and entries a_{ij} , the Fisher exact test begins by calculating the rows and columns sums, r_i and c_j , respectively, and the total sum of the contingency table values, N (equation 2.3).

$$N = \sum_{i=1}^m r_i = \sum_{j=1}^n c_j \quad (2.3)$$

Then, with these values it is possible to calculate the probability of observing the arrangement of data in the considered contingency table, given its specific row and column sums. This probability given by the hypergeometric distribution, can be computed using equation 2.4 and is from now on referred as P_{cutoff} .

$$P_{cutoff} = \frac{(r_1!r_2!\dots r_m!)(c_1!c_2!\dots c_n!)}{N! \prod_{i,j} a_{i,j}!} \quad (2.4)$$

After this step, all matrices with row and column sums equal to the contingency table being tested must be found and their respective probabilities of being observed are computed using equation 2.4. Finally, of these probabilities, the ones less than or equal to P_{cutoff} are summed which leads to computation of the p-value of the test [46].

Usually, a p-value of 0.05 is considered statistically significant for rejecting the null hypothesis that cases and controls are independent of the genotype class. However, testing multiple hypotheses may lead to an increasing family-wise error rate (the probability of rejecting at least one null hypothesis incorrectly in a series of hypothesis tests). For this reason, to correct the problem of multiple comparisons the Bonferroni correction method can be applied in which the significance level of the test, α , is divided by the total number of hypothesis tested, m [47].

$$p - \text{value}_{\text{adjusted}} = \frac{\alpha}{m}. \quad (2.5)$$

2.4.3 Logistic Regression

Logistic regression is a regression model appropriate when the dependent variable is binary. Therefore, it is often used for the analysis of the data when categorical traits are used. This method uses a logistic function to calculate the probability of having case status given the genotype class. Given its flexibility and the fact that allows for adjustment of clinical covariates, logistic regression has been extensively used [37].

Consider y_i for ($i \in \{1, n\}$) a response variable that characterizes the individual i in a cohort of n subjects as case or control, 1 or 0, respectively. Then consider the expected value of the response variable y_i , $E(y_i) = p_i$. In generalized linear models there is a link function g between p_i and the expected value of the response variable $E(y_i)$. In the case of the logistic regression, the link function is the logit function,

$$g(p_i) = \log\left(\frac{p_i}{1 - p_i}\right). \quad (2.6)$$

Then for each subject the mean expected value of the response variable can be represented by a linear combination of explanatory variables,

$$E(y_i) = \beta X_i = \log\left(\frac{p_i}{1 - p_i}\right), \quad i = 1 \dots n \quad (2.7)$$

where β is a vector containing the regression coefficients and X_i is a column vector with the explanatory variables. In this case, each explanatory variable represents the genotype of a given genetic variant. Then, by inverting the logit function we have the logistic function that gives the expected values p_i ,

$$p_i = \frac{e^{\beta X_i}}{1 + e^{\beta X_i}} = \frac{1}{1 + e^{-\beta X_i}}, \quad i = 1 \dots n. \quad (2.8)$$

The computed expected values will range between 0 and 1 and can be treated as a probability. In this case, the probabilities p_i can be interpreted as the genetic risk of individual i for manifesting the phenotype under study.

The model parameters *i.e.* the regression coefficients are unknown and are usually estimated using maximum likelihood estimation. For this, the parameters that maximize the following likelihood function,

$$l(\beta) = \sum_{i=1}^n y_i \log(p_i) + (1 - y_i) \log(1 - p_i), \quad (2.9)$$

are estimated using an iterative approach such as the Newton-Raphson algorithm.

To validate the model built using logistic regression a leave-one-out cross-validation can be used. This process is an iterative method. For each k iterate ($k \in \{1, n\}$, where n is the number of samples) both a test and validation datasets are defined. The validation dataset contains all samples excepting the k -th sample that constitutes the test dataset. Then, for each iteration the validation dataset is tested against the model built with the training dataset.

To evaluate the performance of the model a receiver operating characteristic (ROC) curve and its respective area under the curve (AUC) are used. The ROC curve plots true positive rates (TPR) versus false positive rates (FPR) which are calculated using the following equations,

$$\text{TPR} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (2.10)$$

$$\text{FPR} = \frac{\text{False Positives}}{\text{False Positives} + \text{True Negatives}} \quad (2.11)$$

In the case of a logistic regression used to classify hypertensive individuals, the TPR also known as sensitivity refers to the proportion of cases who are correctly identified as hypertensive. In opposition, the FPR refers to the proportion of non-hypertensive individuals wrongly classified as hypertensive. However, given the fact that the output of the logistic regression model for each instance is a probability of manifesting the trait under study we cannot compute the TPR and FPR without converting this probabilities to a binary outcome. For this, it is necessary to define a discrimination threshold between the binary classes. By considering various thresholds it is possible to compute a set of TPR and FPR pairs that constitute the points of the ROC curve.

Chapter 3

Hypertension Background

3.1 Definition and Impact of Hypertension

Systemic arterial hypertension, from now on referred to as hypertension, is a progressive cardiovascular condition characterized by blood pressure elevation in the systemic arteries. The most recent guidelines for BP classification from the European Society of Cardiology [48] define the thresholds for blood pressure classification (Table 3.1).

Hypertension in adults is defined as a SBP ≥ 140 mmHg and/or a DBP ≥ 90 mmHg. Hypertension is divided in three grades: grade 1 (SBP 140-159 mmHg and/or DBP 90-99 mmHg), the most common that accounts for almost 80% of the cases; grade 2 (SBP ≥ 160 -179 mmHg and/or DBP ≥ 100 -109 mmHg); and grade 3 (SBP ≥ 180 mmHg and/or DBP ≥ 110 mmHg), the most severe form of hypertension. In addition, individuals with elevated systolic blood pressure in combination with a normal or low diastolic blood pressure are said to have isolated systolic blood hypertension [48].

The underlying cause of hypertension allows its classification as primary, also known as essential, or secondary. The first accounts for approximately 95% of all hypertension cases and it can be understood as a multifactorial condition resulting from the combination of genetic, environmental, and lifestyle factors. In opposition, secondary hypertension forms are rare and have a clear identifiable cause. [49, 50].

Despite being easily identified through BP measurement, hypertension may be misdiagnosed if differences between office BP and out-of-office BP are involved. During office visits, if BP levels are higher than normal due to a conditional response involved with being in a medical setting. This is referred to as white-coat hypertension [51]. By contrary, the opposite may also happen when office BP is lower than out-of-office BP which is seen more frequently in younger individuals, in males, in patients living in stressful conditions or with increased physical activity during the day, and in patients who smoke and have bad drinking habits. This condition is termed masked hypertension. Individuals with masked hypertension have been associated frequently with increased prevalence of cardiovascular risk factors and target organ damage mostly due to the fact that in most patients this condition is not detected [52].

Globally, it is estimated that 3.5 billion individuals have non-optimal systolic blood pressure levels, of at least 110 to 115 mmHg, and 874 million adults have hypertension, meaning that approximately 20%

of all adults are affected by this condition [2].

From 1990 to 2015, the prevalence of nonoptimal blood pressure and hypertension increased from 73.1% to 81.3% and from 17.3% to 20.5%, respectively, mostly due to population ageing and increased prevalence of obesity over the years. Nonoptimal blood pressure remains the largest risk factor contributing to the global burden of disease and it is estimated to lead to approximately 9.4 million deaths per year, of which 41% are related to cardiovascular diseases [2].

Table 3.1: Blood pressure categories in adults*

Blood Pressure Category	SBP		DBP
Optimal	<120 mmHg	and	<80 mmHg
Normal	120-129 mmHg	and/or	80-84 mmHg
High Normal	130-139 mmHg	and/or	85-89 mmHg
Grade 1 Hypertension	140-159 mmHg	and/or	90-99 mmHg
Grade 2 Hypertension	160-179 mmHg	and/or	100-109 mmHg
Grade 3 Hypertension	≥180 mmHg	and/or	≥110 mmHg
Isolated systolic hypertension	≥140 mmHg	and	<90 mmHg

* When SBP and DBP fall into different categories, it should be designated the higher BP category to define the individual's blood pressure status. Adapted from the 2013 ESH/ESC Guidelines for the management of arterial hypertension [48].

3.2 Pathophysiology of Hypertension

Circulating blood creates pressure on the walls of blood vessels. The ejection of blood into the aorta by the left ventricle results in a characteristic aortic pressure pulse, which has two important values: a maximal and a minimal aortic pressure. Following ejection, the maximal aortic pressure, systolic pressure, is reached. Shortly after, as the left ventricle relaxes and refills, the pressure falls toward its minimal value, the diastolic pressure. The difference between the systolic and diastolic blood pressures is termed pulse pressure (PP) [53].

Although these values are clinically important they are not an adequate measure of tissue perfusion. The mean arterial pressure (MAP) is the primary pressure responsible for driving blood flow into organs and as a result it is considered as the perfusion pressure seen by tissues. It is determined by the cardiac output (CO), systemic vascular resistance (SVR) and central venous pressure (CVP). Considering that CVP is usually 0 mmHg, the mean arterial blood pressure is commonly defined as the product of CO and SVR [53].

CO represents the volume of blood pumped by the heart per unit time *i.e.* the total blood flow of the circulation. It is described as the product of stroke volume, the volume pumped by the left ventricle per beat, and heart rate, the number of heart beats per unit time. Stroke volume is dependent on cardiac pumping ability and on peripheral circulatory factors that impact venous return to the heart. Those peripheral circulatory factors influence tissue blood flow by modulating changes in the diameter of venous vessels. Likewise, SVR changes involve adjustments in vessel lumen diameter. In addition, factors affecting viscosity of blood and the anatomy of the vascular network also influence SVR [50, 53].

To sum up, BP is determined by several parameters of the cardiovascular system and by the balance of arterial tone, which in turn is influenced by intravascular volume and neurohormonal systems.

Although this formula is suitable for describing short-term control of BP, long-term regulation is not as easily defined since it is necessary to consider the involvement of an integrated neurohormonal system whose elements are connected by complex interactions [50, 54, 55]. Figure 3.1 presents all the key players responsible for the maintenance of BP levels, whose mechanisms are explained in detail further below.

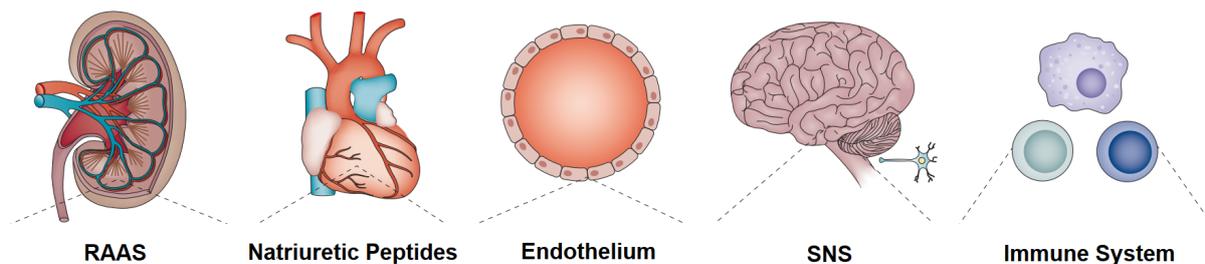


Figure 3.1: The major neuroendocrine systems involved in blood pressure regulation. Adapted from Oparil et al. [56]. RAAS = Renin-Angiotensin-Aldosterone System, SNS = Sympathetic Nervous System.

3.2.1 Sodium Homeostasis Regulation

As discussed above, intravascular volume directly affects BP. Being sodium a key regulator of blood volume, its role in the maintenance of BP levels must be taken in consideration. High serum sodium concentration activates fluid retention mechanisms in the kidney, which increases blood volume and consequently the blood pressure.

In normotensive individuals, upon increase of serum sodium concentration, BP is maintained constant through compensatory haemodynamic mechanisms. These mechanisms include reduction in renal and peripheral vascular resistance and increased production of a vasodilator named nitric oxide (NO) from the endothelium, with both being responsible for lowering the pressure exerted by the blood on arterial walls [56].

In opposition, salt-sensitive individuals show poor functioning of the normal compensatory mechanisms activated after high salt ingestion. For this reason, ingestion of a sodium load higher than 5g is followed by a marked elevation in BP within a few hours after ingestion [56].

The development of salt sensitivity is a result of endothelial dysfunction caused by genetic and environmental factors. In response to a high salt load, salt sensitive individuals set in motion molecular and biochemical events manifested as an increase in oxidative stress, reducing bioavailable NO and consequently reducing endothelium-dependent vasodilation. By promoting endothelium dysfunction, chronic salt intake is a known risk factor for the development of salt sensitivity and as a result hypertension [57].

3.2.2 Renin-Angiotensin-Aldosterone System (RAAS)

Although being present at the cellular level in many organs, the most relevant role of the RAAS occurs in the kidney by regulating its pressure-volume homeostasis. This is accomplished by maintaining renal perfusion when there is decreased extracellular fluid volume due to sodium and fluid loss or by suppressing perfusion when fluid overload is present [56].

Decreased perfusion pressure at the renal afferent arterioles, reduced sodium delivery to the macula densa and activation of renal sympathetic nerves are all stimulus in response of lowered blood pressure for the release of renin from the juxtaglomerular cells of the kidney [58]. By cleaving angiotensinogen, a hormone synthesised in the liver and secreted in the plasma, renin allows the formation of angiotensin I. Angiotensin converting enzyme (ACE), subsequently removes a peptide from angiotensin I to produce angiotensin II, which has been proven to contribute for the pathogenic role of the RAAS in hypertension [59].

Angiotensin II enhances sodium reabsorption and water retention by direct actions on the kidneys and by inducing aldosterone secretion from the adrenal glands, leading to elevation of blood pressure. It is also associated with endothelial dysfunction, a risk factor for hypertension, and profibrotic effects, caused by increase oxidative stress, that ultimately leads to injury in renal, cardiac and vascular tissues. Therefore, angiotensin II is responsible for promoting target organ damage in hypertension via these mechanisms [60].

Through metabolising vasoconstrictive angiotensin II into favourable angiotensin (1-7), ACE has an important role in modulating the pathophysiology of hypertension and cardiovascular and renal diseases. Angiotensin (1-7), contrary to angiotensin II, induces vasodilation, diuresis and natriuresis which are important mechanisms for decreasing the pressure of circulating blood on the walls of blood vessel [61].

As stated before, aldosterone stimulates sodium absorption in the cortical collecting duct, but it also promotes proliferation of vascular smooth muscle cells that contributes to vasoconstriction. In addition, aldosterone also promotes fibrosis and oxidative stress whose role in endothelial dysfunction and hypertension has been explained before in this section [62].

In summary, the RAAS is responsible for several effects, such as mediation of sodium retention, control of pressure-volume homeostasis, vasoconstriction and endothelial dysfunction, that are related with BP regulation and relevant in the origin and development of hypertension.

3.2.3 Natriuretic Peptides

Natriuretic peptides such as the atrial natriuretic peptide (ANP) and the brain natriuretic peptide (BNP) are hormones that induce natriuresis. ANP and BNP are secreted by cardiac muscle cells upon atrial and ventricular stretch due to high blood pressure. These peptides increase glomerular filtration rate and renal excretion of sodium and water, which eventually decreases blood volume, lowering blood pressure. In addition, several studies have demonstrated ANP and BNP vasodilatory effects on the vasculature which are also responsible for decreasing BP [63, 64].

ProANP and proBNP, precursors of ANP and BNP respectively, are converted into its active forms

by the atrial natriuretic peptide-converting enzyme also known as Corin. By harming the conversion of the precursors of these natriuretic peptides into its mature forms, deficiency in Corin activity has been associated with the development of hypertension and sodium and water retention [65].

Natriuretic peptide deficiency has been also observed in obesity. Considering this and the fact that obesity is a risk factor for hypertension, it can be hypothesized that natriuretic peptide deficiency might be one of the factors responsible for an increased risk of hypertension development in overweight individuals, however more studies are necessary to prove this association [66].

3.2.4 Endothelium

Endothelial cells produce a diverse range of vasoactive substances, most of them used locally to adjust vascular tone and consequently blood flow, to match metabolic needs of the surrounding tissue [67].

Nitric oxide is secreted regularly by endothelial cells in response to mechanical stimulus, such as shear stress created by blood flow, and promotes endothelium-dependent vasodilation. It is synthesized by a family of enzymes called nitric oxide synthases (NOS) that when inhibited results in a decrease of NO bioavailability, causing elevation of blood pressure and development of hypertension [68, 69].

Besides NO, other endothelial derived relaxing factors, such as prostacyclin and endothelium-derived hyperpolarizing factors, and contracting factors, such as endothelin 1 (ET-1), and prostanoids, such as thromboxane A2 and prostaglandin A2, are secreted from the endothelium. Regulation of vascular tone is determined by balancing all these factors.

In some forms of hypertension, tissue concentrations of ET-1 can be elevated however circulating ET-1 levels are not consistently increased in hypertension. Nevertheless, circulating levels may be misleading and do not accurately reflect its vascular production since ET-1 acts nearby its production site in a paracrine way. The role of this vasodilator in the pathogenesis of hypertension is still unclear, however in individuals with hypertension there is a trend for increased sensitivity to the ET-1 effects [60, 70].

As seen previously, endothelial dysfunction is a risk factor for hypertension. A combination of pressure-induced injury and multiple sources of oxidative stress is a central characteristic of endothelial dysfunction in the setting of chronic hypertension. By binding to NO, reactive oxygen species reduce NO bioavailability resulting, as stated before, in elevation of blood pressure and development of hypertension [71].

3.2.5 Sympathetic Nervous System (SNS)

The sympathetic nervous system is a major regulator of blood pressure. Its sympathetic fibres innervate various locations of the vasculature and the heart, allowing the elevation of blood pressure almost instantaneously by increasing cardiac output and through its vasoconstrictor effects [60].

SNS activity is modulated by the pressure exerted on arterial walls, this stimulus is sensed by baroreceptors that are stimulated when pressure changes occur in blood vessels. These mechanoreceptors

are located throughout the arterial tree, however the most important receptors are in the carotid sinus, located above the bifurcation of the common carotid artery and at the base of the internal carotid artery. Stretching of the carotid artery when pressure rises, increases the frequency of action potentials generated at the baroreceptors of the carotid sinus. As a result, nervous fibres projecting from these receptors provide a negative feedback mechanism in which the activity of sympathetic nerves is reduced, leading to a decrease in blood pressure [72].

Comparing SNS activity between hypertension and normotensive individuals, it has been observed an increased activity in individuals with hypertension which results, for most of them, in an imbalance between the sympathetic and the parasympathetic activity, being the first increased and the second decreased. Hyperactivity of the SNS is not merely responsible for the development of hypertension but also for its maintenance, where the severity of hypertension has been correlated with increasing levels of SNS activity [73, 74].

Excessive renal sympathetic nerve activation can cause renin release and consequent increase in sodium reabsorption in the kidney. Therefore, the renal nerves play an important role in the pathogenesis of hypertension by connecting the body fluid volume homeostasis with blood pressure regulation [75, 76].

Besides affecting the development and maintenance of hypertension through direct influence in blood pressure control, SNS hyperactivity leads to endothelial dysfunction, whose mechanisms involved in the pathogenesis of hypertension were explained before. In addition, it also causes proliferation of vascular smooth muscle cells and arterial stiffness that are responsible for creating resistance to blood flow and elevation of blood pressure [77].

There is also evidence of the association between sympathetic overactivity and the development of salt-sensitive hypertension. High salt ingestion is enhanced by SNS hyperactivity, resulting in suppression of serine/threonine-protein kinase (WNK4) expression which leads to inhibition of Na-Cl cotransporters and consequently increased sodium retention at the distal convoluted tubule [78].

3.2.6 Inflammation and the immune system

Inflammation as an essential biological response, when in excess, can lead to degenerative processes and contribute to development of hypertension and related target-organ damage. Inflammatory processes are associated with vasodilation, vascular permeability and migration of immune cells from capillary vessels into the interstitial spaces of the affect tissue. Inflammation also involves the release of biological mediators such as reactive oxygen species, nitric oxide, cytokines, and metalloproteinases [79].

Alterations in the vasculature by formation of a thick layer of arterial lumen and by promotion of vascular fibrosis, resulting in increased vascular resistance and stiffness, is caused by inflammatory cytokines. In addition, cytokines enhance angiotensinogen synthesis promoting angiotensinogen II production, resulting in increased sodium and fluid retention. Via these mechanisms, whether by vasculature alterations or changes in renal function, cytokines promote hypertension [79].

Matrix metalloproteinases by degrading extracellular matrix allow migration of immune cells, recruited

to the inflammation site, into the interstitial spaces of the affected tissue. These cells promote cell apoptosis and tissue fibrosis via enhancement of collagen and fibronectin synthesis, resulting in target-organ damage [79].

Innate and adaptive immune responses have also been implicated in the pathogenesis of hypertension induced by angiotensin II and nitric oxide antagonism in several studies using animal models. Studies using human models are limited, therefore the relevance of the immune system in causing hypertension remains an important question to be answered [79].

3.3 Risk Factors for Hypertension

Primary hypertension is a genetically complex condition in which blood pressure is influenced by multiple genetic and environmental risk factors [49].

Genetic heritability is responsible for explaining approximately 40% of BP variance in various populations, while environmental influences explain approximately other 40% of BP variance. Cultural heritability, which comprises lifestyle and dietary factors, explains an additional 10% of the blood pressure variance. The remaining 10% are still unexplained [49, 80].

It has been suggested in several family studies that blood pressure has a heritability component. A greater similarity in blood pressure was found in genetically related individuals than in genetically unrelated individuals living in the same house. This highlights the influence of genetic factors by taking into account the effect of individuals living in the same house [81]. Additionally, in twin studies monozygotic twins had systolic and diastolic blood pressure values more similar than dizygotic twins [82, 83].

Although genetic heritability estimates of BP variance in twin studies are approximately 40%, the known genetic factors, discovered in genome-wide association studies, cover only a small percentage of blood pressure variation which could highlight the fact that many genetic variants associated with blood pressure are still unknown. Several authors identified different loci associated with blood pressure [3–6]. Recently, Evangelou et al. [84] identified 535 new loci associated with risk of high blood pressure, bringing the total number of BP loci to 901. Even though the individual contribution of each genetic variant is small, their cumulative impact explains 11.2% of blood pressure variance [84]. In line with these results, an analysis performed using the UK biobank dataset point to the complex genetic influences underlying hypertension. This analysis consisted on association tests performed on genotype data of approximately 400 000 individuals in which were included 65 000 essential hypertension patients. Summary statistics of the results obtained, collected from the Oxford Brain Imaging Genetics (BIG) Server [85], a browser with GWAS results performed using the UK biobank dataset, are presented in figure 3.2. This Manhattan plot depicts that different genetic variants with p-values close to p-value threshold considered in the analysis are associated with hypertension. The range of the p-values of the significant associations may indicate that the effect size of these genetic variants is small. Given this, it is clear that the genetic factors influencing blood pressure and, consequently, hypertension include a broad number of genetic variants, each with small effect size.

Besides genetic factors, several other factors influence blood pressure, age being one of the strongest

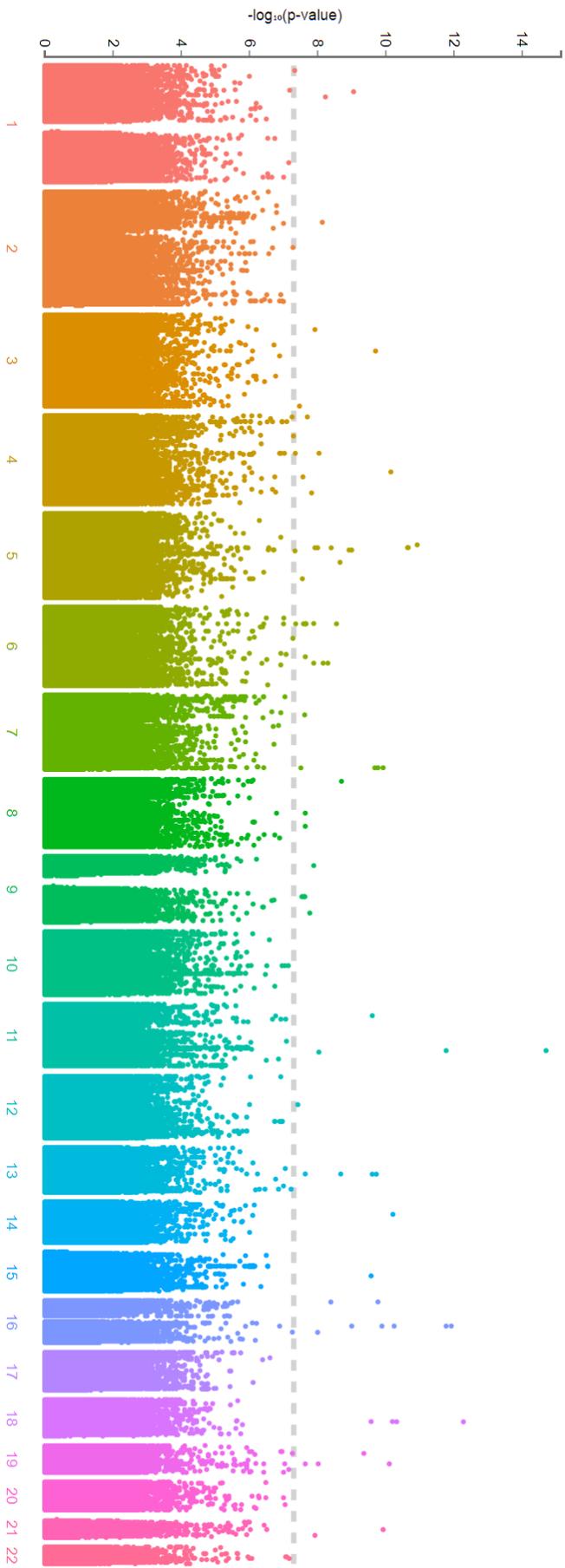


Figure 3.2: Genome wide association for essential hypertension using the UK biobank dataset. The plot represents the significance of the association between a genetic variant and hypertension, depicted by the negative log of the p-values. The threshold for suggestive association is represented using a dashed line ($\text{p-value} = 10^{-7}$). Adapted from the Oxford Brain Imaging Genetics (BIG) Server [85].

due to the vasculature changes that occur with ageing. It is known that hypertension prevalence increases sharply with age. According to a report from the American Heart Association from 2016 [86], while hypertension prevalence was 7.3% in individuals with ages between 18 and 39 years of age, the prevalence among individuals aged 60 and older was 65%.

With ageing, SBP levels increase at a higher rate than DBP levels until the midlife years. After this life stage, approximately between the fifth and sixth decades of life, DBP levels tend to reach a plateau followed by a substantial decline until the end of life. As a result, pulse pressure levels, given by the difference between SBP and DBP, increase progressively with ageing and at a higher rate after the fifth decade. These age-related changes in blood pressure are most likely explained by modifications in large artery stiffness [87].

A number of studies have identified weight as one of the most relevant risk factor for hypertension development. The 1999-2000 NHANES survey [88] showed that weight is an important predictor of hypertension: prevalence in obese individuals (body mass index (BMI) ≥ 30 kg/m²) was 42.5%, compared with a prevalence of 27.8% and 15.3% in overweight (BMI ≥ 25 kg/m² to < 30 kg/m²) and non-overweight (BMI < 25 kg/m²) individuals, respectively. Wilson et al. [89], using data from the Framingham cohort, also revealed that higher BMI is strongly associated with hypertension development. In this study, the relative risk for hypertension development, adjusted considering different variables, was 1.48 and 1.70 for overweight men and women, and 2.23 and 2.63 for obese men and women, respectively.

As discussed above, with ageing blood pressure tends to rise. However, this adverse progression of blood pressure may be avoided in young adults if a stable BMI is maintained into middle age [90].

Hypertension prevalence is also influenced by sex in an age-dependent way. NHANES survey reports and higher prevalence in males than in females until 45 years of age, yet between 45 and 64 years of age the prevalence in men and women are similar. After the sixth decade a higher percentage of women have hypertension than men [86].

Regarding ethnicity, African individuals develop high blood pressure earlier in life and have higher blood pressure values than Caucasians. Other ethnic groups such as Hispanics and Chinese have similar prevalence of hypertension to whites. Therefore, race/ethnicity is also considered as risk factor for hypertension [91].

Education seems to have an influence in hypertension development, with less educated individuals having higher hypertension prevalence. However, this association seems to be explained by dietary habits and weight differences between individuals with distinct education levels [92].

Dietary risk factors for hypertension include sodium, potassium, calcium and magnesium intakes. Although high sodium intake alone is not sufficient to develop hypertension, it has been associated with higher rates of hypertension in different populations. Higher potassium, calcium and magnesium intakes have a protective influence against the development of hypertension, being associated with lower prevalence of hypertension in several populations [93].

Excessive alcohol intake has been linked, in more than 50 epidemiologic studies, to hypertension. In brief, intake of three or more drinks increases both SBP and DBP. The BP-elevating effect of alcohol is seen both in men and women but is stronger in elderly people than in young adults [94].

3.4 Hypertension as a Risk Factor for Several Diseases

There is strong evidence for the relationship between blood pressure and cardiovascular diseases. High blood pressure is responsible for a continuous and independent increased risk for total and CVD related mortality, stroke (ischaemic and haemorrhagic), coronary artery disease (CAD), heart failure (HF), left ventricular hypertrophy, peripheral vascular disease and end-stage renal disease. Although this relationship applies to both systolic and diastolic blood pressure, elevated SBP confers in most cases a higher risk for CVD than elevated DBP [95, 96].

Individuals with hypertension or following a BP-lowering treatment at 30 years of age had a lifetime risk for cardiovascular disease of 63.3% while normotensive individuals of the same sex and age had a risk of 46.1%. In addition, cardiovascular diseases manifest approximately 5 years earlier in individuals with hypertension than in normotensive individuals [95].

Elevated blood pressure increases the relative risk for several CVD, however its impact is greater for stroke [97]. At ages 40 to 49 years, a 20 mmHg increase in systolic blood pressure is associated with a two-fold increased risk for stroke. Conversely, a small reduction of 5 mmHg in SBP levels decreases stroke and CVD mortality by 14% and 9%, respectively [98].

Renal disease is also a recognized complication of hypertension. It is estimated that hypertension is responsible for approximately 25% of all cases of end-stage renal disease, while diabetes mellitus is responsible for more than 40% of the cases [99]. The data from this study only considers a single diagnostic cause and since hypertension is usually present in individuals with diabetes mellitus, its contribution in the development of renal complications may be underestimated.

3.5 Hypertension Treatment

Once determined that the patient has hypertension or uncontrolled blood pressure, initiation of treatment is recommended either by adopting a nonpharmacological approach or a combined non-pharmacological and pharmacological approach.

Treatment strategy focus is the patient health and the reduction of adverse cardiovascular (CV) outcomes. Therefore, the adopted strategy differs according to the patient's total cardiovascular risk, which is dependent not only on the SBP and DBP values but also on the presence of other cardiovascular risk factors and on the presence of chronic kidney disease, diabetes or symptomatic cardiovascular disease [48]. Table 3.2 summarizes the recommended treatment strategies for each stratum of total CV risk.

As seen in table 3.2, a non-pharmacological approach or, in other words, lifestyle changes are recommended for all individuals with hypertension independently of its grade. In addition, lifestyle changes are also recommended for those with more than one risk factor for hypertension or those with asymptomatic organ damage, chronic kidney disease, diabetes or cardiovascular diseases either in combination with other risk factors or not. For these patients without hypertension, lifestyle changes are used primarily as a way of preventing hypertension and its adverse cardiovascular outcomes [56]. In opposition, the pharmacological approach is recommended only for hypertensive patients.

Table 3.2: Recommended treatment strategies for each stratum of total cardiovascular risk.

Other risk factors, asymptomatic organ damage or disease	High normal BP SBP 130-139 mmHg or DBP 85-89 mmHg	Grade 1 HT SBP 140-159 mmHg or DBP 90-99 mmHg	Grade 2 HT SBP 160-179 mmHg or DBP 100-109 mmHg	Grade 3 HT SBP \geq 180 mmHg or DBP \geq 110 mmHg
No other RF	No BP intervention	Lifestyle changes for several months; Then add BP drugs targeting <140/90	Lifestyle changes for several weeks; Then add BP drugs targeting <140/90	Lifestyle changes; Immediate BP drugs targeting <140/90
1-2 RF	Lifestyle changes; No BP intervention	Lifestyle changes for several weeks; Then add BP drugs targeting <140/90	Lifestyle changes for several weeks; Then add BP drugs targeting <140/90	Lifestyle changes; Immediate BP drugs targeting <140/90
\geq3 RF	Lifestyle changes; No BP intervention	Lifestyle changes for several weeks; BP drugs targeting <140/90	Lifestyle changes; BP drugs targeting <140/90	Lifestyle changes; Immediate BP drugs targeting <140/90
OD, CKD stage 3 or diabetes	Lifestyle changes; No BP intervention	Lifestyle changes; BP drugs targeting <140/90	Lifestyle changes; BP drugs targeting <140/90	Lifestyle changes; Immediate BP drugs targeting <140/90
Symptomatic CVD, CKD stage \geq4 or diabetes with OD/RFs	Lifestyle changes; No BP intervention	Lifestyle changes; BP drugs targeting <140/90	Lifestyle changes; BP drugs targeting <140/90	Lifestyle changes; Immediate BP drugs targeting <140/90

Low risk (green); Moderate risk (yellow); High risk (orange); Very high risk (red). BP = blood pressure; CKD = chronic kidney disease; CV = cardiovascular; CVD = cardiovascular disease; DBP = diastolic blood pressure; HT = hypertension; OD = organ damage; RF = risk factor; SBP = systolic blood pressure. Please note that for patients with diabetes the DBP target is 80-85 mmHg. Adapted from the 2013 ESH/ESC Guidelines for the management of arterial hypertension [48].

3.5.1 Non-Pharmacological Approach

As seen previously in section 3.3, non-genetic factors such as age, weight, ethnicity, dietary habits and alcohol intake strongly influence blood pressure and increase the risk for hypertension development. Contrary to age and ethnicity, the other non-genetic risk factors for hypertension are modifiable. Therefore, appropriate lifestyle changes are an important approach to prevent hypertension, particularly in adults with elevated BP, and to treat hypertension in hypertensive patients, usually in combination with pharmacological therapy [100–102].

Recommended lifestyle changes begin with reduction of salt intake, adoption of a healthier diet, restriction of alcohol consumption and weight reduction alone or in combination with regular physical

activity [103]. Additionally, given its association with cardiovascular risk, smoking cessation is also strongly recommended [48].

Reduction of Salt Intake

As seen in section 3.2, salt intake is associated with blood pressure levels and excessive salt ingestion can lead to the development of hypertension. Thus, control of salt intake is of most importance when considering non-pharmacological approaches for the prevention and treatment of hypertension.

In many countries, the usual salt intake ranges between 9 and 12g per day [104]. According to the World Health Organization (WHO), salt intake should be reduced to a daily intake of 5g or less [48]. In line with the WHO recommendation and the most recent studies that have demonstrated that sodium reduction leads to decrease of BP in normotensive and specially in hypertensive individuals, 1% vs 3% respectively [103, 105], the ESH/ESC guidelines also recommend 5-6 g/day of salt [48]. Nonetheless, a reduction of salt intake to 3 g/day would lead to a greater effect than the current recommendations. MacGregor et al. have shown that reducing salt intake to 3 g/day would reduce SBP, DBP, stroke deaths and ischemic heart disease death by 7.5%, 4.2%, 33% and 25%, respectively. In contrast, a reduction of salt intake to 5 g/day would lead to a reduction in SBP, DBP, stroke deaths and ischemic heart disease death of 5.0%, 2.8%, 32% and 17%, respectively [104]. It has also been shown that sodium reduction is a more effective BP reduction approach in black individuals, older people and patients with diabetes or chronic kidney disease [105].

Restriction of Alcohol Consumption

Epidemiology studies have demonstrated the association between alcohol intake and the raise of both systolic and diastolic blood pressure [106, 107]. Hence, being alcohol consumption one of the reversible risk factors contributing to hypertension, intake reduction is a crucial step in the maintenance of BP levels and prevention and treatment of hypertension.

It has been shown that reduction of alcohol consumption is responsible for lowering BP in individuals who drank more than three drinks per day. This association was stronger in those consuming six drinks/day. For these individuals, a reduction of alcohol intake of 50% lead to a reduction of systolic and diastolic blood pressure of 5.50 and 3.98 mmHg, respectively [108].

Thus, the ESH/ESC Guidelines recommend keeping alcohol intake to 20-30 and 10-20 g per day for hypertensive men and women respectively [48].

Dietary Changes

Several diets such as carbohydrate-restricting diets [109], diets rich in protein [110], vegetarian diets [111] and Mediterranean diets [112] have been associated with lower blood pressure values. It is advised for hypertensive individuals the adoption of a healthier diet consisting in a higher consumption of fresh fruits and vegetables, fibre, lean proteins, and a lower consumption of foods rich in saturated fats [48].

The value of the Mediterranean diet has been shown in different studies. In recent years, this type of diet has been reported to have a protective effect against cardiovascular diseases [113, 114]. In line with this, Estruch et al. have demonstrated that high cardiovascular risk individuals adopting a Mediterranean diet supplemented with olive oil had a reduction of cardiovascular events of 31% [114].

The effect size of dietary changes in reducing blood pressure is substantially increased when in combination with other lifestyle changes such as exercise or weight loss. In overweight or obese individuals with high blood pressure, the adoption of DASH (Dietary Approaches to Stop Hypertension) diet alone or in combination with supervised exercise lead to 11.2/7.5 mmHg (SBP/DBP) and 16.1/9.9 mmHg (SBP/DBP) blood pressure reduction, respectively [115]. Therefore, the current recommendations point to dietary changes combined with regular physical activity as a way of decreasing blood pressure [48].

Weight Reduction

One of the identified risk factors for hypertension is excess body weight and its effect on blood pressure has been demonstrated before [116]. The effect of weight reduction in blood pressure as also been confirmed in a recent meta-analysis [117]. Hypertensive individuals who experienced a weight reduction of 5.1 Kg improved its blood pressure levels with a reduction of 4.44 mmHg and 3.57 mmHg of systolic and diastolic blood pressure, respectively [117]. As a result, weight loss is a fundamental recommendation for hypertensive patients [48].

Weight reduction should be accomplished by adopting a combined approach of regular exercise and dietary changes. For those individuals, whose traditional methods used for weight reduction are not sufficient, pharmacological interventions or even bariatric surgical procedures, which has been associated with reduction of cardiovascular risk in obese patients [118], can be considered [48].

Regular Physical Activity

Different studies have demonstrated the BP-lowering effect of physical activity, especially in aerobic exercise but also in static isometric or dynamic resistance training [119, 120].

The stronger effect of aerobic exercise on blood pressure reduction was shown in a recent meta-analysis. Cornelissen et al. have shown that aerobic training was associated with a blood pressure reduction of 6.9/4.9 mmHg (SBP/DBP) in hypertensive patients [119]. In addition, it was shown that aerobic exercise is associated with lower CV risk by favourably affecting different cardiovascular risk factors [119]. Thus, for hypertensive individuals, it is recommended by the ESH/ESC guidelines the participation in aerobic exercises of moderate intensity in at least five days per week [48].

3.5.2 Pharmacological Approach

As discussed previously, a pharmacological approach is recommended for hypertensive patients usually in combination with a nonpharmacological approach.

Over the years, several antihypertensive medications have been used for blood pressure lowering and the evaluation of its efficacy and its benefits in reducing cardiovascular outcomes has been a topic of

study. The ESH/ESC guidelines confirm that, although a plethora of different of different classes of pharmacological agents exists, diuretics, β -adrenoreceptor blockers, calcium-channel blockers, angiotensin-converting enzyme inhibitors, and angiotensin receptor blockers are all appropriate for reducing blood pressure, due to their similar efficacy that differs only by a small percentage in the overall population [48]. Given this, the choice of drug should take into consideration its contra-indications and the fact that some pharmacological classes show greater efficacy in individuals with special conditions (Table 3.3). For example, in hypertensive pregnant women antihypertensive treatment with ACE inhibitors and angiotensin II receptor blocker should be avoid because of increased risk of congenital malformations [121]. Alternatively, methyldopa (a centrally acting agent), labetalol (β -adrenoreceptor blocker) and nifedipine (calcium-channel blocker) are recommended [122].

Table 3.3: Drugs for the management of hypertension preferred in specific health conditions.

Condition	Drug
Asymptomatic organ damage	
LVH	ACE inhibitor, CCB, ARB
Asymptomatic atherosclerosis	CCB, ACE inhibitor
Microalbuminuria	ACE inhibitor, ARB
Renal dysfunction	ACE inhibitor, ARB
Clinical CV event	
Previous stroke	Any agent effectively lowering BP
Previous myocardial infarction	BB, ACE inhibitor, ARB
Angina pectoris	BB, CCB
Heart Failure	Diuretic, BB, ACE inhibitor, ARB, mineralocorticoid receptor antagonists
Aortic aneurysm	BB
Atrial fibrillation, prevention	Consider ARB, ACE inhibitor, ARB, BB or mineralocorticoid receptor antagonists
Atrial fibrillation, ventricular rate control	BB, non-dihydropyridine CCB
ESRD/proteinuria	ACE inhibitor, CCB
Other	
ISH (elderly)	Diuretic, CCB
Metabolic syndrome	ACE inhibitor, ARB, CCB
Diabetes mellitus	ACE inhibitor, ARB
Pregnancy	Methyldopa, BB, CCB
Blacks	Diuretic, CCB

ACE = angiotensin-converting enzyme; ARB = angiotensin receptor blocker; BB = β -adrenoreceptor blocker; CCB = calcium-channel blocker; BP = blood pressure; CV = cardiovascular; ESRD = end-stage renal disease; ISH = isolated systolic hypertension; LVH = left ventricular hypertrophy. Adapted from the 2013 ESH/ESC Guidelines for the management of arterial hypertension [48].

Efficacy of antihypertensive medication is also influenced by ethnicity. For black individuals the response to β -blockers and ACE inhibitors is lower than in patients of other ethnic groups [123]. Thus, for these individuals, calcium-channel blockers and diuretics are recommended as a first line of treatment

either in monotherapy or in combination therapy [48].

Furthermore, regarding age and gender, current evidence demonstrates that reduction of blood pressure through pharmacological treatment is not influenced by any of those two factors [124, 125].

Diuretics

Diuretics, particularly thiazide diuretics, have remained for decades a first-line class of drugs for hypertension management. Regarding its molecular structure, a benzothiadiazine ring is only present in thiazide-type diuretics, whereas thiazide-like diuretics lack this molecular structure. Even though having different structures, thiazide diuretics and thiazide-like diuretics are among the groups of drugs that have a primary action in the distal tubule. Both subclasses of diuretics promote natriuresis by inhibiting sodium and chloride co-transporters in the renal tubules, which leads to a decrease in blood volume and consequently to blood pressure decrease [126].

It has been discussed the efficacy of hydrochlorothiazide, the most commonly used thiazide-type diuretic. In a recent systematic review, it was demonstrated that thiazide-like diuretics, indapamide and chlorthalidone, are more effective than hydrochlorothiazide in reducing blood pressure at similar prescribed doses [127]. Despite being more effective, both these diuretics subclasses have possible adverse effects related with electrolyte disturbances, such as hypokalaemia and hyponatraemia that can lead to cardiac arrhythmias and confusion, seizures and coma, respectively [48]. However, these potassium imbalances can be minimized or even corrected by taking potassium supplements or potassium-sparing agents that do not promote potassium secretion [56].

β -Adrenoreceptor Blockers

β -adrenoreceptor blockers work by blocking the binding of norepinephrine and epinephrine to β -adrenoreceptors. Consequently, normal sympathetic activity is inhibited resulting in a decrease in cardiac output and heart rate which is responsible for lowering blood pressure [128].

This class of drugs are mostly used by hypertensive patients who have suffered a previous myocardial infarction and in patients with heart failure or reduced ventricular rate control. However, although recommended by the ESH/ESC guidelines, some studies point to an inferior effect in reducing cardiovascular outcomes in hypertensive individuals as compared to other hypertensive drugs [129]. Additionally, this class of drugs tend to increase body weight [130].

These disadvantages of traditional β -adrenoreceptor blockers seem to not be shared by newer vasodilators β -blockers such as celiprolol, carvedilol and nebivolol [131]. However, large trials with hypertensive patients to assess if these newer vasodilators β -blockers have better clinical outcomes are needed [56].

Calcium-Channel Blockers

Calcium-channel blockers (CCBs) have been widely used for hypertension management. They act by blocking calcium entrance in the cell leading to a vasodilatory effect due to relaxation of vascular smooth

muscle. This vasodilatory effect decreases systemic vascular resistance and consequently lowers blood pressure [132].

Three distinct calcium-channel blockers subclasses exist such as phenylalkylamines, benzothiazepines and dihydropyridines, being the last the one with a more potent vasodilatory effect [132].

The early debate concerning CCBs efficacy and associated risks has been silenced by recent evidence showing that calcium-channel blockers are associated with a risk reduction for heart failure of 19% [133]. It was also demonstrated in other studies that calcium-channel blockers were not inferior to other drugs in reducing heart failure risk [134, 135].

A clear advantage of using CCBs for BP lowering is that it can be used in combination therapy with all other drug classes used as a first choice for hypertension treatment. In opposition, possible adverse effects of calcium-channel blockers are peripheral oedema, especially in obese individuals, and heart rate and cardiac contractility reduction [56, 136].

Angiotensin-Converting Enzyme Inhibitors and Angiotensin Receptor Blockers

As seen in section 3.2.2, the renin-angiotensin-aldosterone system has a crucial role in blood pressure maintenance. Therefore, ACE inhibitors and angiotensin II receptor blockers are both among the drug classes frequently used for hypertension treatment mostly due to their inhibitory effect on RAAS components.

ONTARGET investigators have demonstrated that in individuals with vascular disease or high-risk diabetes ACE inhibitors and angiotensin receptor blockers were both equivalent in reducing blood pressure and cardiovascular outcomes [137]. Additionally, both classes increase regression to normoglycemia in individuals with impaired fasting glucose levels which supports current recommendations for these pharmacological classes in individuals with diabetes mellitus or with predisposition for diabetes development [48, 56].

Although considered first-line antihypertensives, both these classes have potential adverse effects. ACE inhibitors have been associated with kidney function impairment, elevated potassium levels and cough. In addition, angioedema, characterized by swelling of the deep layers of the dermis or submucosal tissue due to vascular leakage, could occur, particularly in black individuals [138]. In turn, angiotensin II receptor blockers are also associated with elevated potassium levels and kidney function reduction. However, cough or angioedema have not been associated with this drug class [56].

3.5.3 Monotherapy and Combination Therapy

Wald et al. has clearly demonstrated that combination therapy is more effective in reducing blood pressure in hypertensive patients than monotherapy. Their results showed that a combination therapy of two drugs of different classes results in a 5-fold reduction in blood pressure as compared with doubling the dose of one drug [139]. Consequently, adopting a combination therapy as the initial approach leads to higher odds of achieving the target BP value and to a reduced risk of treatment discontinuation due to lower patient adherence [140]. Nonetheless, monotherapy is preferred for patients with low to

moderate cardiovascular risk and mild BP elevation, since the potential side-effects associated with the combination of more than one drug do not justify the greater BP reduction [48]. Therefore, the ESH/ESC guidelines recommend combination therapy for patients with high to very high cardiovascular risk or patients with marked BP elevation. In addition, combination therapy is also recommended for patients who do not respond to monotherapy even at full dose. Figure 3.3, represents the different therapeutic strategies used to achieve target BP [48].

Since different combination therapy strategies can be used, the choice of the class of antihypertensive drugs must consider that certain combinations are more suitable than others. When adopting a combination therapy, it is important to consider the additive and adverse effects of each combination. Preferred drug combinations involve one of the following drug classes: thiazide diuretics, calcium-channel blockers and angiotensin converting enzyme inhibitors or angiotensin II receptor blockers; that can be combined either in pairs or in groups of three [48]. In contrast, the combination of two different blockers of the renin-angiotensin system, ACE inhibitors and angiotensin II receptor blockers, is not recommended [48]. ONTARGET investigators have shown that combination therapy of one drug of each of the previously mentioned class of drugs was associated with more adverse events, particularly with increased risk of renal impairment [137].

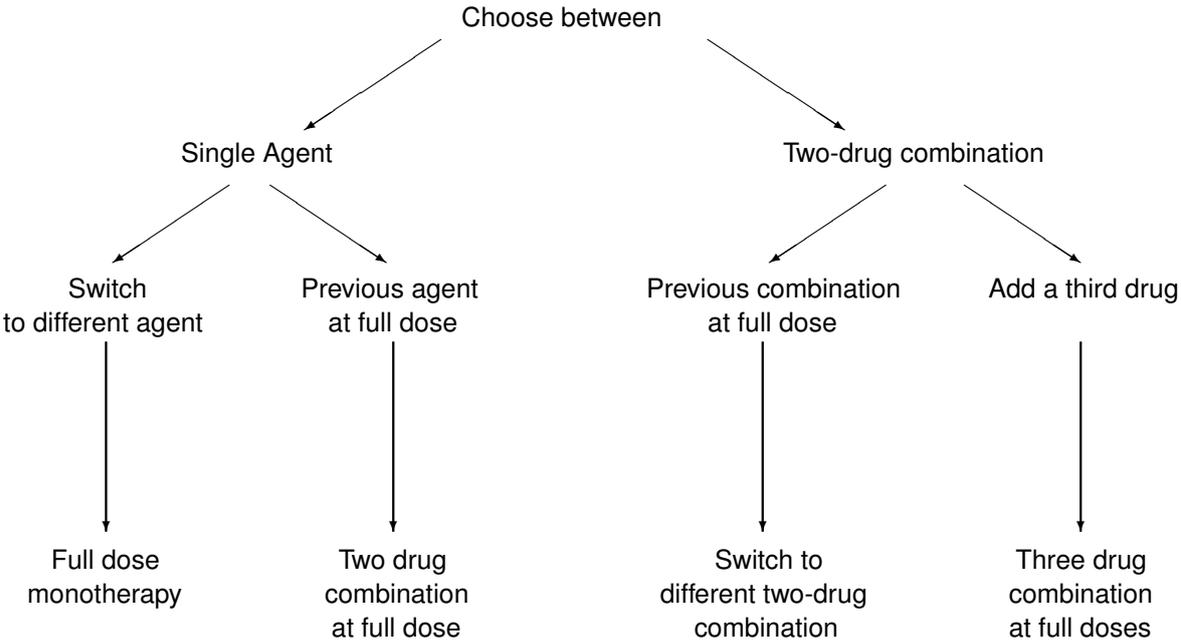


Figure 3.3: Recommended therapeutic strategies used to achieve blood pressure target. Adapted from the 2013 ESH/ESC Guidelines for the management of arterial hypertension [48].

3.5.4 Treatment-Resistant Hypertension

Hypertension is defined as treatment-resistant when the target BP (<140/90 mmHg SBP/DBP) is not achieved after lifestyle changes and a combination therapy of three antihypertensive drugs of different classes, in which a diuretic is included [48, 141]. True prevalence of treatment-resistant hypertension

remains unknown. However, it has been estimated that occurs in 10-30% of hypertensive patients [141].

When diagnosing this condition, it is important to rule out if it is a case of apparent or true resistant hypertension. Poor adherence to treatment, cuff related artefacts due the use of small cuffs and consequent overestimation of blood pressure or white-coat hypertension are all common causes of apparent resistant hypertension. By contrast, true resistant hypertension may occur when pharmacological treatment is not accompanied by reduction of lifestyle factors such as obesity, excessive alcohol consumption and high salt intake. Other causes such as obstructive sleep apnoea, volume overload, undiagnosed secondary hypertension also contribute to the development of this condition [141].

In these patients, administration of a fourth or even fifth drug could lead to BP reduction to the recommended values. The PATHWAY trial aimed to evaluate the efficacy of adding a fourth non-diuretic drug to a regimen of three drugs in which a diuretic was included. It was demonstrated that mineralocorticoid receptor antagonists, particularly spironolactone, was superior as a fourth drug than other antihypertensive medication [142]. Despite being a good choice, mineralocorticoid receptor antagonists when used together with ACE inhibitors are associated with increased risk of hyperkalaemia. For this reason, closer monitoring of serum potassium levels should be made [143].

More invasive approaches can be considered for BP maintenance in individuals with resistant hypertension. Baroreflex activation therapy, characterized by electric stimulation of the carotid sinus via implanted devices, can be used to effectively reduce blood pressure in most resistant-hypertensive patients. The effects of this procedure are usually maintained over a long-term period, 22 to 53 months [144].

Catheter-based renal denervation is another invasive approach that has also been shown to lead to sustained reduction of blood pressure [145]. By disrupting renal nerves, this approach helps reducing renal sympathetic hyperactivity that has been associated with the development and progression of hypertension [146].

Nevertheless, for both of these procedures, it is still necessary to fully document its efficacy and to compare it with the best treatment currently available in long-term observational studies [48].

Chapter 4

Materials and Methods

4.1 Data Selection and Characterization

The data used in this thesis was obtained from two different sources. Genotype data was collected from the HeartGenetics biobank whose samples were gathered in collaboration with the Coimbra Hospital and University Centre. This institution contributed with the selection, the recruitment and the demographic and clinical characterization of the subjects included in the cohort.

The HeartGenetics originally contained subjects with cardiovascular disease and/or traditional cardiovascular risk factors, individuals with family history of heart failure and/or traditional cardiovascular risk factors and centenarians with no hereditary diseases or cardiovascular diseases. From this biobank, subjects who had either congenital heart disease, myocardopathy, autoimmune disease, hypothyroidism, diabetes mellitus, renal insufficiency and chronic liver disease or were under therapeutics with either anticonvulsants, antipsychotics, immunosuppressors or antiretrovirals were excluded.

From the HeartGenetics biobank, characterized by having a total of 3365 samples with genotype data of 111 genetic variants, two datasets were created to be used in different analyses. It is important to mention that not all samples of the HeartGenetics biobank have genotype data for the 111 genetic variants, this is mostly due to the fact that the genetic panel analysed changed over the years until a final genetic panel was defined. For this reason, the selection of the data used in the analyses performed in this thesis involved a careful process of data cleaning.

The first dataset, from now on referred to as HTN dataset, is a dataset with genotype data of 187 hypertensive (cases) and 230 healthy (controls) individuals. It is a subset of the HeartGenetics biobank from which was selected genotype data from the variants genetic panel of the DNArterial test (total of 57 genetic variants described in tables A.1 and A.2) because most of the samples were genotyped for these variants. After this, given the fact that most samples were not genotyped for all genetic variants a data cleaning procedure was used. For each sample and each genetic variant the missing data percentage was computed. Then, first the samples and then the genetic variants with missing data percentage larger than 10% were removed. Lastly, for each variant, missing data was replaced by the most common genotype. After all these steps, the HTN dataset yielded a total of 417 samples with genotype data from

53 variants.

The second dataset, from now on referred as PHARMA dataset, is a dataset with genotype data of 13 individuals with treatment-resistant hypertension (cases) and 309 healthy or non treatment-resistant hypertensive individuals (controls). Clinical information available for each sample allowed the indirect classification of treatment-resistant hypertensive individuals by selection of individuals under combination therapy for hypertension with 3 different classes of drugs, in which a diuretic is included, or of individuals who underwent a renal denervation procedure used as treatment for treatment-resistant hypertension. The PHARMA dataset is a subset of the HeartGenetics biobank from which was selected genotype data from genetic variants known to have pharmacogenetic associations with antihypertensive drugs (total of 24 genetic variants described in tables A.1 and A.2). A different approach for sample selection was used for this dataset, only samples without missing data were selected. This was possible because it were used less variants in this dataset as compared to the HTN dataset. Therefore, a larger percentage of samples had genotype data for all the considered variants in this dataset. This approach removes all missing data in the dataset and eliminates the necessity of using missing data imputation. Is important to mention that this approach was not used in the HTN dataset because only two samples in this dataset had genotype data for all of the 57 genetic variants selected. Table 4.1 presents detailed characteristics of the two cohorts under study.

Table 4.1: Demographic information of the two cohorts under study.

	HTN Cohort - 417 individuals		PHARMA Cohort - 322 individuals	
	Cases	Controls	Cases	Controls
Number of Subjects	187 (44.8%)	230 (55.2%)	13 (4%)	309 (96%)
Age	72.0 ± 21.4	60.4 ± 25.9	55.5 ± 13.2	71.7 ± 24.6
Sex (%)				
Female	71.8%	63.8%	66.7%	66.1%
Male	28.2%	36.2%	33.3%	33.9%

Finally, for the identification of genetic variants known to be associated with a drug phenotype, from now on referred as pharmacogenetic variants, the Pharmacogenomics Knowledgebase (PharmGKB) [147] was used, which collects and curates pharmacogenetic associations found in the literature. Two different types of searches were made on PharmGKB: per variant in which were considered the 24 pharmacogenetic variants of the DNArterial test genetic panel which are presented in tables A.1 and A.2; and per drug or class of drug in which were considered 58 antihypertensive drugs and 5 drug classes (diuretics, ACE inhibitors, angiotensin II receptor blockers, calcium channel blockers and β -blockers) identified as relevant drugs used for hypertension treatment in the most recent guidelines from the American Heart Association [148]. Then, the results from the two searches performed were joined and duplicated pharmacogenetic associations were removed. After this, it were selected associations with statistical significance, as defined in the PharmGKB database, which lead to a total of 445 statistically significant pharmacogenetic associations.

For each association between a genetic variant and a drug phenotype relevant information was

extracted by PharmGKB curators from the studies where the association was identified. This information includes the drug molecules and genetic variants (mapped to a common standard, usually dbSNP rs number) studied, the study size, population characteristics, allele frequencies and statistical values, usually p-values from the hypothesis tested. For each association is also presented a description of the association found between a given variant and a drug molecule.

From all the statistically significant pharmacogenetic associations identified it was necessary to select which were the relevant associations that should be included for analysis. An initial approach was to remove associations which had missing data, for instance associations where study characteristics were not identified. Knowing that a search per variant was made on PharmGKB, associations between those genetic variants with drug molecules not used for hypertension treatment were unintentionally included. For this reason, after removing associations with missing data, only pharmacogenetic associations between genetic variants and drugs used for hypertension or other cardiovascular conditions treatment were selected. An exploratory approach was considered in this step. Since most of the patients under antihypertensive therapy also have other cardiovascular risk factors or diseases and, consequently, take medication for these conditions, a less strict selection of the drugs under study was considered. By doing this, it is possible to verify if genetic variants with relevant pharmacogenetic associations with antihypertensive drugs are also associated with other drugs used for cardiovascular diseases. After all these steps, a total of 396 pharmacogenetic associations involving a total of 203 genetic variants and 87 drugs or classes of drugs were included. Figure 4.1 presents a diagram which depicts the flow of information between each step used for selecting pharmacogenetic associations to be analysed.

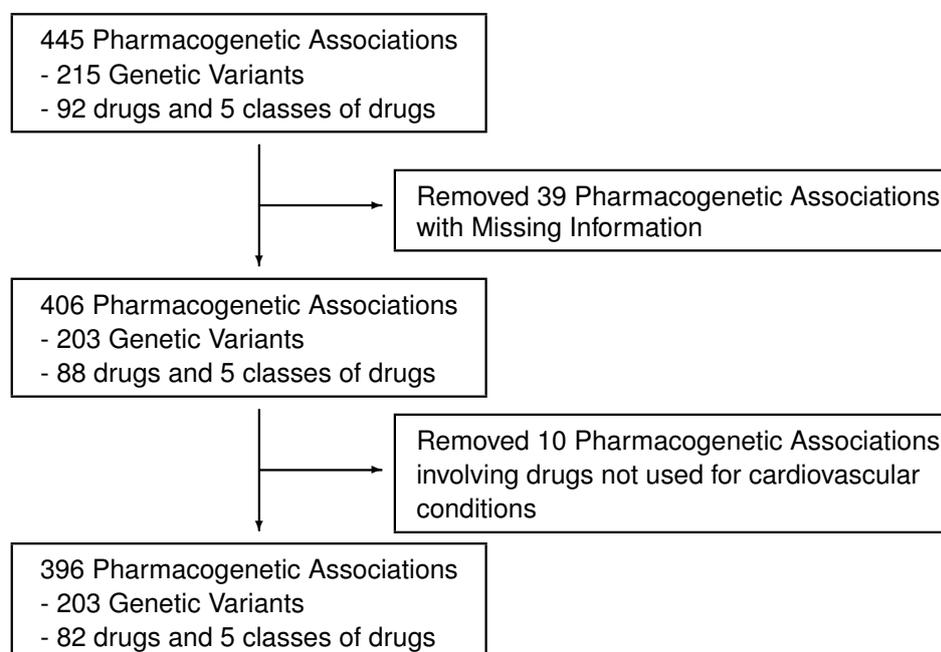


Figure 4.1: Flow diagram of pharmacogenetic associations selection.

In PharmGKB variant-drug associations are assigned a level of evidence which is based on well-defined criteria [147]. Four levels of evidence are used ranging from 1, the highest level of evidence, to

4, the lowest. Of these levels, levels 1 and 2 are divided into A and B sublevels. Therefore, for each of the 396 pharmacogenetic associations a level of evidence retrieved from PharmaGKB was also attributed.

4.2 Genotype Data Encoding and Genetic Models

Before encoding the genotype data, the HTN and PHARMA datasets were each organized into a $m \times n$ matrix, where m and n represent the number of samples and the number of genetic variants in the dataset, respectively. In these matrices, each entry i, j of the matrix represents the genotype of sample i and genetic variant j . Since each genotype is characterized by a string it was necessary to encode the data into a numerical form so that the necessary analyses could be more easily performed.

Genotype data can be encoded in a variety of ways. For each genetic variant the different genotypes (dominant and recessive homozygous genotypes and the heterozygous genotype) are encoded according to genetic models such as the dominant, recessive and additive models [38].

Each genetic model has different underlying assumptions about the genetic effects of the genotype in the phenotype. Therefore, using the best-fitting genetic model is crucial when genetic variants are used for prediction of a given phenotype [149]. For this reason, the three genetic models were used to encode the genotype data of the HTN and PHARMA datasets, which lead to three different genotype data codifications for each one of the datasets.

4.3 Genotype Data Quality Control

For the two datasets, HTN and PHARMA, a data quality control was performed. At a first stage, the minor allele frequencies for each variant was computed. Then, it was tested if all the genetic variants were under Hardy-Weinberg equilibrium in the control population. For this, the *HWExact* function from the “HardyWeinberg” R package which performs an exact Fisher test for HWE was used.

After this analysis, genetic variants with MAF smaller than 5% were removed. Additionally, genetic markers for which there was a deviation from the Hardy-Weinberg equilibrium were flagged for future analyses if proven to have a significant association with the traits studied in this work. This approach was considered, as an alternative of removing these variants, because deviations from the HWE occur frequently mostly due to the fact that this concept is based on multiple assumptions that are not always verified.

Another data quality control procedure is the analysis for population stratification. The approach used in this thesis for this analysis consists of a principal component analysis (PCA) followed by a k-means clustering. For this, the *eigenstrat* function from the “AssocTests” R package and the function *kmeans* from the “Stats” R package were used to perform the PCA and the k-means, respectively.

4.4 Univariate Association Analysis for the Identification of Relevant Genetic Variants in Hypertension

Association tests are used to study the association between genetic variants and phenotypes. The association with the hypertension trait was studied for each one of the genetic variants, not removed during the quality control procedure, considered in the HTN cohort. For this, three Fisher exact tests were performed per variant, one for each genetic model used, using the function *fisher.test* from the “Stats” R package. For each association test an adjusted p-value of 0.001 computed using the Bonferroni correction was considered as statistically significant for rejecting the null hypothesis that cases and controls are independent of the genotype class.

4.5 Construction of a Genetic Risk Score for Hypertension

In this thesis, using the *glm* function from the “Stats” R package a logistic regression applied to the HTN dataset was used which allowed the building of a genetic risk score model for hypertension. To validate the model built a leave-one-out cross-validation was performed. To evaluate the performance of the model a receiver operating characteristic curve also referred as ROC curve and its respective area under the curve (AUC) were used. This procedure was done using the function *roc* from the R package “pROC”.

In this analysis an additive, a dominant, and a recessive model were considered for the genotype data of the HTN dataset. In addition, a mix genetic model in which each genetic variant is encoded by a different genetic model was used to encode the HTN dataset used in the logistic regression. For this, in each iteration of the leave-one-out cross-validation, a different mix model dataset was constructed using the training data in which each genetic variant was encoded with the most adequate genetic model (additive, recessive or dominant). For the identification of the best-fitting genetic model, three Fisher exact tests per variant were performed. The genetic model that achieved the smallest p-value was selected.

4.6 Criteria for the Selection of Relevant Genetic Variants with Pharmacogenetic Associations

As previously mentioned a search for genetic variants with significant pharmacogenetic associations was performed in PharmGKB from which 396 associations were identified. After extracting the relevant information from PharmGKB, different criteria were used to select the most relevant genetic variants that should be included in a genetic panel for a pharmacogenetic test for hypertension.

Firstly, genetic variants with at least one pharmacogenetic association with a level of evidence equal or lower than 2 were selected. Those levels of evidence 1 and 2 correspond, respectively, to a high and a moderate level of evidence. In the first level are included variant-drug associations from which the

majority of evidence confirms the association. In addition, these associations are replicated in different studies with a significant p-value and strong effect size. By comparison, the second level of evidence includes associations that are replicated in different studies although some studies may not show statistical significance or a small effect size.

Then, from the list of pharmacogenetic associations identified a subset of genetic variants with a FDA drug label containing pharmacogenetic information approved by the FDA were selected [150].

Additionally, a third criterion was considered, where selected genetic variants with at least one pharmacogenetic association with a p-value smaller than 0.001 were also selected. This threshold was used because it was defined that it was preferable to have a genetic panel with genetic variants whose pharmacogenetic associations had a strong statistical significance than including genetic variants with associations with weaker statistical significance. To conclude, given the high risk of complications in patients with treatment-resistant hypertension and the fact that optimization of medical treatment still remains an important concern, it were also selected genetic variants with known pharmacogenetic associations identified in cohorts whose individuals had treatment-resistant hypertension.

Figure 4.2 represents a decision diagram with the different criteria considered for the selection of genetic variants with relevant pharmacogenetic associations.

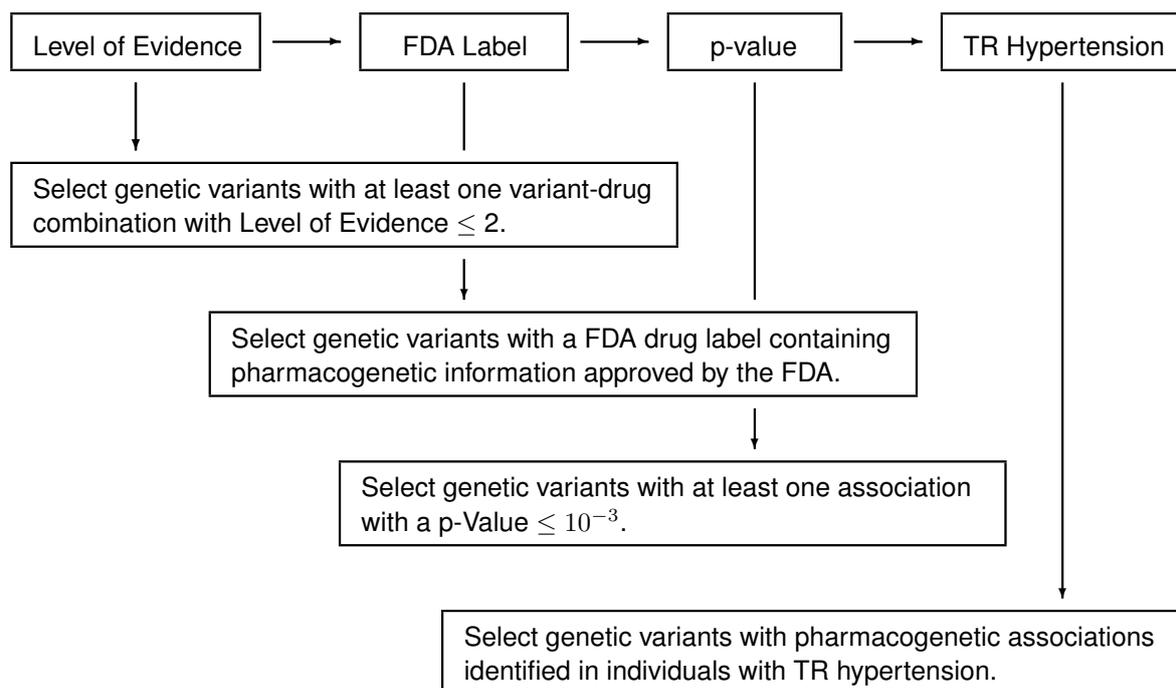


Figure 4.2: Decision diagram with the different criteria considered for the selection of genetic variants with relevant pharmacogenetic associations. TR = treatment-resistant.

4.7 Analysis of the Genetic Influence on Treatment-Resistant Hypertension

The PHARMA cohort was used to assess the underlying genetic background of treatment-resistant hypertension.

Firstly, a genetic risk score for each individual to compare cases and controls for all the 24 genetic variants with known pharmacogenetic association with antihypertensive drugs was computed. For this analysis three different genetic models were considered (additive, dominant and recessive models) in which the genotypes of each variant were numerically encoded as explained in section 4.2. The genetic risk score for each individual was computed using an unweighted approach in which the score is the direct sum of the number of risk alleles. For instance, considering the additive model in which genotypes were encoded as 0,1, or 2 which indicates the number of risk alleles per genotype, the genetic risk score of a given individual is computed by summing the coded genotypes.

An additional approach was used to compute the genetic risk score. Similar to the procedure used in section 4.5 a mix dataset in which each genetic variant was encoded with its most adequate genetic model was constructed. Then, the genetic risk score of each individual was computed using the same unweighted approach mentioned previously.

The genetic risk score of each cohort subject was then used to compare the controls with the cases. For this, a Mann–Whitney U test was used to compare the distribution of genetic risks scores between the two groups by testing the null hypotheses that is equally likely that a random observation, in this case the genetic risk score, in one group will be less or greater than a random observation of the second group.

The Mann–Whitney U test [151] begins by joining all the observations in one set which are then assigned with numerical ranks. Considering the case where there are a group of observations with equal values, the assigned rank to each observation is the midpoint of the unadjusted rankings. Then, the ranks of each group are summed separately by group, and the null hypothesis is tested by calculating a statistic for each of the considered groups using the following formula:

$$U = R - \frac{n(n+1)}{2} \quad (4.1)$$

where:

n = number of observations;

R = sum of the ranks of the considered group.

Knowing that for large samples U is approximately normally distributed, the minimum value of the U calculated for each group is used to compute the standardized value or Z-value given by equation 4.2 which gives the p-value of the test. For this analysis, a Bonferroni correction was also used to define the significance level of the tests which was calculated using equation 2.5.

$$z = \frac{U - \frac{n_x n_y}{2}}{\sqrt{\frac{n_x n_y (N+1)}{12}}} \quad (4.2)$$

Finally, associations tests for each of the 24 genetic variants in the PHARMA dataset to study its association with the treatment-resistant hypertension trait were also performed. For this, three Fisher exact tests were performed per variant, one for each genetic model used (dominant, recessive, and additive).

Chapter 5

Results

5.1 Genotype Data Quality Control

Quality control results and association statistics per genetic model for each genetic variant are presented in table B.1. Of the 53 genetic variants analysed, 4 had a minor allele frequency less than the considered threshold of 5% (rs75770792, rs111253292, rs5063, and rs5068). As a result, these variants were excluded from the posterior analyses performed on the HTN cohort. In addition, 2 genetic variants were not in Hardy-Weinberg equilibrium (rs11091046 and rs1403543). The results of another quality control analysis can be seen in Figure 5.1. By doing a PCA and a k-means it was possible to verify that, in both cohorts, only one cluster is present and consequently there is no population stratification.

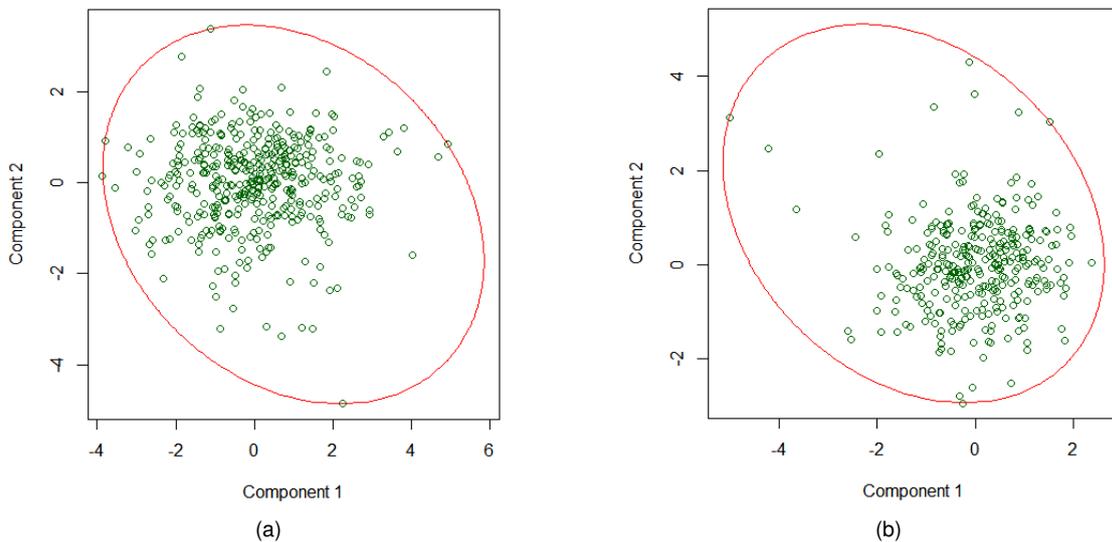


Figure 5.1: Population stratification analysis results using a PCA followed by a k-means. Scatter plot of principal component axis one (component 1) and axis two (component 2) based on genotype data. Individual data points are clustered in groups according to their genotype which indicates their ancestry. (a) HTN cohort. (b) PHARMA cohort.

5.2 Genetic Variants Associated with Hypertension

An univariate analysis was performed on the HTN cohort to identify genetic variants associated with the hypertension trait. Association statistics per genetic model are presented in table B.1.

Of the set of genetic variants considered, a total of 6 variants showed a trend for association with hypertension (p -value ≤ 0.05). The SNP rs4646994 from the *ACE* gene showed a trend for association in all genetic models, being stronger in the additive model with a p -value = $1.3E-03$. This trend for association was also verified for the variants rs3918226 (*NOS3*), rs5065 (*NPPA*), and rs6693954 (*REN*) for both additive and dominant models. In addition, the locus rs880054 (*WNK1*) considering a dominant model and locus rs72811418 (*CYBA*) considering an additive and a recessive model (lower p -value for this model) were shown to have a trend for association with hypertension. However, of the 49 genetic variants analysed, considering as significant a p -value ≤ 0.001 (Bonferroni adjusted p), none showed a statistical significant association with hypertension.

5.3 Genetic Risk Score Model Evaluation for Essential Hypertension

Logistic regression was used to create a genetic risk score to evaluate the impact of the combination of all 49 genetic variants considered in the HTN cohort on hypertension risk. An additive, a dominant and recessive model were considered for this analysis. In addition, a mix model (for more details see section 4.5) was also used. The discriminative power of the logistic regression of each model considered was measured by the AUC of the ROC curve. Figure 5.2 presents the 4 ROC curves obtained and the respective AUCs. Although being higher when using a mix model (AUC = 0.568), the AUCs showed that the logistic regression achieved a modest result using the 49 genetic variants.

5.4 Genetic Variants with Relevant Pharmacogenetic Associations

Using the selection criteria presented previously in section 4.6, from a total of 396 pharmacogenetic associations involving 203 genetic variants and 87 drugs or classes of drugs were selected 33 genetic variants with relevant pharmacogenetic associations¹. Of these 33 genetic variants, 6 variants (rs1799752, rs4961, rs1801253, rs5051, rs13306673) are present in the current genetic panel of the HeartGenetics test for hypertension.

Considering the first criterion in which are included genetic variants with at least one pharmacogenetic association with level of evidence less than or equal to 2, 17 genetic variants were selected from a total of 6 different genes. Is important to mention that one of these variants (rs4961) was selected taken into account a pharmacogenetic association with level of evidence 2B that was observed only in individuals with liver cirrhosis.

¹Detailed information relative to these genetic variants is presented in confidential appendices.

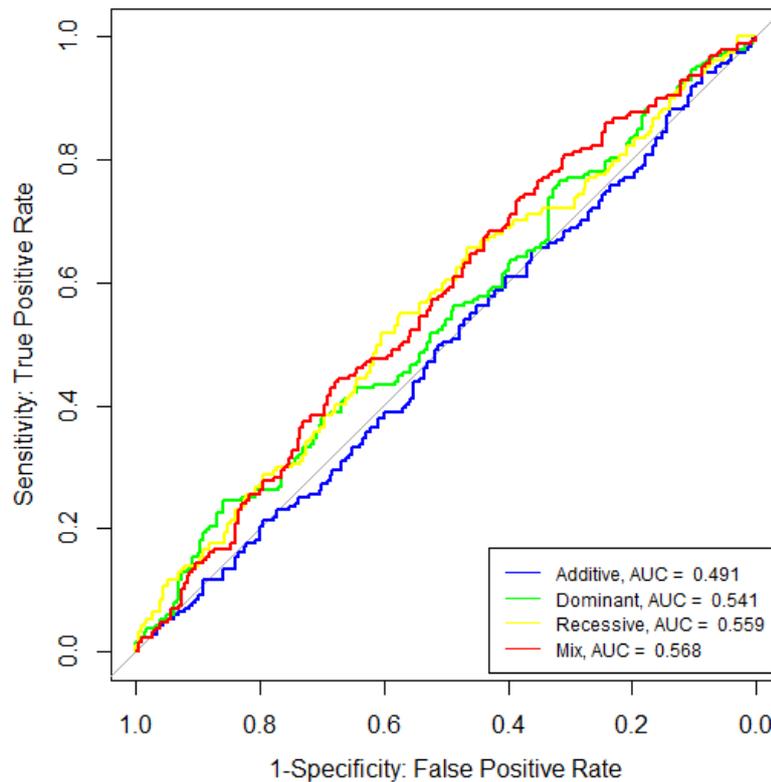


Figure 5.2: Receiver operator characteristics (ROC) curves of the logistic regressions used for hypertension discrimination using different genotype codifications: additive model, dominant model, recessive model and mix model.

The remaining genetic variants selected had pharmacogenetic associations with weak levels of evidence (level 3), from which 9 were labelled by the FDA as relevant pharmacogenetic variants and 2 had a pharmacogenetic association verified in individuals with treatment-resistant hypertension.

5.5 Genetic Risk Score Model Evaluation for Treatment-Resistant Hypertension

The genetic risk score was computed for each subject of the PHARMA cohort. Its distribution between individuals with treatment-resistant hypertension and controls was compared using a Mann-Whitney test. Figure 5.3 displays, using boxplots, the genetic risk scores, for each genetic model used, of cases and controls and the respective p-values computed using the Mann-Whitney test.

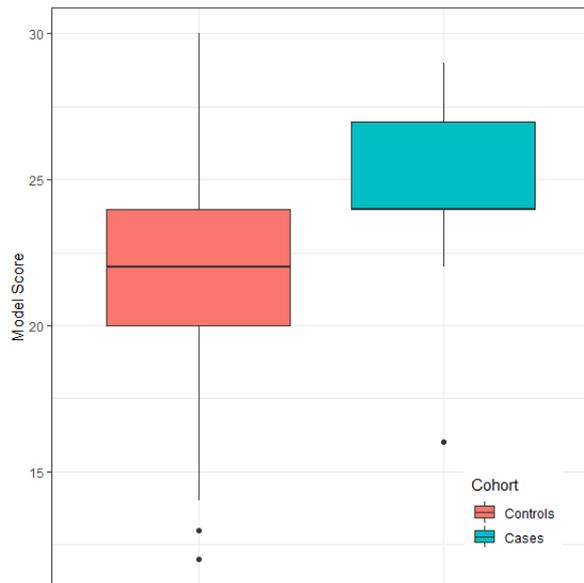
For the first analysis, in each were considered all 24 pharmacogenetic variants, a statistical difference in the risk score between the control group and the case group for all the genetic models used was found. When an additive model was considered an improved statistical difference between groups with a p-value = 0.003 was achieved. Considering this genetic model, the upper quartile of the control group is equal to median and lower quartile of the cases group which means that 75% of the individuals in this group had a higher genetic risk score for treatment-resistant hypertension than 75% of the controls. An

equivalent interpretation can also be made considering the recessive and mix models. For both these genetic models, the upper quartile of the controls group is equal to the median of the cases group which means that 50% of the cases had a higher genetic risk score than 75% of the controls.

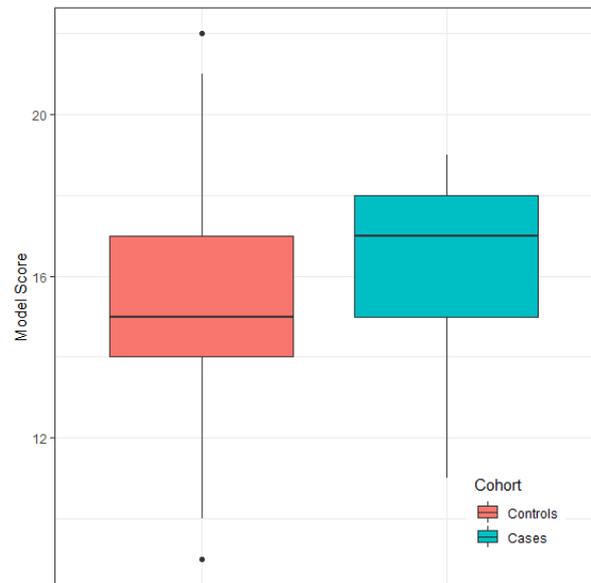
An identical second analysis was performed by considering from the HeartGenetics pharmacogenetic panel a total of six genetic variants (rs1799752, rs4961, rs1801253, rs5051, rs13306673) identified as relevant pharmacogenetic variants in section 5.4. It was found a statistical difference in the risk score between the controls group and the cases group for the additive, recessive, and mix models. Considering these models, 50% of the cases had a higher genetic risk score than 75% of the controls. When a recessive model was considered it was achieved a higher statistical difference between groups with a p-value = 0.029. Under a dominant model the non-significant p-value of 0.214 was obtained.

Overall, when all variants were considered versus the six selected variants a better discrimination between cases and controls with lower p-values was achieved.

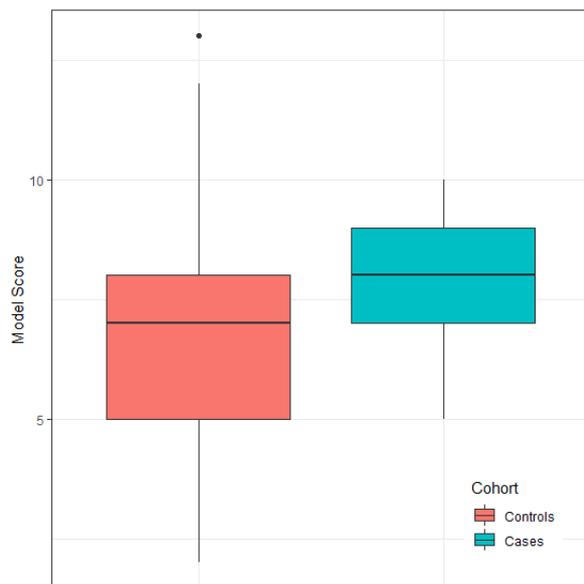
The results from the univariate analysis performed on this cohort to identify genetic variants associated with the treatment-resistant hypertension trait are presented in the confidential appendices. Of the 24 genetic variants considered, a total of 8 variants showed a trend for association with treatment-resistant hypertension (rs1801252 (*ADRB1*), rs699 (*AGT*), rs5186 (*AGTR1*), rs1048101 (*ADRA1A*), rs1801253 (*ADRB1*), rs5051 (*AGT*), rs1799722 (*BDKRB2*), and rs2228576 (*SCNN1A*)). Under the additive model the genetic variants rs1801252 (*ADRB1*), rs699 (*AGT*), and rs5186 (*AGTR1*) had a p-value ≤ 0.05 . This trend for association was also verified under a dominant model for variants rs1048101 (*ADRA1A*) and rs1801252 (*ADRB1*). When a recessive genetic model was considered this trend for association was verified for the variants rs1801253 (*ADRB1*), rs699 (*AGT*), rs5051 (*AGT*), rs1799722 (*BDKRB2*), and rs2228576 (*SCNN1A*). However, of the 24 genetic variants analysed, considered as significant a p-value ≤ 0.002 (Bonferroni adjusted *p*), none showed a statistical significant association with treatment-resistant hypertension.



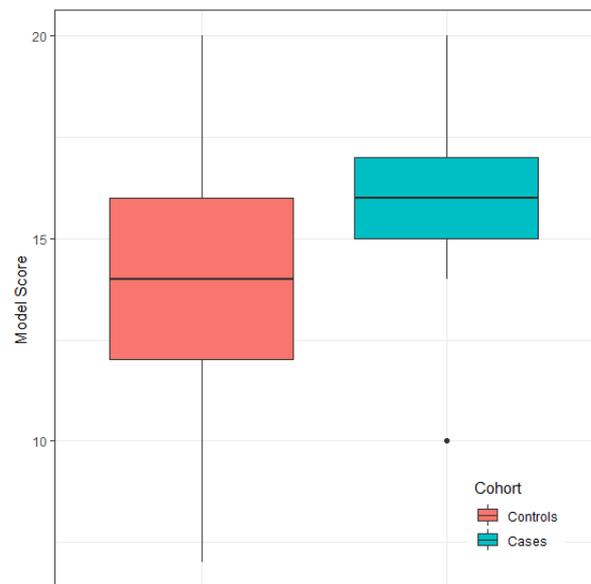
(a) Additive Model



(b) Dominant Model



(c) Recessive Model



(d) Mix Model

Figure 5.3: Boxplot of the genetic risk scores calculated using a total of 24 genetic variants for the cases and controls groups by genetic model used. (a) Additive Model: p-value = 0.003, (b) Dominant Model: p-value = 0.015, (c) Recessive Model: p-value = 0.009, (d) Mix Model: p-value = 0.005.

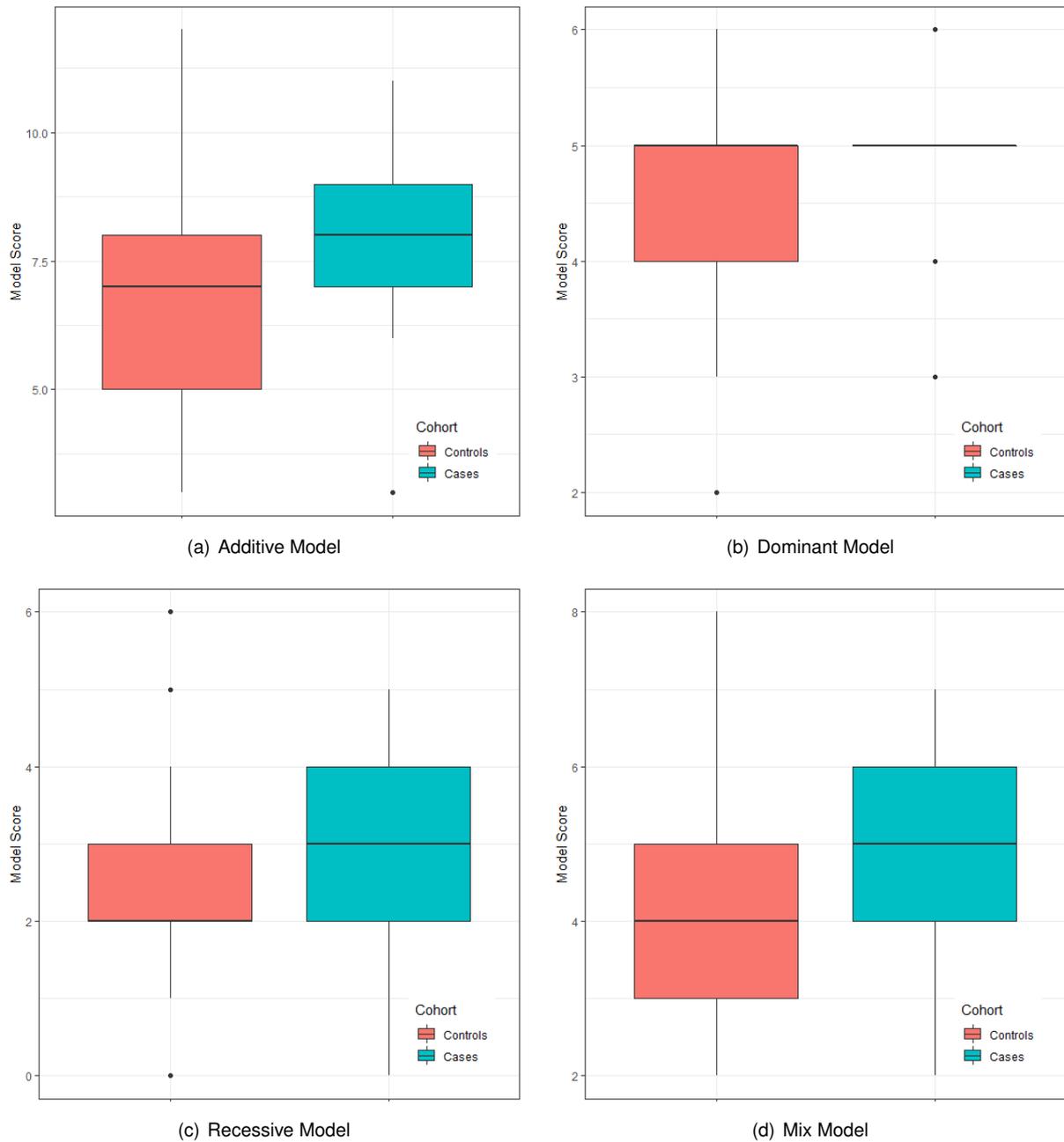


Figure 5.4: Boxplot of the genetic risk scores calculated using a total of six genetic variants for the cases and controls groups by genetic model used. (a) Additive Model: p-value = 0.035, (b) Dominant Model: p-value = 0.214, (c) Recessive Model: p-value = 0.029, (d) Mix Model: p-value = 0.035.

Chapter 6

Discussion and Conclusions

The main objective of this thesis was to develop a genetic test for hypertension with a clear clinical value. For this, two different approaches were considered. Firstly, it was evaluated the feasibility of having a genetic test with the aim of predicting hypertension genetic risk. Then, it was considered a second approach in which it was evaluated the clinical validity of a genetic test with the aim of offering a pharmacogenetic-guided treatment.

Using an in-house genetic dataset, from a total of 53 genetic variants, the variants rs4646994 (*ACE*), rs3918226 (*NOS3*), rs5065 (*NPPA*), rs6693954 (*REN*), rs880054 (*WNK1*), and rs72811418 (*CYBA*) showed a trend for association with the hypertension trait. However, none achieved statistical significance.

The *ACE* polymorphism, rs4646994 (also referred as rs1799752), has been previously mentioned in different studies in which a positive correlation of the del allele with hypertension was verified in a Indian and Chinese population, however this finding was not verified in juvenile American, Turkish or Sardinian populations [152]. These findings, in line with the results obtained in this thesis, show that the association between this *ACE* polymorphism and hypertension is not robust.

Although, the results of this work show a trend for correlation with hypertension for the locus rs880054, previously reported associations of this SNP with hypertension were not replicated in the British Genetics of Hypertension (BRIGHT) study [A6].

From the remaining six genetic variants mentioned previously, the loci rs3918226 (allele T), rs5065 (allele T), and rs6693954 (allele A) have been associated with hypertension with OR = 1.54 (95% CI: 1.37-1.73), OR = 0.94 (95% CI: 0.88-1.00), and OR = 1.98 (95% CI: 1.40-2.80), respectively [153–155]. Lastly, the polymorphism rs72811418 of *CYBA* gene was shown to be associated with hypertension by Moreno et al., however the findings were not replicated in a different cohort [156].

The differences observed between the results obtained in this work and current evidence may be the consequence of using a cohort with a small sample size, at least in comparison with the sample sizes currently used in genome-wide association studies.

Regarding the genetic risk score model created to predict hypertension risk, it can be said that it did achieved a modest result which is confirmed by the AUCs of the 4 models created using an additive,

a dominant, a recessive and mix genetic model. However, the results obtained seem to be in line with the state of the art. Held et al. compared the performance of the logistic regression method to predict the risk of hypertension using non-genetic factors alone and in combination with genetic data [157]. Reported AUCs were 0.785 and 0.778 for the model considering non-genetic factors alone and the model considering both genetic and non-genetic factors, respectively. Although, were only considered 10 causal genes for the genetic component of the model used in this study, the results reported point that the inclusion of genotype data for risk prediction may not improve substantially the prediction ability of these models. Additionally, these results clearly show, as discussed previously in this work, that non-genetic factors play a crucial role in hypertension development which also limits the accuracy of a genetic test for hypertension risk that uses only genotype data.

As mentioned previously, the analysis of the UK biobank dataset shows that the genetic factors influencing hypertension include a broad number of genetic variants, each with small effect size. These results in line with the results obtained in this work point that it is not clear that developing a genetic test with the aim of predicting hypertension risk is feasible and therefore its clinical value may not be sufficient to be adopted by clinical practitioners.

A total of 33 genetic variants with relevant pharmacogenetic associations were selected with the aim of evaluating if they could guide hypertension treatment by considering the underlying interactions between the genotype and the response to antihypertensive drugs.

Different criteria were used for the selection of genetic variants with interesting pharmacogenetic associations. As a result, the selected genetic variants have pharmacogenetic associations with different levels of evidence, from which only a few present pharmacogenetic associations with a good level of evidence *i.e.* a level of evidence smaller than or equal to 2.

This is the case of the association between the *CYP2D6* polymorphisms and the β -blocker metoprolol. It has been shown in multiple studies that these polymorphism may lead to a decreased metabolism or clearance of metoprolol [158–160]. Despite the reported differences in metoprolol pharmacokinetics in individuals carrying these polymorphisms, this may not translate into differences in adverse event rate. In fact, it was shown that there is no evidence for association between the variable pharmacokinetics of metoprolol, a result of the differences in *CYP2D6* metabolic activity, and a higher rate of adverse events or a lower efficacy in patients with hypertension [159].

In opposition, it has been shown that patients carrying an insertion of an Alu element, a short stretch of repetitive DNA, in the *ACE* polymorphism rs1799752 may have a decrease response to captopril treatment. This result have been observed in patients with heart failure, chronic obstructive pulmonary disease or diabetes mellitus type 2 [161–164].

Regarding the rs4961 *ADD1* polymorphism, it was verified in patients with liver cirrhosis that individuals carrying allele T are more likely to have a bad response (OR = 2.89, 95% CI: 1.28-2.81) to diuretic treatment using furosemide and spironolactone than those carrying allele G [165]. Despite these interesting results, this association was only verified in patients with liver cirrhosis in which diuretics were used for the treatment of ascites and not as an antihypertensive therapy. In other studies, the association between this polymorphism and diuretics has been found in hypertensive patients however with

contradictory results. Turner et al. have shown in a cohort of 259 hypertensive adults that allele T of this polymorphism is not related to hydrochlorothiazide response [166]. In opposition, in a study with similar sample size it was observed a higher reduction in SBP and DBP in hypertensive patients carrying allele T of the rs4961 genetic variant [167].

Another polymorphism with a pharmacogenetic association whose level of evidence is at least 2 is the rs4149601 from the gene *NEDD4L*. In white patients carrying the AA genotype it was observed a poorer response to hydrochlorothiazide therapy [168]. However, in individuals with Asian ancestry, the opposite result was observed, at a lower level of evidence, in which allele A was associated with increased blood pressure reduction [169]. A decreased response to hydrochlorothiazide it was also reported for hypertensive patients carrying allele T of the rs7297610 polymorphism (*YEATS4*) and patients carrying allele G of the rs16960228 polymorphisms (*PRKCA*). The association between this diuretic and each of these polymorphisms, rs7297610 and rs16960228, was observed in a European cohort and in a cohort with black individuals, respectively [169–173].

Concerning the genetic variants with FDA labelled pharmacogenetic associations, *NAT2* polymorphisms have been correlated with adverse effects after hydralazine therapy in patients with treatment-resistant hypertension [174] and the rs1057910 and *CYP2D6* polymorphisms (*CYP2D6*1*, *CYP2D6*10*, *CYP2D6*4*, *CYP2D6*5*) have been associated with decreased clearance of losartan and carvedilol, respectively [175–177]. Despite having an FDA label, these associations were observed in a small number of studies with a small study size. Given this, a level of evidence 3 was attributed to these associations by the PharmGKB which indicates that the reported pharmacogenetic associations may need to be replicated in different studies. The same applies to the remaining genetic variants selected in this work with pharmacogenetic associations whose level of evidence is smaller than 2.

The pharmacogenetic associations involving these last genetic variants may indicate that these polymorphisms can have a relevant role in the pharmacokinetics and pharmacodynamics of antihypertensive drugs. Additionally, it could also be hypothesized that the selected genetic variants may be involved in some underlying genetic factors responsible for the development of treatment-resistant hypertension.

Given this, a final analysis was performed in which using the PHARMA cohort it was studied if genetic variants with known pharmacogenetic associations could predict the risk of treatment-resistant hypertension. Using 24 pharmacogenetic variants from the HeartGenetics genetic panel it was possible to discriminate with statistical significance individuals with treatment-resistant hypertension (p-value = 0.003, additive model). A worse result, despite being statistical significant, was achieved using a subset of those initial genetic variants with relevant pharmacogenetic associations identified in section 4.6 (p-value = 0.035, additive model). These results support the idea that there are underlying genetic factors involved in the development of treatment-resistant hypertension [178]. Of the 24 genetic variants analysed, 8 showed a trend for association with the treatment-resistant hypertension trait. However, none reached statistical significance. Of these polymorphisms the SNPs rs5051 and rs699 have been previously associated with resistant hypertension. In the GenHAT study, a study with 2203 treatment resistant cases and 2354 controls, it was found that carriers of the T allele of rs699 and carriers of the G allele of rs5051 were associated with resistant hypertension with OR = 1.27(95% CI: 1.12-1.44, p-value

= 0.0001) and OR = 1.36 (95% CI: 1.20-1.53, p-value < 0.0001), respectively [179].

Given this, it can be hypothesized that these results may indicate that an individual carrying risk alleles of these genetic variants may have a worse response to treatment or could have a higher risk for treatment-resistant hypertension. Therefore, this information can assist the clinician in the definition of the most adequate pharmacological approach for hypertension treatment for each individual.

Different limitations were present in the analyses performed in this thesis. Firstly, the lack of positive results from the association tests performed using the data from the HTN and PHARMA cohorts and the logistic regression done using the data from the HTN cohort may be a consequence of the small sample size of the cohorts. Hong et al. mentioned that for testing a single genetic marker it would be required at least 248 samples, under the assumption of a 5% MAF, equal ratio of cases and controls, and a 5% error rate during genotyping, to achieve an adequate statistical power [180]. By comparison with the sample size used in this thesis, it is clear that the small number of samples may have contributed for the differences between the results obtained and current evidence. These same analyses may have also been limited by the fact that it were only considered genetic variants from the HeartGenetics genetic panel. Additionally, in the logistic regression it were considered as variables all genetic markers. An equivalent approach could have been taken if only genetic variants with statistical associations with the hypertension trait were considered. By doing this, it could be evaluated if a genetic risk model constructed using only genetic variants correlated with the given trait would have better results than a genetic risk model constructed using all of the genetic variants under study.

Concerning the selection of relevant pharmacogenetic variants through the information on the PharmGKB repository, it can be said that the criteria used for the selection of genetic variants may have been too strict. After selecting genetic variants with level of evidence smaller than or equal to 2 or variants with a FDA label, polymorphisms with at least one pharmacogenetic association with a p-value ≤ 0.001 were selected as relevant. By adopting this threshold for the p-value, genetic variants with important pharmacogenetic associations and significant p-values reported in several studies may have been not considered. However, this limitation was mitigated by the fact that for all variant/drug association pairs analysed the level of evidence was taken into consideration. Given this, if these genetic variants were not selected under the first inclusion criteria (level of evidence ≤ 2) it can be hypothesized that the pharmacogenetic associations involving these polymorphisms still lack a strong supporting evidence. Regarding the PHARMA cohort used to discriminate treatment-resistant hypertension patients, the small percentage of cases used was not ideal. It can be hypothesized that a more balanced cohort could have achieved higher discriminative power however further studies are needed to prove this hypothesis.

The results from this thesis are preliminary and would benefit, considering an academic point of view, from further studies involving a larger cohort of patients with hypertension and treatment-resistant hypertension so that it could be performed new association tests with a stronger statistical power. In addition, it would be also interesting to perform a study using a large cohort of individuals with treatment-resistant hypertension under antihypertensive therapy to understand the underlying genetic factors involved in this condition and in the response to antihypertensive treatment.

Overall, the results obtained point that a genetic test has a modest capacity for predicting hyperten-

sion using only genetic data, mostly because hypertension is a complex condition involving a variety of non-genetic and genetic factors. Given the results obtained in the pharmacogenetic analyses performed, it may be interesting to develop a test with the aim of offering a pharmacogenetic-guided treatment. Therefore, despite the modest capacity for predicting the genetic risk for hypertension, the genetic panel used demonstrates the potential for identifying patients with treatment-resistant hypertension.

Bibliography

- [1] H. Wang, M. Naghavi, C. Allen, R. M. Barber, Z. A. Bhutta, A. Carter, D. C. Casey, F. J. Charlson, A. Z. Chen, M. M. Coates, et al. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the global burden of disease study 2015. *The lancet*, 388(10053):1459–1544, 2016.
- [2] M. H. Forouzanfar, P. Liu, G. A. Roth, M. Ng, S. Biryukov, L. Marczak, L. Alexander, K. Estep, K. H. Abate, T. F. Akinyemiju, et al. Global burden of hypertension and systolic blood pressure of at least 110 to 115 mm hg, 1990-2015. *Jama*, 317(2):165–182, 2017.
- [3] D. Klarin, Q. M. Zhu, C. A. Emdin, M. Chaffin, S. Horner, B. J. McMillan, A. Leed, M. E. Weale, C. C. Spencer, F. Aguet, et al. Genetic analysis in uk biobank links insulin resistance and transendothelial migration pathways to coronary artery disease. *Nature genetics*, 49(9):1392, 2017.
- [4] H. R. Warren, E. Evangelou, C. P. Cabrera, H. Gao, M. Ren, B. Mifsud, I. Ntalla, P. Surendran, C. Liu, J. P. Cook, et al. Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. *Nature genetics*, 49(3):403, 2017.
- [5] P. Surendran, F. Drenos, R. Young, H. Warren, J. P. Cook, A. K. Manning, N. Grarup, X. Sim, D. R. Barnes, K. Witkowska, et al. Trans-ancestry meta-analyses identify rare and common variants associated with blood pressure and hypertension. *Nature genetics*, 48(10):1151, 2016.
- [6] A. T. Kraja et al. New blood pressure–associated loci identified in meta-analyses of 475 000 individualsclinical perspective. *Circulation: Genomic and Precision Medicine*, 10(5), 2017. ISSN 1942-325X. doi: 10.1161/CIRCGENETICS.117.001778. URL <http://circgenetics.ahajournals.org/content/10/5/e001778>.
- [7] G. Bakris, P. Sarafidis, R. Agarwal, and L. Ruilope. Review of blood pressure control rates and outcomes. *Journal of the American Society of Hypertension*, 8(2):127–141, 2014.
- [8] B. Alberts, A. Johnson, J. Lewis, P. Walter, M. Raff, and K. Roberts. *Molecular Biology of the Cell 4th Edition: International Student Edition*. ISBN 9780815332886.
- [9] D. T. Suzuki, A. J. Griffiths, et al. *An introduction to genetic analysis*. 1976.
- [10] B.-S. Jo and S. S. Choi. Introns: the functional benefits of introns in genomes. *Genomics & informatics*, 13(4):112–118, 2015.

- [11] G. Cooper. *The cell: A molecular approach*, 2nd edn. the cell: A molecular approach. sunderland, ma, 2000.
- [12] R. W. Carthew and E. J. Sontheimer. Origins and mechanisms of mirnas and sirnas. *Cell*, 136(4): 642–655, 2009.
- [13] P. R. Burton, M. D. Tobin, and J. L. Hopper. Key concepts in genetic epidemiology. *The Lancet*, 366(9489):941–951, 2005.
- [14] K. A. Frazer, S. S. Murray, N. J. Schork, and E. J. Topol. Human genetic variation and its contribution to complex traits. *Nature Reviews Genetics*, 10(4):241–251, 2009.
- [15] J. Alghamdi and S. Padmanabhan. *Fundamentals of complex trait genetics and association studies*. 2014.
- [16] W. Kalow. Pharmacogenetics and pharmacogenomics: origin, status, and the hope for personalized medicine. *The pharmacogenomics journal*, 6(3):162, 2006.
- [17] M. P. Doogue and T. M. Polasek. *The abcd of clinical pharmacokinetics*, 2013.
- [18] K. J. Karczewski, R. Daneshjou, and R. B. Altman. Pharmacogenomics. *PLoS computational biology*, 8(12):e1002817, 2012.
- [19] B. J. Pleuvry. Receptors, agonists and antagonists. *Anaesthesia & Intensive Care Medicine*, 5 (10):350–352, 2004.
- [20] S.-C. Chow. Bioavailability and bioequivalence in drug development. *Wiley Interdisciplinary Reviews: Computational Statistics*, 6(4):304–312, 2014.
- [21] C. Dietrich, A. Geier, and R. O. Elferink. Abc of oral bioavailability: transporters as gatekeepers in the gut. *Gut*, 52(12):1788–1795, 2003.
- [22] M. Ingelman-Sundberg, S. C. Sim, A. Gomez, and C. Rodriguez-Antona. Influence of cytochrome p450 polymorphisms on drug therapies: pharmacogenetic, pharmacoepigenetic and clinical aspects. *Pharmacology & therapeutics*, 116(3):496–526, 2007.
- [23] I. W. P. Consortium. Estimation of the warfarin dose with clinical and pharmacogenetic data. *New England Journal of Medicine*, 360(8):753–764, 2009.
- [24] A. E. Rettie, R. L. Haining, M. Bajpai, and R. H. Levy. A common genetic basis for idiosyncratic toxicity of warfarin and phenytoin. *Epilepsy research*, 35(3):253–255, 1999.
- [25] A. Zimmermann and J. T. Matschiner. Biochemical basis of hereditary resistance to warfarin in the rat. *Biochemical pharmacology*, 23(6):1033–1040, 1974.
- [26] R. P. Owen, R. B. Altman, and T. E. Klein. Pharmgkb and the international warfarin pharmacogenetics consortium: the changing role for pharmacogenomic databases and single-drug pharmacogenetics. *Human mutation*, 29(4):456–460, 2008.

- [27] N. Nguyen, P. Anley, Y. Y. Margaret, G. Zhang, A. A. Thompson, and L. J. Jennings. Genetic and clinical determinants influencing warfarin dosing in children with heart disease. *Pediatric cardiology*, 34(4):984–990, 2013.
- [28] P. Sunnucks. Efficient genetic markers for population biology. *Trends in ecology & evolution*, 15(5):199–203, 2000.
- [29] S. Raza, M. W. Shoaib, and H. Mubeen. Genetic markers: Importance, uses and applications.
- [30] J. S. Bader. The relative power of snps and haplotype as genetic markers for association tests. *Pharmacogenomics*, 2(1):11–24, 2001.
- [31] L. Monserrat. Genetic and genomic technologies: Next generation sequencing for inherited cardiovascular conditions. In *Cardiovascular Genetics and Genomics*, pages 97–117. Springer, 2018.
- [32] Q. Wang, Q. Lu, and H. Zhao. A review of study designs and statistical methods for genomic epidemiology studies using next generation sequencing. *Frontiers in genetics*, 6:149, 2015.
- [33] A. S. Foulkes. *Applied statistical genetics with R: for population-based association studies*. Springer Science & Business Media, 2009.
- [34] R. Patnala, J. Clements, and J. Batra. Candidate gene association studies: a comprehensive guide to useful in silico tools. *BMC genetics*, 14(1):14–39, 2013.
- [35] S. L. Spain and J. C. Barrett. Strategies for fine-mapping complex traits. *Human molecular genetics*, 24(R1):R1111–R1119, 2015.
- [36] J. Lin and K. Musunuru. From genotype to phenotype. *Circulation: Genomic and Precision Medicine*, 11(2), 2018. ISSN 1942-325X. doi: 10.1161/CIRCGEN.117.001946.
- [37] W. S. Bush and J. H. Moore. Chapter 11: Genome-wide association studies. *PLOS Computational Biology*, 8(12):1–11, 12 2012. doi: 10.1371/journal.pcbi.1002822.
- [38] C. M. Lewis. Genetic association studies: design, analysis and interpretation. *Briefings in bioinformatics*, 3(2):146–153, 2002.
- [39] J. E. Wigginton, D. J. Cutler, and G. R. Abecasis. A note on exact tests of hardy-weinberg equilibrium. *The American Journal of Human Genetics*, 76(5):887–893, 2005.
- [40] C. M. Lewis and J. Knight. Introduction to genetic association studies. *Cold Spring Harbor Protocols*, 2012(3):pdb-top068163, 2012.
- [41] D. J. Schaid, A. J. Batzler, G. D. Jenkins, and M. A. Hildebrandt. Exact tests of hardy-weinberg equilibrium and homogeneity of disequilibrium across strata. *The American Journal of Human Genetics*, 79(6):1071–1080, 2006.
- [42] S. Turner, L. L. Armstrong, Y. Bradford, C. S. Carlson, D. C. Crawford, A. T. Crenshaw, M. de Andrade, K. F. Doheny, J. L. Haines, G. Hayes, et al. Quality control procedures for genome-wide association studies. *Current protocols in human genetics*, 68(1):1–19, 2011.

- [43] A. L. Price, N. J. Patterson, R. M. Plenge, M. E. Weinblatt, N. A. Shadick, and D. Reich. Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics*, 38(8):904, 2006.
- [44] L. Liu, D. Zhang, H. Liu, and C. Arendt. Robust methods for population stratification in genome wide association studies. *BMC bioinformatics*, 14(1):132, 2013.
- [45] H.-Y. Kim. Statistical notes for clinical researchers: chi-squared test and fisher’s exact test. *Restorative dentistry & endodontics*, 42(2):152–155, 2017.
- [46] P. Sprent. Fisher exact test. In *International encyclopedia of statistical science*, pages 524–525. Springer, 2011.
- [47] E. W. Weisstein. Bonferroni correction. 2004.
- [48] G. Mancia, R. Fagard, K. Narkiewicz, J. Redón, A. Zanchetti, M. Böhm, T. Christiaens, R. Cifkova, G. De Backer, A. Dominiczak, et al. Task force members (2013) 2013 esh/esc guidelines for the management of arterial hypertension: the task force for the management of arterial hypertension of the european society of hypertension (esh) and of the european society of cardiology (esc). *J Hypertens*, 31(7):1281–1357, 2013.
- [49] S.-Y. Ahn and C. Gupta. Genetic programming of hypertension. *Frontiers in pediatrics*, 5:285, 2017.
- [50] J. E. Hall, J. P. Granger, J. M. do Carmo, A. A. da Silva, J. Dubinion, E. George, S. Hamza, J. Speed, and M. E. Hall. Hypertension: physiology and pathophysiology. *Comprehensive Physiology*, 2012.
- [51] G. Mancia, G. Grassi, G. Pomidossi, L. Gregorini, G. Bertinieri, G. Parati, A. Ferrari, and A. Zanchetti. Effects of blood-pressure measurement by the doctor on patient’s blood pressure and heart rate. *The Lancet*, 322(8352):695–698, 1983.
- [52] G. Bobrie, P. Clerson, J. Menard, N. Postel-Vinay, G. Chatellier, and P.-F. Plouin. Masked hypertension: a systematic review. *Journal of hypertension*, 26(9):1715–1725, 2008.
- [53] R. Klabunde. *Cardiovascular physiology concepts*, pages 93–123. Lippincott Williams & Wilkins, 2011.
- [54] J. E. Hall, A. C. Guyton, and M. W. Brands. Pressure-volume regulation in hypertension. *Kidney International*, 49, 1996.
- [55] J. D. H. M. Hall J.E., Granger J.P. *Hurst’s The Heart*, chapter Chapter 69: Pathophysiology of hypertension. McGraw-Hill, Inc., 2011.
- [56] S. Oparil, M. C. Acelajado, G. L. Bakris, D. R. Berlowitz, R. Cifková, A. F. Dominiczak, G. Grassi, J. Jordan, N. R. Poulter, A. Rodgers, and P. K. Whelton. Hypertension. *Nature Reviews Disease Primers*, 4, Mar 2018. Primer.

- [57] W. Feng, L. J. Dell'Italia, and P. W. Sanders. Novel paradigms of salt and hypertension. *Journal of the American Society of Nephrology*, 28(5):1362–1369, 2017.
- [58] F. Schweda, U. Friis, C. Wagner, O. Skott, and A. Kurtz. Renin release. *Physiology*, 22(5):310–319, 2007.
- [59] S. Melmed. *Williams textbook of endocrinology*, chapter Chapter 16: Endocrine Hypertension. Elsevier Health Sciences, 2016.
- [60] G. L. Bakris and M. Sorrentino. *Hypertension: A Companion to Braunwald's Heart Disease*, pages 33–51. Elsevier, 2018.
- [61] J. Varagic, S. Ahmad, S. Nagata, and C. M. Ferrario. Ace2: angiotensin ii/angiotensin-(1–7) balance in cardiac and renal injury. *Current hypertension reports*, 16(3):420, 2014.
- [62] A. McCurley and I. Z. Jaffe. Mineralocorticoid receptors in vascular function and disease. *Molecular and cellular endocrinology*, 350(2):256–265, 2012.
- [63] G. E. Woodard and J. A. Rosado. Natriuretic peptides in vascular physiology and pathology. *International review of cell and molecular biology*, 268:59–93, 2008.
- [64] F.-R. E. Curry. Atrial natriuretic peptide: an essential physiological regulator of transvascular fluid, protein transport, and plasma volume. *The Journal of clinical investigation*, 115(6):1458–1461, 2005.
- [65] Z. Armaly, S. Assady, and Z. Abassi. Corin: a new player in the regulation of salt–water balance and blood pressure. *Current opinion in nephrology and hypertension*, 22(6):713–722, 2013.
- [66] N. Schlueter, A. de Sterke, D. M. Willmes, J. Spranger, J. Jordan, and A. L. Birkenfeld. Metabolic actions of natriuretic peptides and therapeutic potential in the metabolic syndrome. *Pharmacology & therapeutics*, 144(1):12–27, 2014.
- [67] R. Khaddaj Mallat, C. Mathew John, D. J. Kendrick, and A. P. Braun. The vascular endothelium: A regulator of arterial tone and interface for the immune system. *Critical reviews in clinical laboratory sciences*, 54(7-8):458–470, 2017.
- [68] T. Ayub, S. N. Khan, S. G. Ayub, R. Dar, and K. I. Andrabi. Reduced nitrate level in individuals with hypertension and diabetes. *Journal of cardiovascular disease research*, 2(3):172–176, 2011.
- [69] J. A. Panza, P. R. Casino, D. M. Badar, and A. A. Quyyumi. Effect of increased availability of endothelium-derived nitric oxide precursor on endothelium-dependent vascular relaxation in normal subjects and in patients with essential hypertension. *Circulation*, 87(5):1475–1481, 1993.
- [70] I. Lazich and G. L. Bakris. Endothelin antagonism in patients with resistant hypertension and hypertension nephropathy. In *Endothelin in Renal Physiology and Disease*, volume 172, pages 223–234. Karger Publishers, 2011.

- [71] K. Dharmashankar and M. E. Widlansky. Vascular endothelial function and hypertension: insights and directions. *Current hypertension reports*, 12(6):448–455, 2010.
- [72] C. Heymans and A. Delaunois. Fundamental role of the tone and resistance to stretch of the carotid sinus arteries in the reflex regulation of blood pressure. *Science*, 114(2969):546–547, 1951.
- [73] G. Mancia and G. Grassi. The autonomic nervous system and hypertension. *Circulation research*, 114(11):1804–1814, 2014.
- [74] P. A. Smith, L. N. Graham, A. F. Mackintosh, J. B. Stoker, and D. A. Mary. Relationship between central sympathetic activity and stages of human hypertension. *American journal of hypertension*, 17(3):217–222, 2004.
- [75] G. F. DiBona. Physiology in perspective: The wisdom of the body. neural control of the kidney. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 289(3):R633–R641, 2005.
- [76] G. F. Dibona. Sympathetic nervous system and hypertension. *Hypertension*, 61(3):556–560, 2013.
- [77] T. Fujita. Mechanism of salt-sensitive hypertension: focus on adrenal and sympathetic nervous systems. *Journal of the American Society of Nephrology*, 25(6):1148–1155, 2014.
- [78] S. Mu, T. Shimosawa, S. Ogura, H. Wang, Y. Uetake, F. Kawakami-Mori, T. Marumo, Y. Yatomi, D. S. Geller, H. Tanaka, et al. Epigenetic modulation of the renal β -adrenergic–wnk4 pathway in salt-sensitive hypertension. *Nature medicine*, 17(5):573, 2011.
- [79] G. L. Bakris and M. Sorrentino. *Hypertension: A Companion to Braunwald’s Heart Disease*, pages 60–69. Elsevier, 2018.
- [80] P. B. Munroe, M. R. Barnes, and M. J. Caulfield. Advances in blood pressure genomics. *Circulation Research*, 112(10):1365–1379, 2013.
- [81] J. L. ANNEST, C. F. SING, P. BIRON, and J.-G. MONGEAU. Familial aggregation of blood pressure and weight in adoptive families: II. estimation of the relative contributions of genetic and common environmental factors to blood pressure correlations between family members. *American Journal of Epidemiology*, 110(4):492–503, 1979.
- [82] R. J. Havlik, R. J. Garrison, M. Feinleib, W. B. Kannel, W. P. Castelli, and P. M. Mcnamara. Blood pressure aggregation in families. *American Journal of Epidemiology*, 110(3):304–312, 1979.
- [83] P. Stocks. A biometric investigation of twins and their brothers and sisters. *Annals of Human Genetics*, 4(1-2):49–108, 1930.

- [84] E. Evangelou et al. Genetic analysis of over one million people identifies 535 novel loci for blood pressure. *bioRxiv*, 2017. doi: 10.1101/198234. URL <https://www.biorxiv.org/content/early/2017/10/11/198234>.
- [85] Diagnoses - main icd10: I10 essential (primary) hypertension. URL http://big.stats.ox.ac.uk/pheno/biobank_I10.
- [86] D. Mozaffarian, E. J. Benjamin, A. S. Go, D. K. Arnett, M. J. Blaha, M. Cushman, S. R. Das, S. de Ferranti, J.-P. Després, H. J. Fullerton, et al. Heart disease and stroke statistics—2016 update: a report from the american heart association. *Circulation*, 133(4):e38–e360, 2016.
- [87] S. S. Franklin, W. Gustin, N. D. Wong, M. G. Larson, M. A. Weber, W. B. Kannel, and D. Levy. Hemodynamic patterns of age-related changes in blood pressure: the framingham heart study. *Circulation*, 96(1):308–315, 1997.
- [88] Y. Wang and Q. J. Wang. The prevalence of prehypertension and hypertension among us adults according to the new joint national committee guidelines: new challenges of the old problem. *Archives of Internal Medicine*, 164(19):2126–2134, 2004.
- [89] P. W. Wilson, R. B. D’agostino, L. Sullivan, H. Parise, and W. B. Kannel. Overweight and obesity as determinants of cardiovascular risk: the framingham experience. *Archives of internal medicine*, 162(16):1867–1872, 2002.
- [90] D. M. Lloyd-Jones, K. Liu, L. A. Colangelo, L. L. Yan, L. Klein, C. M. Loria, C. E. Lewis, and P. Savage. Consistently stable or decreased body mass index in young adulthood and longitudinal changes in metabolic syndrome components: the coronary artery risk development in young adults study. *Circulation*, 115(8):1004–1011, 2007.
- [91] A. P. Carson, G. Howard, G. L. Burke, S. Shea, E. B. Levitan, and P. Muntner. Ethnic differences in hypertension incidence among middle-aged and older adults: the multi-ethnic study of atherosclerosis. *Hypertension*, 57(6):1101–1107, 2011.
- [92] J. Stamler, P. Elliott, L. Appel, Q. Chan, M. Buzzard, B. Dennis, A. R. Dyer, P. Elmer, P. Greenland, D. Jones, et al. Higher blood pressure in middle-aged american adults with less education—role of multiple dietary factors: the intermap study. *Journal of human hypertension*, 17(9):655, 2003.
- [93] J. Stamler, G. Rose, P. Elliott, A. Dyer, M. Marmot, H. Kesteloot, and R. Stamler. Findings of the international cooperative intersalt study. *Hypertension*, 17(1 Suppl):I9, 1991.
- [94] M. Hillbom, P. Saloheimo, and S. Juvela. Alcohol consumption, blood pressure, and the risk of stroke. *Current hypertension reports*, 13(3):208–213, 2011.
- [95] E. Rapsomaniki, A. Timmis, J. George, M. Pujades-Rodriguez, A. D. Shah, S. Denaxas, I. R. White, M. J. Caulfield, J. E. Deanfield, L. Smeeth, 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. *The Lancet*, 383(9932):1899–1911, 2014.

- [96] J. Stamler, R. Stamler, and J. D. Neaton. Blood pressure, systolic and diastolic, and cardiovascular risks: Us population data. *Archives of internal medicine*, 153(5):598–615, 1993.
- [97] G. L. Bakris and M. Sorrentino. *Hypertension: A Companion to Braunwald's Heart Disease*, pages 1–14. Elsevier, 2018.
- [98] P. S. Collaboration et al. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *The Lancet*, 360(9349):1903–1913, 2002.
- [99] A. J. Collins, R. N. Foley, D. T. Gilbertson, and S.-C. Chen. United states renal data system public health surveillance of chronic kidney disease and end-stage renal disease. *Kidney international supplements*, 5(1):2–7, 2015.
- [100] P. K. Whelton, L. Appel, J. Charleston, A. T. Dalcin, C. Ewart, L. Fried, D. Kaidy, M. J. Klag, S. Kumanyika, L. Steffen, et al. The effects of nonpharmacologic interventions on blood pressure of persons with high normal levels: results of the trials of hypertension prevention, phase i. *Jama*, 267(9):1213–1220, 1992.
- [101] T. of Hypertension Prevention Collaborative Research Group et al. Effects of weight loss and sodium reduction intervention on blood pressure and hypertension incidence in over-weight people with high normal blood pressure: the trials of hypertension prevention, phase ii. *Arch. Intern. Med.*, 157:657–667, 1997.
- [102] F. M. Sacks, L. P. Svetkey, W. M. Vollmer, L. J. Appel, G. A. Bray, D. Harsha, E. Obarzanek, P. R. Conlin, E. R. Miller, D. G. Simons-Morton, et al. Effects on blood pressure of reduced dietary sodium and the dietary approaches to stop hypertension (dash) diet. *New England journal of medicine*, 344(1):3–10, 2001.
- [103] H. O. Dickinson, J. M. Mason, D. J. Nicolson, F. Campbell, F. R. Beyer, J. V. Cook, B. Williams, and G. A. Ford. Lifestyle interventions to reduce raised blood pressure: a systematic review of randomized controlled trials. *Journal of hypertension*, 24(2):215–233, 2006.
- [104] F. J. He and G. A. MacGregor. How far should salt intake be reduced? *Hypertension*, 42(6):1093–1099, 2003.
- [105] N. A. Graudal, T. Hubeck-Graudal, and G. Jürgens. Effects of low-sodium diet vs. high-sodium diet on blood pressure, renin, aldosterone, catecholamines, cholesterol, and triglyceride (cochrane review). *American journal of hypertension*, 25(1):1–15, 2012.
- [106] A. L. Klatsky and E. Gunderson. Alcohol and hypertension: a review. *Journal of the American Society of Hypertension*, 2(5):307–317, 2008.
- [107] I. B. Puddey, L. J. Beilin, and R. Vandongen. Regular alcohol use raises blood pressure in treated hypertensive subjects: a randomised controlled trial. *The Lancet*, 329(8534):647–651, 1987.

- [108] M. Roerecke, J. Kaczorowski, S. W. Tobe, G. Gmel, O. S. Hasan, and J. Rehm. The effect of a reduction in alcohol consumption on blood pressure: a systematic review and meta-analysis. *The Lancet Public Health*, 2(2):e108–e120, 2017.
- [109] L. A. Bazzano, T. Hu, K. Reynolds, L. Yao, C. Bunol, Y. Liu, C.-S. Chen, M. J. Klag, P. K. Whelton, and J. He. Effects of low-carbohydrate and low-fat diets: a randomized trial. *Annals of internal medicine*, 161(5):309–318, 2014.
- [110] J. He, M. R. Wofford, K. Reynolds, J. Chen, C.-S. Chen, L. Myers, D. L. Minor, P. J. Elmer, D. W. Jones, and P. K. Whelton. Effect of dietary protein supplementation on blood pressure: a randomized, controlled trial. *Circulation*, pages CIRCULATIONAHA-110, 2011.
- [111] Y. Yokoyama, K. Nishimura, N. D. Barnard, M. Takegami, M. Watanabe, A. Sekikawa, T. Okamura, and Y. Miyamoto. Vegetarian diets and blood pressure: a meta-analysis. *JAMA internal medicine*, 174(4):577–587, 2014.
- [112] A. J. Nordmann, K. Suter-Zimmermann, H. C. Bucher, I. Shai, K. R. Tuttle, R. Estruch, and M. Briel. Meta-analysis comparing mediterranean to low-fat diets for modification of cardiovascular risk factors. *The American journal of medicine*, 124(9):841–851, 2011.
- [113] F. Sofi, R. Abbate, G. F. Gensini, and A. Casini. Accruing evidence on benefits of adherence to the mediterranean diet on health: an updated systematic review and meta-analysis. *The American journal of clinical nutrition*, 92(5):1189–1196, 2010.
- [114] R. Estruch, E. Ros, J. Salas-Salvadó, M.-I. Covas, D. Corella, F. Arós, E. Gómez-Gracia, V. Ruiz-Gutiérrez, M. Fiol, J. Lapetra, et al. Primary prevention of cardiovascular disease with a mediterranean diet supplemented with extra-virgin olive oil or nuts. *New England Journal of Medicine*, 2018.
- [115] J. A. Blumenthal, M. A. Babyak, A. Hinderliter, L. L. Watkins, L. Craighead, P.-H. Lin, C. Caccia, J. Johnson, R. Waugh, and A. Sherwood. Effects of the dash diet alone and in combination with exercise and weight loss on blood pressure and cardiovascular biomarkers in men and women with high blood pressure: the encore study. *Archives of internal medicine*, 170(2):126–135, 2010.
- [116] R. Romero, J. Bonet, A. de la Sierra, M. T. Aguilera, E. S. Investigators, et al. Undiagnosed obesity in hypertension: clinical and therapeutic implications. *Blood pressure*, 16(6):347–353, 2007.
- [117] J. E. Neter, B. E. Stam, F. J. Kok, D. E. Grobbee, and J. M. Geleijnse. Influence of weight reduction on blood pressure: a meta-analysis of randomized controlled trials. *Hypertension*, 42(5):878–884, 2003.
- [118] J. Jordan, V. Yumuk, M. Schlaich, P. M. Nilsson, B. Zahorska-Markiewicz, G. Grassi, R. E. Schmieder, S. Engeli, and N. Finer. Joint statement of the european association for the study of obesity and the european society of hypertension: obesity and difficult to treat arterial hypertension. *Journal of hypertension*, 30(6):1047–1055, 2012.

- [119] V. A. Cornelissen and R. H. Fagard. Effects of endurance training on blood pressure, blood pressure-regulating mechanisms, and cardiovascular risk factors. *Hypertension*, 46(4):667–675, 2005.
- [120] V. A. Cornelissen, R. H. Fagard, E. Coeckelberghs, and L. Vanhees. Impact of resistance training on blood pressure and other cardiovascular risk factors: a meta-analysis of randomized, controlled trials. *Hypertension*, pages HYPERTENSIONAHA–111, 2011.
- [121] M. Pucci, N. Sarween, E. Knox, G. Lipkin, and U. Martin. Angiotensin-converting enzyme inhibitors and angiotensin receptor blockers in women of childbearing age: risks versus benefits. *Expert review of clinical pharmacology*, 8(2):221–231, 2015.
- [122] P. R. James and C. Nelson-Piercy. Management of hypertension before, during, and after pregnancy. *Heart*, 90(12):1499–1504, 2004.
- [123] K. Jamerson and V. DeQuattro. The impact of ethnicity on response to antihypertensive therapy. *The American journal of medicine*, 101(3):22S–32S, 1996.
- [124] B. P. L. T. Trialists' Collaboration. Effects of different regimens to lower blood pressure on major cardiovascular events in older and younger adults: meta-analysis of randomised trials. *Bmj*, 336(7653):1121–1123, 2008.
- [125] F. Turnbull, M. Woodward, B. Neal, F. Barzi, T. Ninomiya, J. Chalmers, V. Perkovic, N. Li, S. MacMahon, and B. P. L. T. T. Collaboration. Do men and women respond differently to blood pressure-lowering treatment? results of prospectively designed overviews of randomized trials. *European heart journal*, 29(21):2669–2680, 2008.
- [126] M. E. Ernst and M. Moser. Use of diuretics in patients with hypertension. *New England Journal of Medicine*, 361(22):2153–2164, 2009.
- [127] G. C. Roush, M. E. Ernst, J. B. Kostis, S. Tandon, and D. A. Sica. Head-to-head comparisons of hydrochlorothiazide with indapamide and chlorthalidone: antihypertensive and metabolic effects. *Hypertension*, pages HYPERTENSIONAHA–114, 2015.
- [128] B. Pitt. Cardiovascular therapeutics: A companion to Braunwald's heart disease. *Circulation*, 96(3), 1997.
- [129] C. S. Wiysonge, H. A. Bradley, J. Volmink, B. M. Mayosi, and L. H. Opie. Beta-blockers for hypertension. *The Cochrane Library*, 2017.
- [130] A. M. Sharma, T. Pischon, S. Hardt, I. Kunz, and F. C. Luft. Hypothesis: β -adrenergic receptor blockers and weight gain: a systematic analysis. *Hypertension*, 37(2):250–254, 2001.
- [131] G. L. Bakris, V. Fonseca, R. E. Katholi, J. B. McGill, F. H. Messerli, R. A. Phillips, P. Raskin, J. T. Wright, R. Oakes, M. A. Lukas, et al. Metabolic effects of carvedilol vs metoprolol in patients with type 2 diabetes mellitus and hypertension: a randomized controlled trial. *Jama*, 292(18):2227–2236, 2004.

- [132] B. P. L. T. T. Collaboration et al. Effects of different blood-pressure-lowering regimens on major cardiovascular events: results of prospectively-designed overviews of randomised trials. *The Lancet*, 362(9395):1527–1535, 2003.
- [133] M. Law, J. Morris, and N. Wald. 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*, 338:b1665, 2009.
- [134] P. A. Poole-Wilson, J. Lubsen, B.-A. Kirwan, F. J. Van Dalen, G. Wagener, N. Danchin, H. Just, K. A. Fox, S. J. Pocock, T. C. Clayton, et al. Effect of long-acting nifedipine on mortality and cardiovascular morbidity in patients with stable angina requiring treatment (action trial): randomised controlled trial. *The Lancet*, 364(9437):849–857, 2004.
- [135] B. Dahlöf, P. S. Sever, N. R. Poulter, H. Wedel, D. G. Beevers, M. Caulfield, R. Collins, S. E. Kjeldsen, A. Kristinsson, G. T. McInnes, et al. Prevention of cardiovascular events with an antihypertensive regimen of amlodipine adding perindopril as required versus atenolol adding bendroflumethiazide as required, in the anglo-scandinavian cardiac outcomes trial-blood pressure lowering arm (ascot-bpla): a multicentre randomised controlled trial. *The Lancet*, 366(9489):895–906, 2005.
- [136] M. D. Guazzi, C. Cipolla, P. Della Bella, F. Fabbiocchi, P. Montorsi, and P. Sganzerla. Disparate unloading efficacy of the calcium channel blockers, verapamil and nifedipine, on the failing hypertensive left ventricle. *American heart journal*, 108(1):116–123, 1984.
- [137] O. Investigators. Telmisartan, ramipril, or both in patients at high risk for vascular events. *New England Journal of Medicine*, 358(15):1547–1559, 2008.
- [138] N. J. Brown, W. A. Ray, M. Snowden, and M. R. Griffin. Black americans have an increased rate of angiotensin converting enzyme inhibitor-associated angioedema. *Clinical Pharmacology & Therapeutics*, 60(1):8–13, 1996.
- [139] D. S. Wald, M. Law, J. K. Morris, J. P. Bestwick, and N. J. Wald. Combination therapy versus monotherapy in reducing blood pressure: meta-analysis on 11,000 participants from 42 trials. *The American journal of medicine*, 122(3):290–300, 2009.
- [140] G. Corrao, A. Parodi, A. Zambon, F. Heiman, A. Filippi, C. Cricelli, L. Merlino, and G. Mancia. Reduced discontinuation of antihypertensive treatment by two-drug combination as first step. evidence from daily life practice. *Journal of hypertension*, 28(7):1584–1590, 2010.
- [141] A. Myat, S. R. Redwood, A. C. Qureshi, J. A. Spertus, and B. Williams. Resistant hypertension. *Bmj*, 345(7884):e7473, 2012.
- [142] B. Williams, T. M. MacDonald, S. Morant, D. J. Webb, P. Sever, G. McInnes, I. Ford, J. K. Cruickshank, M. J. Caulfield, J. Salsbury, 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. *The Lancet*, 386(10008):2059–2068, 2015.

- [143] D. N. Juurlink, M. M. Mamdani, D. S. Lee, A. Kopp, P. C. Austin, A. Laupacis, and D. A. Re-delmeier. Rates of hyperkalemia after publication of the randomized aldactone evaluation study. *New England Journal of Medicine*, 351(6):543–551, 2004.
- [144] G. L. Bakris, M. K. Nadim, H. Haller, E. G. Lovett, J. E. Schafer, and J. D. Bisognano. Baroreflex activation therapy provides durable benefit in patients with resistant hypertension: results of long-term follow-up in the rheos pivotal trial. *Journal of the American Society of Hypertension*, 6(2): 152–158, 2012.
- [145] H. Krum, M. Schlaich, R. Whitbourn, P. A. Sobotka, J. Sadowski, K. Bartus, B. Kapelak, A. Walton, H. Sievert, S. Thambar, et al. Catheter-based renal sympathetic denervation for resistant hypertension: a multicentre safety and proof-of-principle cohort study. *The Lancet*, 373(9671): 1275–1281, 2009.
- [146] G. Grassi, B. M. Cattaneo, G. Seravalle, A. Lanfranchi, and G. Mancia. Baroreflex control of sympathetic nerve activity in essential and secondary hypertension. *Hypertension*, 31(1):68–72, 1998.
- [147] M. Whirl-Carrillo, E. M. McDonagh, J. Hebert, L. Gong, K. Sangkuhl, C. Thorn, R. B. Altman, and T. E. Klein. Pharmacogenomics knowledge for personalized medicine. *Clinical Pharmacology & Therapeutics*, 92(4):414–417, 2012.
- [148] P. Whelton, R. Carey, W. Aronow, D. Casey Jr, K. Collins, C. Dennison Himmelfarb, et al. 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*, 2017.
- [149] G. Salanti, L. Southam, D. Altshuler, K. Ardlie, I. Barroso, M. Boehnke, M. C. Cornelis, T. M. Frayling, H. Grallert, N. Grarup, et al. Underlying genetic models of inheritance in established type 2 diabetes associations. *American journal of epidemiology*, 170(5):537–545, 2009.
- [150] C. for Drug Evaluation and Research. Science & research (drugs) - table of pharmacogenomic biomarkers in drug labeling. URL <https://www.fda.gov/Drugs/ScienceResearch/ucm572698.htm>.
- [151] *Mann–Whitney U and Wilcoxon Signed-Rank*, chapter 7, pages 297–352. Wiley-Blackwell, 2017. ISBN 9781119454205. doi: 10.1002/9781119454205.ch7. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119454205.ch7>.
- [152] P. R. Gard. Implications of the angiotensin converting enzyme gene insertion/deletion polymorphism in health and disease: a snapshot review. *International journal of molecular epidemiology and genetics*, 1(2):145, 2010.
- [153] E. Salvi, Z. Kutalik, N. Glorioso, P. Benaglio, F. Frau, T. Kuznetsova, H. Arima, C. Hoggart, J. Tichet, Y. P. Nikitin, et al. Genome-wide association study using a high-density snp-array and

- case-control design identifies a novel essential hypertension susceptibility locus in the promoter region of *enos*. *Hypertension*, 59(2):248, 2012.
- [154] D. Conen, R. J. Glynn, J. E. Buring, P. M. Ridker, and R. Y. Zee. Natriuretic peptide precursor a gene polymorphisms and risk of blood pressure progression and incident hypertension. *Hypertension*, 50(6):1114–1119, 2007.
- [155] B. Sun, J. S. Williams, L. Pojoga, B. Chamarthi, J. Lasky-Su, B. A. Raby, P. N. Hopkins, X. Jeunemaitre, N. J. Brown, C. Ferri, et al. Renin gene polymorphism: its relationship to hypertension, renin levels and vascular responses. *Journal of the Renin-Angiotensin-Aldosterone System*, 12(4):564–571, 2011.
- [156] M. U. Moreno, G. San José, A. Fortuño, Ó. Beloqui, J. Díez, and G. Zalba. The c242t cyba polymorphism of nadph oxidase is associated with essential hypertension. *Journal of hypertension*, 24(7):1299–1306, 2006.
- [157] E. Held, J. Cape, and N. Tintle. Comparing machine learning and logistic regression methods for predicting hypertension using a combination of gene expression and next-generation sequencing data. In *BMC proceedings*, volume 10, page 34. BioMed Central, 2016.
- [158] J. Batty, A. Hall, H. White, J. Wikstrand, R. de Boer, D. van Veldhuisen, P. van der Harst, F. Waagstein, Å. Hjalmarsen, J. Kjekshus, et al. An investigation of *cyp2d6* genotype and response to metoprolol cr/xl during dose titration in patients with heart failure: A merit-hf substudy. *Clinical Pharmacology & Therapeutics*, 95(3):321–330, 2014.
- [159] I. Zineh, A. L. Beitelshees, A. Gaedigk, J. R. Walker, D. F. Pauly, K. Eberst, J. S. Leeder, M. S. Phillips, C. A. Gelfand, and J. A. Johnson. Pharmacokinetics and *cyp2d6* genotypes do not predict metoprolol adverse events or efficacy in hypertension. *Clinical Pharmacology & Therapeutics*, 76(6):536–544, 2004.
- [160] S. Jin, H. Chung, M. Chung, J.-I. Kim, J.-H. Kang, S. Woo, S. Bang, S. Lee, H. Lee, and J. Roh. Influence of *cyp2d6** 10 on the pharmacokinetics of metoprolol in healthy korean volunteers. *Journal of clinical pharmacy and therapeutics*, 33(5):567–573, 2008.
- [161] B. Volkan-Salanci, S. Dagdelen, M. Alikasifoglu, T. Erbas, M. Hayran, and B. Erbas. Impact of renin-angiotensin system polymorphisms on renal haemodynamic responsiveness to acute angiotensin-converting enzyme inhibition in type 2 diabetes mellitus. *Journal of the Renin-Angiotensin-Aldosterone System*, 10(1):41–50, 2009.
- [162] S. Mizuiri, H. Hemmi, A. Inoue, M. Takano, S. Kadomatsu, H. Tanimoto, M. Tanegashima, I. Hayashi, T. Fushimi, and A. Hasegawa. Renal hemodynamic changes induced by captopril and angiotensin-converting enzyme gene polymorphism. *Nephron*, 75(3):310–314, 1997.
- [163] P. Jacobsen, K. Rossing, P. Rossing, L. Tarnow, C. Mallet, O. Poirier, F. Cambien, and H.-H. Parving. Angiotensin converting enzyme gene polymorphism and ace inhibition in diabetic nephropathy. *Kidney international*, 53(4):1002–1006, 1998.

- [164] H. Kanazawa, K. Hirata, and J. Yoshikawa. Effects of captopril administration on pulmonary haemodynamics and tissue oxygenation during exercise in ace gene subtypes in patients with copd: a preliminary study. *Thorax*, 58(7):629–631, 2003.
- [165] Y.-Y. Yang, H.-C. Lin, M.-W. Lin, C.-J. Chu, F.-Y. Lee, M.-C. Hou, S.-D. Lee, W.-P. Lee, T.-T. Liu, and T.-S. Jap. Identification of diuretic non-responders with poor long-term clinical outcomes: a 1-year follow-up of 176 non-azotaemic cirrhotic patients with moderate ascites. *Clinical Science*, 121(11):509–521, 2011.
- [166] S. T. Turner, A. B. Chapman, G. L. Schwartz, and E. Boerwinkle. Effects of endothelial nitric oxide synthase, α -adducin, and other candidate gene polymorphisms on blood pressure response to hydrochlorothiazide. *American journal of hypertension*, 16(10):834–839, 2003.
- [167] P. Manunta, G. Lavery, C. Lanzani, P. S. Braund, M. Simonini, C. Bodycote, L. Zagato, S. Delli Carpini, C. Tantardini, E. Brioni, et al. Physiological interaction between α -adducin and wnk1-nedd4l pathways on sodium-related blood pressure regulation. *Hypertension*, 52(2):366–372, 2008.
- [168] S. T. Turner, E. Boerwinkle, J. R. O’connell, K. R. Bailey, Y. Gong, A. B. Chapman, C. W. McDonough, A. L. Beitelshes, G. L. Schwartz, J. G. Gums, et al. Genomic association analysis of common variants influencing antihypertensive response to hydrochlorothiazide. *Hypertension*, 62(2):391–397, 2013.
- [169] F. Luo, Y. Wang, X. Wang, K. Sun, X. Zhou, and R. Hui. A functional variant of nedd4l is associated with hypertension, antihypertensive response, and orthostatic hypotension. *Hypertension*, 54(4):796–801, 2009.
- [170] J. D. Duarte, S. T. Turner, B. Tran, A. B. Chapman, K. R. Bailey, Y. Gong, J. G. Gums, T. Y. Langaee, A. L. Beitelshes, R. M. Cooper-Dehoff, et al. Association of chromosome 12 locus with antihypertensive response to hydrochlorothiazide may involve differential yeats4 expression. *The pharmacogenomics journal*, 13(3):257, 2013.
- [171] S. T. Turner, K. R. Bailey, B. L. Fridley, A. B. Chapman, G. L. Schwartz, H. S. Chai, H. Sicotte, J.-P. Kocher, A. S. Rodin, and E. Boerwinkle. Genomic association analysis suggests chromosome 12 locus influencing antihypertensive response to thiazide diuretic. *Hypertension*, 52(2):359–365, 2008.
- [172] P. Svensson-Färbom, B. Wahlstrand, P. Almgren, J. Dahlberg, C. Fava, S. Kjeldsen, T. Hedner, and O. Melander. A functional variant of the nedd4l gene is associated with beneficial treatment response with β -blockers and diuretics in hypertensive patients. *Journal of hypertension*, 29(2):388–395, 2011.
- [173] C. W. McDonough, S. E. Burbage, J. D. Duarte, Y. Gong, T. Y. Langaee, S. T. Turner, J. G. Gums, A. B. Chapman, K. R. Bailey, A. L. Beitelshes, et al. Association of variants in nedd4l with blood

- pressure response and adverse cardiovascular outcomes in hypertensive patients treated with thiazide diuretics. *Journal of hypertension*, 31(4):698, 2013.
- [174] L. B. Spinasse, A. R. Santos, P. N. Suffys, E. S. Muxfeldt, and G. F. Salles. Different phenotypes of the *nat2* gene influences hydralazine antihypertensive response in patients with resistant hypertension. *Pharmacogenomics*, 15(2):169–178, 2014.
- [175] Ü. Yasar, C. Forslund-Bergengren, G. Tybring, P. Dorado, A. LLerena, F. Sjöqvist, E. Eliasson, and M.-L. Dahl. Pharmacokinetics of losartan and its metabolite e-3174 in relation to the *cyp2c9* genotype. *Clinical Pharmacology & Therapeutics*, 71(1):89–98, 2002.
- [176] Y. Takekuma, T. Takenaka, M. Kiyokawa, K. Yamazaki, H. Okamoto, A. Kitabatake, H. Tsutsui, and M. Sugawara. Evaluation of effects of polymorphism for metabolic enzymes on pharmacokinetics of carvedilol by population pharmacokinetic analysis. *Biological and Pharmaceutical Bulletin*, 30(3):537–542, 2007.
- [177] M. Honda, Y. Ogura, W. Toyoda, M. Taguchi, T. Nozawa, H. Inoue, and Y. Hashimoto. Multiple regression analysis of pharmacogenetic variability of carvedilol disposition in 54 healthy japanese volunteers. *Biological and Pharmaceutical Bulletin*, 29(4):772–778, 2006.
- [178] N. El Rouby and R. M. Cooper-DeHoff. Genetics of resistant hypertension: a novel pharmacogenomics phenotype. *Current hypertension reports*, 17(9):71, 2015.
- [179] A. I. Lynch, M. R. Irvin, B. R. Davis, C. E. Ford, J. H. Eckfeldt, and D. K. Arnett. Genetic and adverse health outcome associations with treatment resistant hypertension in genhat. *International journal of hypertension*, 2013, 2013.
- [180] E. P. Hong and J. W. Park. Sample size and statistical power calculation in genetic association studies. *Genomics & informatics*, 10(2):117–122, 2012.

Appendix A

HeartGenetics DNA Arterial Genetic Panel

Table A.1 : General genetic information for the 57 analysed variants. EA = effect alleles, † pharmacogenetic variants, ‡ variants excluded from the HTN dataset.

Variants	EA	Genes	Genes Full Name	Variants	EA	Genes	Genes Full Name
rs4646994†	DEL	ACE	Angiotensin I converting enzyme	rs5343	C	EDNRA	Endothelin Receptor Type A
rs4961‡	G	ADD1	Adducin 1 (alpha)	rs1458038	A	FGF5	Fibroblast Growth Factor 5
rs1048101†	C	ADRA1A	Adrenoceptor Alpha 1A	rs841	T	GCH1	GTP Cyclohydrolase 1
rs17333700	C	ADRA1A	Adrenoceptor Alpha 1A	rs2960306†	T	GRK4	G Protein-Coupled Receptor Kinase 4
rs1801252†	G	ADRB1	Adrenoceptor Beta 1	rs1024323†	T	GRK4	G Protein-Coupled Receptor Kinase 4
rs1801253†	C	ADRB1	Adrenoceptor Beta 1	rs1801058	T	GRK4	G Protein-Coupled Receptor Kinase 4
rs1042713	G	ADRB2	Adrenoceptor Beta 2	rs11739136†	A	KCNMB1	Potassium calcium-activated channel subfamily M regulatory beta subunit 1
rs4762†	T	AGT	Angiotensinogen	rs2301149†	C	KCNMB1	Potassium calcium-activated channel subfamily M regulatory beta subunit 1
rs699†	C	AGT	Angiotensinogen	rs2779249	G	NOS2	Nitric Oxide Synthase 2
rs5051†	A	AGT	Angiotensinogen	rs1799983†	G	NOS3	Nitric Oxide Synthase 3
rs5186†	C	AGTR1	Angiotensin II Receptor, Type 1	rs3918226	T	NOS3	Nitric Oxide Synthase 3
rs5182†	C	AGTR1	Angiotensin II Receptor, Type 1	rs5065†	T	NPPA	Natriuretic Peptide A
rs11091046	A	AGTR2	Angiotensin II Receptor, Type 2	rs5063	G	NPPA	Natriuretic Peptide A
rs1403543	G	AGTR2	Angiotensin II Receptor, Type 2	rs5068	T	NPPA	Natriuretic Peptide A
rs1799722†	T	BDKRB2	Bradykinin Receptor B2	rs5268‡	G	NPPC	Natriuretic Peptide C
rs2239101‡	T	CACNA1C	Calcium voltage-gated channel subunit alpha 1 C	rs2070951	G	NR3C2	Nuclear receptor subfamily 3 group C member 2
rs2357928†	G	CACNB2	Calcium voltage-gated channel auxiliary subunit beta 2	rs12750834†	T	REN	Renin
rs3781719	C	CALCA	Calcitonin-Related Peptide Alpha	rs5707	G	REN	Renin
rs848307	T	CLCNKA	Chloride voltage-gated channel Ka	rs11240688†	C	REN	Renin

Table A.2: General genetic information for the 57 analysed variants. EA = effect alleles, † pharmacogenetic variants, ‡ variants excluded from the HTN dataset.

Variants	EA	Genes	Genes Full Name	Variants	EA	Genes	Genes Full Name
rs12140311	T	CLCNKB	Chloride voltage-gated channel Kb	rs6693954	A	REN	Renin
rs75770792	T	CORIN	Corin, Serine Peptidase	rs3745368	G	RETN	Resistin
rs111253292	C	CORIN	Corin, Serine Peptidase	rs2228576†	G	SCNN1A	Sodium channel epithelial 1 alpha subunit
rs9932581	G	CYBA	Cytochrome B-245, Polypeptide	rs3759324	A	SCNN1A	Sodium channel epithelial 1 alpha subunit
rs72811418	T	CYBA	Cytochrome B-245, Polypeptide	rs11643718‡	G	SLC12A3	Solute carrier family 12 member 3
rs11191548†	T	CYP17A1	Cytochrome P450 family 17 subfamily A member 1	rs13306673†	C	SLC12A3	Solute carrier family 12 member 3
rs9332978	G	CYP4A11	Cytochrome P450 family 4 subfamily A member 11	rs3754777	A	STK39	Serine Threonine Kinase 39
rs6280	G	DRD3	Dopamine Receptor D3	rs1468326	C	WNK1	WNK Lysine Deficient Protein Kinase 1
rs212526	G	ECE1	Endothelin Converting Enzyme 1	rs880054†	C	WNK1	WNK Lysine Deficient Protein Kinase 1
rs5370†	T	EDN1	Endothelin 1				

Appendix B

Association Analysis Results for the Hypertension Trait

Table B.1: Association results for hypertension trait and quality control results in the HTN samples. When a given variant has a number of genotype groups, which is dependent on the genetic model used, smaller than 2 is not possible to perform the association test, for this reason NAs appear in the table.

Variants	Add	Dom	Rec	HWE	MAF	Variants	Add	Dom	Rec	HWE	MAF
	<i>P</i>	<i>P</i>	<i>P</i>	<i>P</i>			<i>P</i>	<i>P</i>	<i>P</i>	<i>P</i>	
rs4646994	1.3E-03†	0.01†	1.9E-03†	0.12	0.38	rs5343	0.42	0.36	0.62	0.25	0.28
rs1048101	0.59	0.44	0.47	0.42	0.47	rs1458038	0.87	0.62	1.00	0.74	0.28
rs17333700	0.64	0.42	1.00	0.40	0.20	rs841	0.77	0.53	1.00	1.00	0.18
rs1801252	0.82	0.91	0.59	0.22	0.12	rs2960306	0.33	0.27	0.21	1.00	0.38
rs1801253	0.98	1.00	0.92	0.25	0.28	rs1024323	0.62	0.76	0.33	0.48	0.38
rs1042713	0.17	0.09	0.37	0.18	0.35	rs1801058	0.41	0.84	0.19	0.58	0.41
rs4762	0.90	0.71	1.00	0.46	0.10	rs11739136	0.37	0.27	0.48	0.44	0.11
rs699	0.52	0.30	1.00	0.34	0.42	rs2301149	0.89	0.81	1.00	0.75	0.11
rs5051	0.55	0.29	0.90	0.59	0.43	rs2779249	0.96	0.86	0.84	0.30	0.33
rs5186	0.52	0.38	0.39	1.00	0.30	rs1799983	0.73	0.49	1.00	0.31	0.36
rs5182	0.92	0.71	0.92	0.22	0.41	rs3918226	0.03†	0.03†	NA	0.09	0.09
rs11091046	0.25	0.11	0.48	0.00‡	0.39	rs5065	0.02†	0.03†	0.53	0.09	0.18
rs1403543	0.13	0.06	0.72	0.00‡	0.39	rs5063	0.70	NA	0.70	1.00	0.03*
rs1799722	0.45	0.29	0.38	0.79	0.43	rs5068	0.85	NA	0.85	1.00	0.04*
rs2357928	0.89	0.67	0.91	0.14	0.47	rs2070951	0.81	0.66	0.70	0.23	0.45
rs3781719	0.47	0.56	0.28	0.51	0.29	rs12750834	0.72	0.45	0.76	1.00	0.16
rs848307	0.66	0.52	1.00	0.67	0.09	rs5707	0.66	0.42	1.00	0.68	0.21
rs12140311	1.00	1.00	NA	0.23	0.09	rs11240688	0.48	1.00	0.24	0.80	0.17
rs75770792	0.09	0.09	NA	1.00	0.00*	rs6693954	0.05†	0.03†	0.11	0.07	0.35
rs111253292	0.09	0.09	NA	1.00	0.00*	rs3745368	0.57	0.45	0.63	1.00	0.06
rs9932581	0.43	0.68	0.21	0.22	0.40	rs2228576	0.62	0.87	0.33	0.76	0.30
rs72811418	1.7E-03†	0.20	0.01†	0.23	0.07	rs3759324	0.20	0.07	0.61	0.67	0.21
rs11191548	0.73	0.45	1.00	0.37	0.08	rs13306673	0.79	NA	0.79	0.38	0.08
rs9332978	0.22	0.19	0.25	0.63	0.09	rs3754777	0.25	0.47	0.26	0.28	0.18
rs6280	0.58	0.47	0.68	1.00	0.40	rs1468326	0.14	0.08	0.75	0.63	0.16
rs212526	0.64	1.00	0.36	0.34	0.41	rs880054	0.11	0.04†	0.55	0.42	0.46
rs5370	0.92	0.76	1.00	0.24	0.21						

Add = additive model, Dom = dominant model, Rec = recessive model, HWE = Hardy-Weinberg equilibrium, NA = missing data, † p-value ≤ 0.05 , ‡ HWE is rejected, * MAF ≤ 0.05 .