

MACHINE LEARNING-BASED MULTI-ORGAN MULTI-MODALITY MODELLING OF END- ORGAN CHANGES WITH HYPERTENSION PROGRESSION IN THE UK BIOBANK

Mohanad Alkhodari

BSc., MSc.



Thesis presented for degree of Doctor of Philosophy in Medical Sciences

Supervisors: Prof. Paul Leeson

Dr. Winok Lapidaire

Dr. Abhirup Banerjee

Dr. Turkey Kart

Prof. Ahsan H. Khandoker

Michaelmas Term 2025

Approximate word count: ~43,464 words



WOLFSON
COLLEGE
UNIVERSITY OF OXFORD

Wolfson College

&

Cardiovascular Clinical Research Facility

Radcliffe Department of Medicine – Medical Sciences Division

University of Oxford

ACKNOWLEDGEMENTS

This thesis would not have been possible without the support of many people. I am deeply grateful to my family, whose constant encouragement has guided me through every stage of this journey. To my mother and father, thank you for always listening during difficult moments and for being there whenever I needed you, especially while travelling. To my siblings, who have always stood by me, bringing joy and easing my stress when it was most needed. To my wonderful fiancée, Lana, who started this journey with me from the very beginning and stayed by my side through every challenge. I am also thankful for the support of my supervisory team, Paul Leeson, Winok Lapidaire, Abhirup Banerjee, Turkey Kart, and Ahsan Khandoker, whose guidance has been essential to completing this work. To Paul, in particular, thank you for your steady support, understanding, and care throughout the most demanding phases of my DPhil. Your insight and encouragement helped me reach levels I did not anticipate, and you taught me the value of shaping my own ideas and directing my research with confidence. To Abhi, thank you for being there whenever I needed someone to talk to and for giving me space to express my stress. To Ahsan, heartfelt thanks for believing in me, giving me the opportunity to pursue my dream, and guiding me through this work while helping to resolve every challenge along the way. I am also grateful to our collaborator, Leontios Hadjileontiadis, whose consistent support has helped me achieve my ambitions. To my colleague, Prenali Sattwika, thank you for your constant support, both in the lab, particularly with the systematic review, and in helping me work through personal challenges. I am also very grateful to have supportive colleagues at CCRF, whose encouragement and assistance have helped me throughout this journey. Finally, I give thanks to Allah for granting me the courage and strength to pursue this knowledge.

DEDICATIONS

For my mum, dad, siblings, and fiancée with all my love

ABBREVIATIONS

BP	Blood pressure
SBP	Systolic blood pressure
DBP	Diastolic blood pressure
MR	Magnetic resonance
CMR	Cardiac magnetic resonance
CUS	Carotid ultrasound
BMI	Body mass index
CO	Cardiac output
LVS	Left ventricular stroke volume
LVED	Left ventricular end-diastolic volume
IMT	Intima-media thickness
WMH	White matter hyperintensities
AI	Artificial Intelligence
cPCA	Contrastive principal component analysis
cTI	Contrastive trajectory inference
k-NN	k -nearest neighbours
CBPE	Confidence-based performance estimation
JS	Jensen-Shannon distance
LME	Linear mixed-effects
ARIC	Atherosclerosis risk in communities study
PVS	Pregnancy complications vascular study
NP	Normotensive pregnancy
HDP	Hypertensive disorders of pregnancy
PET	Pre-eclampsia (or preeclamptic toxemia)
EPET	Early preeclampsia
LPET	Late preeclampsia
PIH	Pregnancy-induced hypertension

ABSTRACT

MACHINE LEARNING-BASED MULTI-ORGAN MULTI-MODALITY MODELLING OF END-ORGAN CHANGES WITH HYPERTENSION PROGRESSION IN THE UK BIOBANK

**THESIS PRESENTED FOR THE DEGREE DOCTOR OF PHILOSOPHY IN
MEDICAL SCIENCES AT THE UNIVERSITY OF OXFORD**

Mohanad Alkhodari

Wolfson College, Michaelmas Term 2025

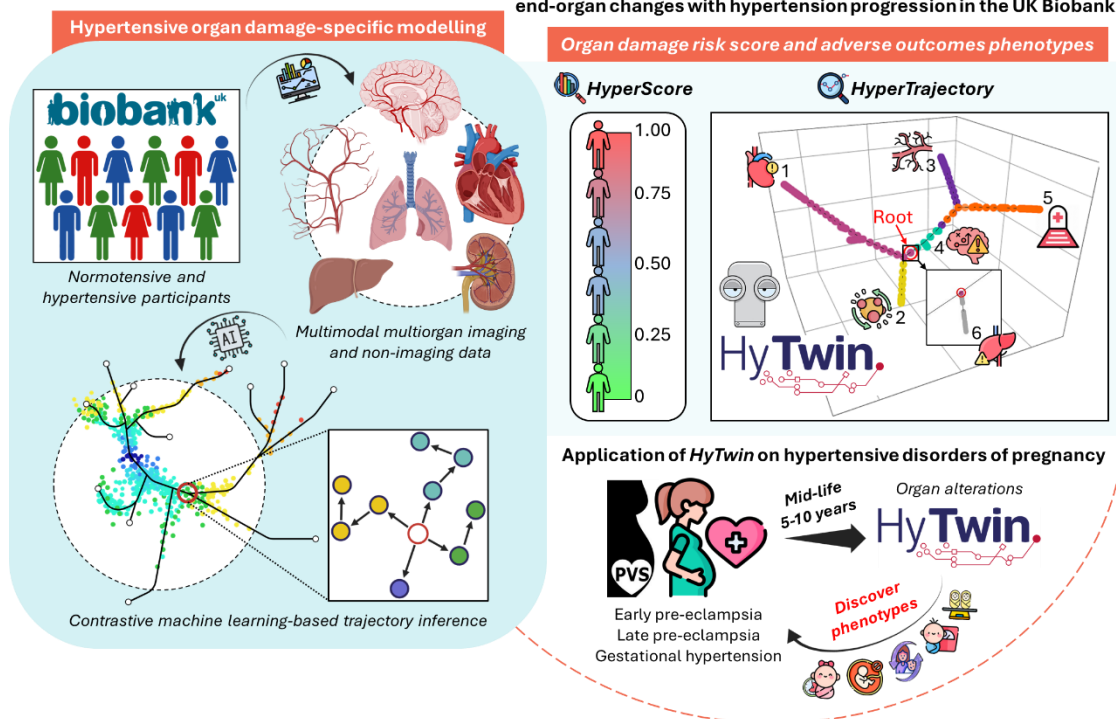
Hypertension, or elevated blood pressure, is the major cause of premature death worldwide, accounting for nearly one-third of global mortality. It increases the risk of adverse outcomes in vital organs including brain, heart, kidneys, and vasculature. However, current clinical management guidelines still primarily focus on controlling blood pressure, with limited attention to the structural and functional alterations that develop in target organs over years of exposure. Moreover, traditional assessments rely on risk factors that do not necessarily reflect the extent of damage to specific organs or guide targeted therapeutic interventions. In this thesis, I sought to provide a comprehensive characterisation of the multi-organ damage caused by hypertension by introducing a global organ damage risk score and identifying organ-specific disease outcome trajectories, thereby laying the foundation for personalised assessments in hypertension management based on organ-specific information.

A large cohort from the UK Biobank imaging substudy was selected as the baseline dataset, comprising both multimodal imaging and non-imaging features. First, I modelled the spectrum of multi-organ damage in this cohort using a machine learning-

based contrastive trajectory inference (cTI) approach, referred to in this thesis as the *HyTwin* model. This framework was used to derive a global organ damage score that maps the continuum from healthy (zero) to severe (one), termed the *HyperScore*, and to identify predominant organ-specific disease progression pathways within the model, referred to as *HyperTrajectory*. These identified phenotypes were associated with key clinical outcomes, including mortality and advanced damage in the heart, brain, and vascular system. Then, I demonstrated the generalisability of the *HyTwin* model by showing that the identified phenotypes were reproducible in an external, out-of-domain cohort from the Atherosclerosis Risk in Communities (ARIC) study. Finally, I evaluated the potential of the *HyTwin* model as a framework to investigate the lifecourse of blood pressure-associated end-organ changes in specific subgroups of individuals with hypertension, specifically, hypertensive disorders of pregnancy (HDP). Here, I used a longitudinal clinical imaging dataset that included women with different pregnancy complications (*HyPrevent* dataset) available in Oxford. I characterised end-organ damage in individuals according to imaging data acquired in mid-life and related this back to their pregnancy characteristics to discover distinct pregnancy-related phenotypes associated with certain *HyperTrajectories*, characterised in my prior work. Future work is needed to investigate the long-term efficacy of the *HyTwin* findings, explore ethnicity-related characteristics of the *HyperTrajectory*, and evaluate the reproducibility of the HDP phenotypes in independent patient cohorts. This could strengthen the *HyTwin*'s clinical utility by enabling early and accurate risk stratification and improving its applicability across diverse patient groups, while supporting personalised treatment pathways according to each individual's organ damage profile.

GRAPHICAL ABSTRACT

Machine learning-based multi-organ multi-modality modelling of end-organ changes with hypertension progression in the UK Biobank



LIST OF PUBLICATIONS

Original Manuscripts Published During DPhil

Published:

1. **Alkhodari M**, Lapidaire W, Xiong Z, Hadjileontiadis L, Khandoker A, Banerjee A, Leeson P. HyperScore: a unified measure to model hypertension progression using multi-modality measurements and semi-supervised learning. 17th IEEE international conference on bioinformatics and biomedicine (BIBM). Istanbul, Turkey. 2023.
2. **Alkhodari M**, Xiong Z, Khandoker A, Hadjileontiadis L, Leeson P, Lapidaire W. The role of artificial intelligence in hypertensive disorders of pregnancy: towards personalized healthcare. Expert Review of Cardiovascular Therapy, Taylor & Francis. 2023;21(7):531-543.

Under-review:

3. **Alkhodari M**, Lapidaire W, Kart T, Xiong Z, Krasner S, Fletcher A, Savage N, Baumeister T, Ohuma E, Namburete A, Lamata P, Iturria-Medina Y, Chappel T, Thilaganathan B, Fraser A, Mackillop L, McManus R, Ntusi N, Hadjileontiadis L, Khandoker A, Lewandowski A, Banerjee A, Leeson P. Contrastive machine learning to quantify hypertensive multi-organ damage and identify new disease phenotypes: a multi-national multi-modal study.
4. **Alkhodari M**, Sattwika P, Cutler H, Milner G, Kart T, Hadjileontiadis L, Khandoker A, Banerjee A, Leeson P, Lapidaire W. Artificial intelligence applied to imaging for prediction, phenotyping, and management of hypertension: a systematic review.
5. **Alkhodari M**, Sattwika PD, Wilmes N, Baktash V, Lapidaire W, Leeson P, Kart T. HyLLM: Explaining Hypertensive End-Organ Damage in Multi-modal Imaging Using a Multi-stage Large Language Model Framework.

Under-preparation:

6. **Alkhodari M**, Krasner S, Cutler HR, Sattwika PD, Kart T, Savage N, Baumeister T, Ohuma E, Namburete A, Lamata P, Iturria-Medina Y, Chappel T, Thilaganathan B, Fraser A, Mackillop L, McManus R, Ntusi N, Hadjileontiadis L, Khandoker A, Banerjee A, Lewandowski A, Lapidaire W, Leeson P. Discovering pregnancy phenotypes associated with mid-life hypertensive multi-organ damage using contrastive machine learning.

Collaborative Manuscripts Published During DPhil

Published:

1. Fletcher A, **Alkhodari M**, Krasner S, Xiong Z, Alsharqi M, Lewandowski A J, Taboada D, Garbi M, Neubauer S, Lapidaire W, Leeson P. Global diastology index: a novel and personalised diastolic function score derived by machine learning from 33104 echocardiograms and validated against invasive haemodynamics. European Heart Journal (EHJ). 2025;46(Supplement_1):ehaf784-025.
2. Kart T, **Alkhodari M**, Lapidaire W, Banerjee A, Lewandowski A, Leeson P. Deep Learning-based Modelling of Complex Hypertensive Multi-Organ Damage with Uncertainty Quantification from Simple Clinical Measures. 18th IEEE international conference on bioinformatics and biomedicine (BIBM). Lisbon, Portugal. 2024.
3. Kart T, **Alkhodari M**, Lapidaire W, Leeson P. Modelling relations between blood pressure, cardiovascular phenotype, and clinical factors using large scale imaging data. European Heart Journal - Cardiovascular Imaging, European Society of Cardiology. 2023;jead161.
4. Sattwika P, Schuermans A, Cutler H, **Alkhodari M**, Anggraeni V, Nurdianti D, Lapidaire W, Leeson P, Lewandowski A. Multi-Organ Phenotypes of Offspring Born Following Hypertensive Disorders of Pregnancy: A Systematic Review. Journal of the American Heart Association (JAHA). 2024;13(21):e033617.
5. Cutler H, Barr L, Sattwika P, Frost A, **Alkhodari M**, Kitt J, Lapidaire W, Lewandowski L, Leeson P. Temporal Patterns of Pre- and Post-Natal Target Organ Damage Associated with Hypertensive Pregnancy: A Systematic Review. European Journal of Preventive Cardiology, European Association of Preventive Cardiology. 2023;zwad275.

Under-review:

6. Fletcher A, **Alkhodari M**, Krasner S, Xiong Z, Alsharqi M, Savage N, Lewandowski A, Costanzo P, Taboada D, Garbi M, Neubauer S, Lapidaire W, Leeson P. Development and Invasive Validation of the Global Diastology Index: A Novel and Personalised Machine Learning Derived Diastolic Function Score.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	2
DEDICATIONS	3
ABBREVIATIONS	4
ABSTRACT	5
GRAPHICAL ABSTRACT	7
LIST OF PUBLICATIONS	8
TABLE OF CONTENTS	10
LIST OF FIGURES	14
LIST OF TABLES	16
1. INTRODUCTION	17
1.1 Hypertension and its assessment	18
1.1.1 Definition and current guidelines	18
1.1.2 Risk factors and assessment parameters	20
1.2 Hypertension-induced end-organ damage	21
1.2.1 Associations with elevated blood pressure	21
1.2.2 Cardiovascular remodelling	22
1.2.3 Cerebrovascular remodelling	24
1.2.4 Kidney dysfunction	25
1.2.5 Other vasculature alterations	26
1.3 The use of artificial intelligence in hypertension assessment	28
1.3.1 Definition and technical background	28
1.3.2 Applications in hypertension assessment	31
1.3.3 Contrastive trajectory inference (cTI)	32
1.4 Thesis aims and objectives	35
2. SYSTEMATIC REVIEW AND META-ANALYSIS	37
2.1 Abstract	38
2.2 Introduction	40
2.3 Methods	41
2.3.1 Study design	41
2.3.2 Search strategy and eligibility criteria	41
2.3.3 Population, intervention, and outcomes	42

2.3.4	Selection of studies.....	43
2.3.5	Data extraction	43
2.3.6	Risk of bias assessment	43
2.3.7	Thematic data synthesis.....	44
2.3.8	Meta-analysis.....	45
2.4	Results	46
2.4.1	Search results.....	46
2.4.2	General characteristics of studies	46
2.4.3	Thematic analysis	52
2.4.3.1	AI Task 1: Identifying hypertension or its stages	55
2.4.3.2	AI Task 2: Differentiating hypertension from other diseases.....	58
2.4.3.3	AI Task 3: Characterising target end-organ changes	59
2.4.3.4	AI Task 4: Identifying phenotypes and hypertension subgroups	61
2.4.3.5	AI Task 5: Defining risk of future hypertension-related events	62
2.5	Discussion and conclusion	63
2.5.1	Study limitations.....	65
2.5.2	Conclusions	66
3.	<i>HYTWIN DEVELOPMENT AND MODELLING</i>	67
3.1	Abstract	68
3.2	Introduction	70
3.3	Methods	71
3.3.1	Dataset for model development.....	71
3.3.2	Development of end-organ model.....	72
3.3.2.1	Pre-processing and data preparation.....	72
3.3.2.2	Feature selection to reduce redundancy	72
3.3.2.3	Adjusting covariates and harmonisation	73
3.3.2.4	Definition of groups for contrastive analysis	73
3.3.2.5	Dimensionality reduction and cPCA	74
3.3.2.6	Contrastive trajectory inference (cTI)	76
3.3.3	Generation of <i>HyperScore</i> and <i>HyperTrajectory</i> predictions	77
3.3.3.1	Definition of <i>HyperScore</i> and <i>HyperTrajectory</i>	77
3.3.3.2	<i>HyperScore</i> and <i>HyperTrajectory</i> unseen predictions.....	78
3.3.4	Internal model validation.....	79
3.3.4.1	Stability of model generation of <i>HyperScores</i>	79

3.3.4.2	Stability of allocation of HyperTrajectory probability	80
3.3.5	Clinical relevance of <i>HyperScore</i>	81
3.3.6	Clinical interpretation of <i>HyperTrajectory</i> allocation.....	83
3.3.7	Statistical and regression analysis	84
3.4	Results	84
3.4.1	Dataset for model development.....	84
3.4.2	Development of end-organ model	86
3.4.3	Internal validation.....	88
3.4.3.1	Stability of HyperScore estimation	88
3.4.3.2	Stability of HyperTrajectory allocation.....	90
3.4.4	Clinical relevance of <i>HyperScore</i>	91
3.4.4.1	Association with demographics and organ phenotypes.....	91
3.4.4.2	Time-to-event analysis	92
3.4.4.3	Comparison to conventional risk prediction clinical tools.....	95
3.4.5	Clinical relevance of <i>HyperTrajectory</i>	96
3.4.5.1	Variation in end-organ phenotypes by trajectory.....	96
3.4.5.2	Variation in events and time-to-event across HyperTrajectories.....	97
3.5	Discussion and conclusion	101
3.5.1	Study limitations.....	104
3.5.2	Conclusions	105
4.	<i>HYTWIN</i> EXTERNAL TESTING.....	106
4.1	Abstract	107
4.2	Introduction	109
4.3	Methods	110
4.3.1	Dataset for external testing.....	110
4.3.2	Variables matching between datasets.....	111
4.3.3	Modality-wise differences and alignment	111
4.3.4	Transformation to contrastive principal components.....	112
4.3.5	Evaluation of <i>HyperScore</i> prediction	112
4.3.6	Clinical relevance of <i>HyperScore</i> and <i>HyperTrajectory</i>	113
4.4	Results	113
4.4.1	External dataset characteristics	113
4.4.1	Allocation of ARIC individuals on <i>HyperTrajectories</i>	115
4.4.2	<i>HyperScore</i> prediction similarities	116

4.4.3	Clinical relevance of <i>HyperScore</i> and <i>HyperTrajectory</i>	117
4.5	Discussion and conclusion	120
4.5.1	Study limitations.....	123
4.5.2	Conclusions	123
5.	<i>HYTWIN</i> AND HYPERTENSIVE DISORDERS OF PREGNANCY	125
5.1	Abstract	126
5.2	Introduction	128
5.3	Methods.....	129
5.3.1	Dataset description and participants.....	129
5.3.2	Data preparation	130
5.3.3	Transformation to contrastive principal components.....	131
5.3.4	Summarisation of information across pregnancy subgroups	132
5.3.5	Summarisation of information based on <i>HyperScore</i> and <i>HyperTrajectory</i> 132	
5.3.6	Odds ratio analysis of <i>HyperTrajectories</i>	133
5.3.7	Characterising the observed phenotypes	133
5.3.8	UK Biobank validation.....	134
5.3.9	Statistical analysis	135
5.4	Results	136
5.4.1	Characteristics of the PVS cohort	136
5.4.2	PVS and the estimated <i>HyperScores</i>	139
5.4.3	PVS and the <i>HyperTrajectory</i> allocation	140
5.4.4	Odds ratios of pregnancy factors.....	144
5.4.5	Predominant characteristics of identified pregnancy phenotypes	145
5.4.6	Reproducible phenotypes on UKB pregnancy subset	146
5.5	Discussion and conclusion	149
5.5.1	Study limitations.....	153
5.5.2	Conclusions	153
6.	CONCLUSIONS	155
6.1	Future questions and research directions.....	158
7.	REFERENCES	162

LIST OF FIGURES

Figure 1.1 Hypertension-induced cardiovascular damage.	23
Figure 1.2 Hypertension-induced cerebrovascular damage.	24
Figure 1.3 Hypertension-induced kidney damage.	26
Figure 1.4 Types and applications of artificial intelligence (AI) in personalised healthcare.	29
Figure 1.5 Contrastive principal component analysis (cPCA).	33
Figure 1.6 Contrastive trajectory inference (cTI) with minimum spanning tree (MST).	34
Figure 2.1 Flow diagram of the study selection process.	47
Figure 2.2 Overall summary of the population included in the selected studies and type of used data.	48
Figure 2.3 Overall characterisation of included studies based on bias assessment and study type.	49
Figure 2.4 Overall characterisation of included studies based on artificial intelligence approaches.	50
Figure 2.5 Thematic tasks of included studies.	51
Figure 2.6 Overall thematic analysis of AI Tasks.	55
Figure 2.7 Thematic analysis of the included studies in AI Task 1.	56
Figure 2.8 Meta-analysis.	57
Figure 2.9 Summary receiver operating characteristics (SROC) curves of the included studies with their corresponding funnel plots.	57
Figure 2.10 Thematic analysis of the included studies in AI Task 2.	58
Figure 2.11 Thematic analysis of the included studies in AI Task 3.	60
Figure 2.12 Thematic analysis of the included studies in AI Task 4.	61
Figure 2.13 Thematic analysis of the included studies in AI Task 5.	63
Figure 3.1 Preparation and selection of the optimal contrastive covariance matrix for the normotensive and hypertensive participants in the overall dataset.	75
Figure 3.2 Contrastive trajectory inference (cTI) and construction of the minimum spanning tree (MST).	76
Figure 3.3 Machine learning-based multi-organ and multi-modality modelling of the end-organ changes associated with hypertension.	87
Figure 3.4 Internal validation of the modelling approach and HyperScore estimation performance.	88

Figure 3.5 Additional internal validation of the HyperScore using confidence-based performance estimation (CBPE) analysis.....	89
Figure 3.6 Internal validation performance of the HyperTrajectory.....	90
Figure 3.7 Association between end-organ phenotypes and hypertension progression.	92
Figure 3.8 Prediction of adverse outcomes and analysis of survival.....	94
Figure 3.9 Comparison of HyperScore and HyperTrajectory relative to conventional and clinically valid scores.	95
Figure 3.10 Characterisation of the HyperTrajectory in the UK Biobank.	96
Figure 3.11 Overall summary of the formed HyperTrajectories in the UK Biobank highlighting their HyperScore range, major characteristics, phenotype, and outcomes.	98
Figure 4.1 Characterisation of the HyperTrajectory in the overall ARIC cohort.	116
Figure 4.2 External validation of HyperScore distribution and progression patterns in ARIC and UK Biobank.	117
Figure 4.3 External testing of HyperScore and HyperTrajectory on the Atherosclerosis Risk in Communities Study (ARIC) dataset (right) and comparison with the UK Biobank dataset (left).	119
Figure 5.1 Major characteristics of the PVS cohort for overall normotensive pregnancy (NP) and hypertensive disorder of pregnancy (HDP) groups, alongside their corresponding subgroups.....	139
Figure 5.2 Distribution of pregnancy subgroups and major pregnancy factors based on HyperScore stages.	140
Figure 5.3 Distribution of pregnancy subgroups and major pregnancy factors based on HyperTrajectories.	141
Figure 5.4 Odds ratios (OR) for pregnancy subgroups distribution and major pregnancy factors across HyperTrajectories in PVS.	144
Figure 5.6 Odds ratios (OR) for available pregnancy subgroups distribution and major pregnancy factors across HyperTrajectories in the UK Biobank pregnancy subset. ..	148
Figure 5.7 Ratio of odds ratios (ROOR) analysis between PVS and the UK Biobank pregnancy subset across HyperTrajectories.	149
Figure 5.8 Overall summary of the HyperTrajectories linking general hypertensive organ damage characteristics by pregnancy-related phenotypes.....	150

LIST OF TABLES

Table 1.1 Blood pressure categorisation according to international management guidelines.....	19
Table 2.1 Key terms used for the search strategy across different databases.	42
Table 2.2 Summary of studies based on major thematic tasks to characterise hypertension.	52
Table 3.1 Baseline characteristics of the UK Biobank cohort included in this study. .	85
Table 3.2 Characteristics of the HyperTrajectory in the advanced disease state (top 5% HyperScores) in the UK Biobank cohort.	99
Table 4.1 Characteristics of the external testing dataset, ARIC, compared with the UK Biobank for selected age ranges.	114
Table 4.2 Characteristics of the HyperTrajectory in the advanced disease state (top 5% HyperScores) in the ARIC cohort.	119
Table 5.1 Summary of pregnancy characteristics of the PVS cohort.	137
Table 5.2 Characteristics of the HyperTrajectory in the overall PVS cohort.	143
Table 5.3 Overall summary of the HyperTrajectories based on pregnancy factors in PVS.....	146
Table 5.4 Characteristics of the HyperTrajectory in the UK Biobank pregnancy subset.	147

1. INTRODUCTION

1.1 Hypertension and its assessment

1.1.1 Definition and current guidelines

Hypertension, or elevated blood pressure, is the major cause of premature death worldwide, affecting over 1.3 billion adults, which accounts for nearly one-third of the global adult population¹. It is the leading cause of acute medical conditions such as heart attacks, strokes, kidney failure, and vision loss². In England, it is estimated that more than 30% of the adult population have hypertension (about 1 in every 4), which accounts for nearly 14 million adults^{3,4}. Moreover, hypertension is the third biggest risk factor for death and disability in England, after smoking and poor diet, costing the healthcare system over £2.1 billion annually to manage its complications⁵.

Blood pressure elevation can manifest in different forms. Primary hypertension, formerly known as essential hypertension, refers to persistently elevated blood pressure that cannot be attributed to a specific secondary condition⁶. It is the most common type, accounting for approximately 90% of all hypertension cases⁷, and it is often linked to genetics, obesity, and other risk factors. Secondary hypertension accounts for the remaining 10% of cases, where the elevation of blood pressure has a clear identifiable underlying cause including diseases such as renal artery stenosis, obstructive sleep apnoea, or cancer⁸.

Hypertension is usually diagnosed by taking average clinical systolic blood pressure (SBP) and diastolic blood pressure (DBP) measures. According to the American Heart Association (AHA) guidelines^{9,10}, stage 1 hypertension is defined as an SBP ≥ 130 mmHg and/or a DBP ≥ 80 mmHg. On the other hand, the UK National Institute for Health and Care Excellence (NICE)¹¹, European Society of Hypertension (ESH)¹²

Table 1.1 | Blood pressure categorisation according to international management guidelines.

ACC/AHA 2017 stages ⁹	ACC/AHA 2025 stages ¹⁰	NICE 2019 stages ¹¹	ESH 2023 stages ¹²	JNC-7 stages ¹³	Blood pressure thresholds
Normal	Normal	Normal	Normal	Normal	<120 SBP and <80 DBP
Elevated	Elevated	Normal	Normal	Pre-hypertension	>120 SBP and/or <90 DBP
Hypertension stage 1	Hypertension stage 1	Normal	High-normal	Pre-hypertension	>130 SBP and/or >80 DBP
Hypertension stage 2	Hypertension stage 2	Hypertension stage 1	Hypertension stage 1	Hypertension stage 1	>140 SBP and/or >90 DBP
Hypertension stage 2	Hypertension stage 2	Hypertension stage 2	Hypertension stage 2	Hypertension stage 2	>160 SBP and/or >100 DBP
Hypertension crisis	Severe Hypertension	Hypertension stage 3	Hypertension stage 3	Hypertension stage 2	>180 SBP and/or >120 DBP

Text in bold indicates agreement across all hypertension management guidelines.

and Seventh Report of Joint National Committee (JNC-7)¹³ define stage 1 hypertension as an SBP $\geq$ 140 mmHg and/or a DBP $\geq$ 90 mmHg. All guidelines agree in defining stage 2 hypertension as an SBP $\geq$ 160 mmHg and/or a DBP $>$ 100 mmHg, and normal as having an SBP/DBP of <120/80 mmHg. The different blood pressure stages according to these guidelines are summarised in **Table 1.1**.

Current guidelines recommend the management of hypertension primarily through pharmacological interventions, such as blood pressure- and cholesterol-lowering medications. However, blood pressure readings can fluctuate with everyday activities and may be distorted by white-coat or masked hypertension, where values differ depending on whether measurements are taken in the clinic or at home¹⁴. Moreover, the persistence of elevated blood pressure is also driven by a range of modifiable and non-modifiable risk factors, underscoring the multifaceted nature of hypertension assessment. Thus, relying solely on blood pressure measurements may be insufficient for clinical decision-making, as levels can be transiently influenced by these risk factors.

1.1.2 Risk factors and assessment parameters

Elevated blood pressure is associated with several risk factors that increase the likelihood of adverse cardiovascular events. Individuals who are overweight or obese are at higher risk of hypertension-related complications, as excess body fat contributes to increased vascular resistance and cardiac workload¹⁵. Additionally, daily physical activity and diet influence the body's response to hypertension, as insufficient physical activity can impair arterial blood flow¹⁶ and excessive sodium intake, *i.e.*, salt, can lead to fluid retention¹⁷. Other modifiable risk factors include smoking, excessive alcohol consumption, and chronic stress, alongside non-modifiable ones such as aging, family history, genetics, and sex⁹.

Sex is a major risk factor influencing blood pressure elevation. A higher incidence of hypertension and poorer outcomes are often reported in men compared to women¹⁸. However, women are at risk of hypertensive disorders of pregnancy (HDP), which affect nearly 10% of pregnant women worldwide^{19,20}. Generally, HDP is associated with an increased risk of developing postpartum cardiovascular disorders, which are the leading cause of maternal death in the United Kingdom²¹. According to the International Society for the Study of Hypertension in Pregnancy (ISSHP) guidelines²², HDP includes gestational hypertension, also known as pregnancy-induced hypertension (PIH), where blood pressure rises above 140/90 mmHg during pregnancy but returns to normal postpartum²³. It also includes preeclamptic toxemia (PET), or preeclampsia, which occurs after 20 weeks of gestation and is characterised by elevated blood pressure and presence of proteinuria (≥ 300 mg/24h)²⁴. Preeclampsia can progress to eclampsia, a severe condition in which pregnant women experience seizures due to markedly elevated blood pressure ($>160/110$ mmHg). In this thesis, I focus on HDP in **Chapter 5**.

Characterising hypertension through associated risk factors has improved the estimation of future adverse events. Several risk prediction models, including QRISK3²⁵, Framingham²⁶, and AHA²⁷ scores, incorporate risk factors to estimate the 10-year risk of developing cardiovascular-related events. QRISK3 score, introduced in 2017, is the preferred tool in England and was validated on a large cohort of individuals aged 25-84 years, predicting events such as heart attacks and strokes²⁵. The Framingham score, developed from examinations conducted between 1971 and 1974 as part of the Framingham Heart Study in the United States, estimates the 10-year risk of developing coronary artery disease (CAD) in individuals aged 30-75 years²⁶. The AHA score, based on the 2013 AHA guidelines, estimates the 10-year risk of atherosclerotic cardiovascular disease (ASCVD) events in individuals aged 40-80 years²⁷. Despite their widespread use in risk estimation, these models have limitations, including the overestimation of risk in certain populations and the reliance on classic risk factors that may not fully reflect an individual's dynamic risk profile²⁸. Therefore, there remains a need for more comprehensive assessments of hypertension, particularly to evaluate the progressive damage it can inflict on vital organs, altering their structure and function.

1.2 Hypertension-induced end-organ damage

1.2.1 Associations with elevated blood pressure

The prolonged, uncontrolled exposure to hypertension leads to structural and functional alterations in the heart, brain, kidneys, and vasculature, predisposing individuals to vascular events and mortality²⁹. Notably, such damage in major organs can be observed

in individuals with mildly elevated blood pressure, suggesting that earlier intervention could benefit those with mild hypertension³⁰. However, because hypertension is often asymptomatic in its early stages, it frequently goes unnoticed and untreated, allowing gradual progression to a symptomatic condition followed by organ dysfunction³¹. Moreover, individuals with the same blood pressure levels may exhibit different degrees of organ damage, reflecting variability in individual response to the disease.

The extended exposure to hypertension, transient blood pressure fluctuations during measurement, and genetic or environmental differences in organ response may explain the variability in organ damage and disease outcomes associated with elevated blood pressure^{32,33}. Treatment based on blood pressure levels alone may under- or over-treat the underlying organ damage³⁴. Additionally, as patients are often asymptomatic, engaging those with significant end-organ damage in medical and lifestyle prevention or treatment strategies remains a significant challenge. Given the complex nature of organ responses to hypertension, a more detailed assessment of structural and functional organ characteristics could offer deeper insights into the extent and severity of damage. Imaging has played a crucial role in the assessment of multiple organs across various diseases, particularly in identifying end-organ changes such as left ventricular hypertrophy (LVH), cerebral haemorrhages, renal fibrosis, and vascular thickening³⁵. Imaging complements baseline evaluation procedures which usually include risk factors, 12-lead electrocardiography (ECG) and laboratory-based biomarkers.

1.2.2 Cardiovascular remodelling

Imaging of the heart and the cardiovascular system reveals early alterations in response to hypertension. Echocardiography is most commonly used as the first-line imaging modality for assessing cardiac damage, owing to its accessibility, non-invasiveness, and

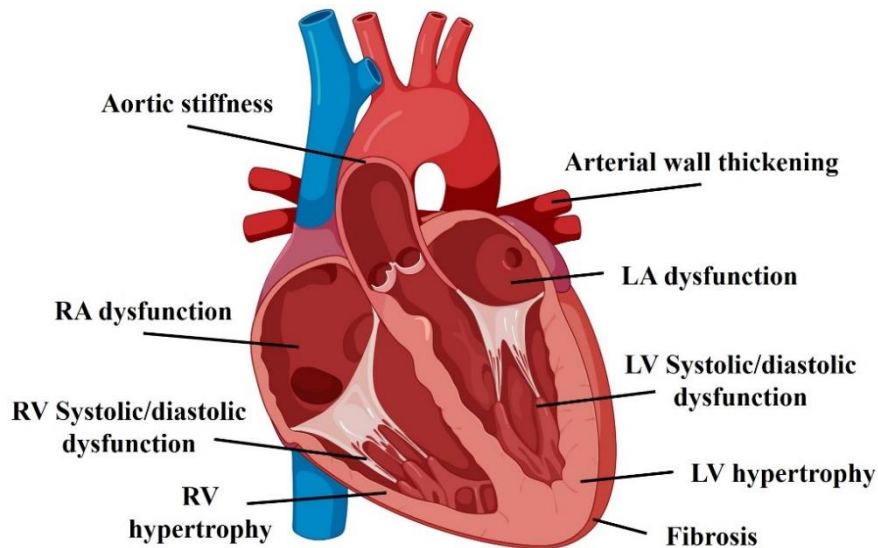


Figure 1.1 | Hypertension-induced cardiovascular damage. Cardiovascular alterations include left ventricular (LV), right ventricular (RV), left atrium (LA), and right atrium (RA) dysfunction and vessel wall thickening.

ability to provide detailed anatomical information³⁵. It is widely recommended by the ESH guidelines³⁶ for identifying LVH, left atrial enlargement, and hypertension-mediated organ damage. There is growing evidence that geometric and structural details, such as left ventricle (LV) and right ventricle (RV) chamber size, myocardial mass, and wall thickness, are closely associated with elevated blood pressure³⁷ (**Figure 1.1**). In particular, they are affected by chronic pressure overload and increased hemodynamic stress due to hypertension, often resulting in increased mass, predominantly in the form of concentric hypertrophy³⁸.

In terms of functional assessment, echocardiography incorporates Doppler technology which enables evaluation of cardiac chamber function and blood flow dynamics within the heart and arteries. It provides insight into cardiac mechanics, particularly LV volume and systolic function, with evidence showing that cardiac output tends to increase with progression of SBP³⁹. However, when higher spatial resolution is required, cardiac magnetic resonance (CMR) imaging is often preferred and is considered the gold-standard modality for assessing cardiac function⁴⁰. Literature has demonstrated its

accuracy and reproducibility in estimating LV volume, mass, and ejection fraction⁴¹, often surpassing echocardiography. CMR provides detailed tissue characterisation of the LV endocardial border in end-diastole and end-systole using techniques such as late gadolinium enhancement (LGE) and T1/T2 mapping⁴². In hypertension, LV stroke and end-diastolic volumes are often elevated due to increased filling pressures or myocardial stiffness⁴³, which are detected using CMR. Moreover, CMR has enabled accurate detection and quantification of diffuse myocardial fibrosis, which is a common pathological endpoint for hypertension and a key cause of diastolic dysfunction⁴⁴.

1.2.3 Cerebrovascular remodelling

The brain is significantly affected by elevated blood pressure, which can damage cerebral blood vessels and disrupt blood flow and oxygen delivery to brain tissue⁴⁵. As hypertension progresses, cerebral blood flow gets interrupted due to increases arterial media thickness, reduced lumen diameter, and the formation of atherosclerotic plaques. These changes (**Figure 1.2**) impair vessel dilation and increase blood-brain barrier permeability and risk of microbleeds⁴⁶. Consequently, hypertension elevates the risk of ischemic and haemorrhagic strokes, as well as cognitive decline and dementia⁴⁷. These

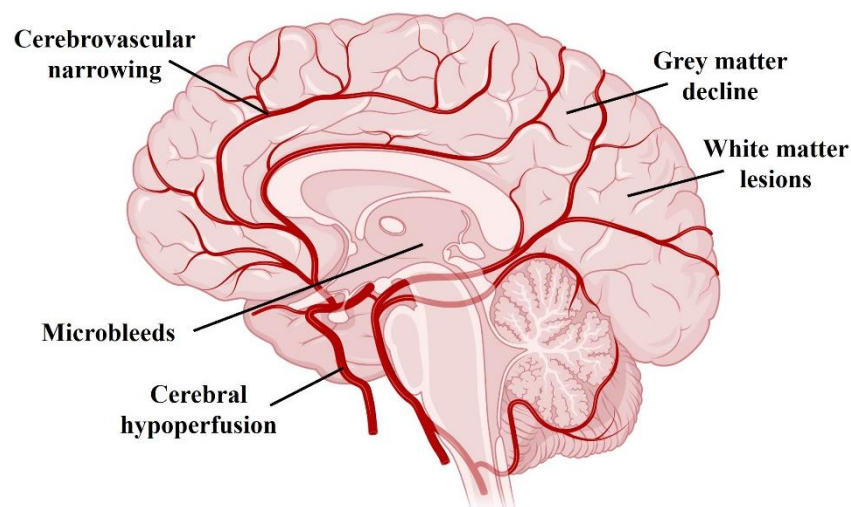


Figure 1.2 | Hypertension-induced cerebrovascular damage. Cerebrovascular alterations include white matter lesions, vessel narrowing, and microbleeds.

conditions are often preceded by cerebral alterations, *e.g.*, cerebral hypoperfusion⁴⁸ and white matter lesions⁴⁹ that can be detected early using brain imaging.

Brain magnetic resonance (MR) imaging is a commonly used modality for assessing cerebral alterations caused by hypertension⁵⁰. Owing to its high spatial resolution, brain MR reveals structural details of grey matter, white matter, cerebral vessels, and cerebrospinal fluid, which are highly affected by elevated blood pressure⁵¹. Several studies have reported increased white matter hyperintensities (WMH), higher prevalence of cerebral microbleeds, enlarged perivascular spaces, and brain atrophy using brain MR when characterising cerebral small vessel disease and hypertension-related damage^{52,53}. Moreover, progressive elevations in blood pressure have been associated with a decline in grey matter volume and impaired white matter microstructure, causing cognitive decline and dementia⁵⁴.

1.2.4 Kidney dysfunction

The renal system gets also disrupted in presence of elevated blood pressure. Hypertension leads to scarring of kidney tissues due to narrowing of the surrounding arteries. This damage impairs their ability to filter blood and regulate fluids, salts, and electrolytes⁵⁵. Moreover, it reduces the kidneys' capacity to remove waste and excess fluids, often resulting in kidney failure, salt retention, and alterations in hormonal systems⁵⁶. Elevated blood pressure can lead to chronic kidney disease (CKD), while CKD itself can further exacerbate hypertension. In England, nearly 7.2 million people are living with CKD, representing approximately 10% of the population⁵⁷.

Kidney dysfunction (**Figure 1.3**) is commonly diagnosed using markers such as creatinine, cystatin-C, proteins, and albumin, all of which have been shown to be elevated in individuals with hypertension^{58,59}. Additionally, estimated glomerular

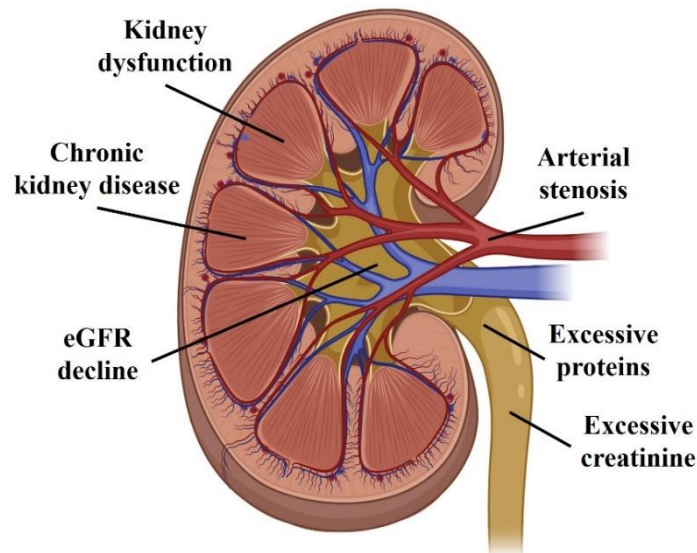


Figure 1.3 | Hypertension-induced kidney damage. Renal alterations include elevated proteins and low filtration capacity, increase creatinine, low estimated glomerular filtration rate (eGFR), arterial narrowing, leading to chronic kidney disease (CKD) and renal dysfunction.

filtration rate (eGFR) is a key diagnostic marker for CKD, with values below 60 mL/min/1.73 m² indicating impaired kidney function⁶⁰. On the imaging side, kidney screening is commonly performed using ultrasound or MR imaging to visualise renal blood vessels and assess for narrowing, particularly in cases of suspected renal artery stenosis, a major cause of renal-associated hypertension⁶¹.

1.2.5 Other vasculature alterations

There is evidence that hypertension causes damage to the retinal microvasculature, often resulting in hypertensive retinopathy⁶². Fundus imaging, particularly optical coherence tomography angiography (OCTA), is commonly used to assess the retinal arteriovenous structure⁶³. OCTA enables detailed quantitative assessment of vessel density and perfusion as well as visualisation of the microvasculature⁶⁴. In hypertensive retinopathy, patients frequently exhibit retinal haemorrhages, arteriovenous nicking, arteriolar narrowing, and cotton wool spots⁶⁵. The severity of retinal vessel damage caused by hypertension varies and correlates with both the duration and the level of elevated blood

pressure, with changes typically classified into four grades of increasing severity⁶⁶. These grades are determined based on the presence of retinal haemorrhages and exudates, optic disc swelling, and the degree of vascular narrowing⁶⁷. Early stages are characterised by mild narrowing of the retinal arterioles, whereas severe stages are associated with pronounced swelling and vision loss. Moreover, the arteriovenous ratio (AVR) is commonly used to assess retinal vasoconstriction and is often markedly reduced below 0.66 as blood pressure progresses, indicating advanced arteriolar narrowing⁶⁸.

Hypertension also contributes to damage across the systemic vasculature, including significant effects on the carotid arteries. Although routine carotid ultrasound (CUS) is no longer recommended by the ESH guidelines¹², the detection of carotid plaque remains a strong marker of atherosclerosis⁶⁹. Elevated blood pressure reduces the elasticity of large arteries and the aorta, leading to increased left ventricular (LV) afterload and higher systolic pressure⁷⁰. Therefore, assessing the presence of carotid atherosclerosis remains crucial for long-term hypertension management. In CUS, the intima-media thickness (IMT) and the presence, size, and characteristics of plaques are commonly evaluated, providing valuable insights into vascular health and cardiovascular risk⁷¹. IMT is measured by B-mode ultrasound as the distance between the intima-lumen and media-adventitia interfaces of the left, right, or both carotid arteries, using different imaging angles and measurement approaches, *i.e.*, mean, maximum, or minimum⁷². Values greater than 900 μm are considered indicative of asymptomatic organ damage in the carotid arteries, with risk increasing further in the presence of elevated blood pressure⁷³.

1.3 The use of artificial intelligence in hypertension assessment

Given the complex nature of hypertension and its systemic impact on multiple organs, traditional methods remain limited in assessing the full spectrum of disease progression. Moreover, in clinical practice, comprehensive evaluation of all affected organs is often impractical due to the increased burden on clinicians, the need for multiple equipment, and the large volume of data requiring analysis. Emerging artificial intelligence (AI) offers a promising solution to overcome these challenges. In this thesis, the term AI is used broadly to encompass simple computational algorithms, *e.g.*, dimensionality reduction, as well as machine learning and advanced deep learning methods.

1.3.1 Definition and technical background

Interest in AI dates back to the 1950s, but over the past decade, it has rapidly expanded across various domains, particularly healthcare, driven by significant advancements in cyber-physical systems, such as surgical robots, and virtual systems, such as machine learning tools⁷⁴. AI can be defined as the ability of machines to acquire intelligence that enables them to perform specific tasks autonomously⁷⁴. Its strong impact in healthcare lies in its potential to serve as a clinical decision-support tool, capable of identifying hidden patterns and complex relationships within large patient datasets⁷⁵. Although AI remains in the early stages of development, its techniques and algorithms have advanced considerably. Current AI algorithms can employ supervised, unsupervised, or semi-supervised approaches during the learning process. Moreover, AI includes machine

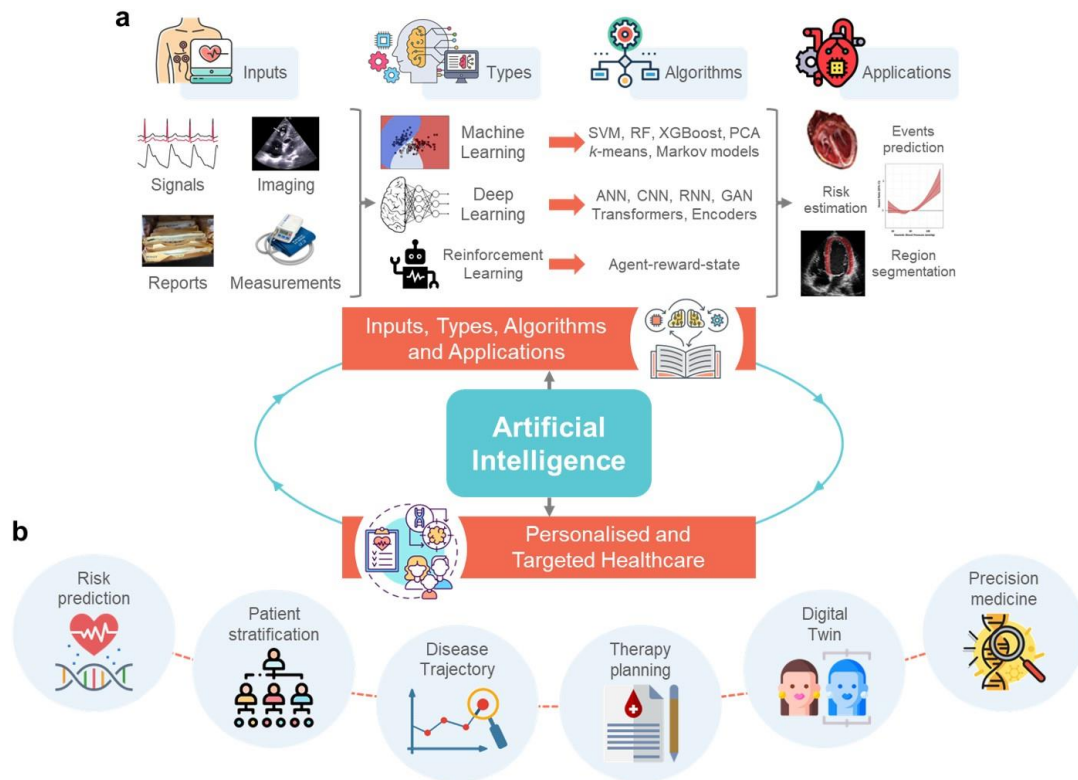


Figure 1.4 | Types and applications of artificial intelligence (AI) in personalised healthcare. a, Various data inputs can be used for AI algorithms, including biosignals, medical imaging, omics, patient electronic health reports, and other clinical measures. AI is divided into machine learning, deep learning, and reinforcement learning which can be applied to multiple applications for classification, regression, or segmentation. **b,** AI can enhance personalised healthcare for predicting risks, stratification of patients based on disease level, construction of disease trajectories, planning for therapy, advancing digital twins, and precision medicine.

learning, within which deep learning is a specialised subset, each designed to perform specific tasks (**Figure 1.4**).

Machine learning is a branch of AI focused on enabling machines to learn from experience, with or without predefined programming rules⁷⁶. These rules determine whether the learning process follows supervised or unsupervised mechanisms⁷⁷. In supervised learning, the model learns from labelled data, where each set of training features is associated with a corresponding label, such as healthy or diseased. In contrast, unsupervised learning involves deriving patterns and labels from unlabelled data through methods such as cluster analysis or dimensionality reduction⁷⁸. Semi-supervised learning combines aspects of both approaches and is particularly valuable in applications

where labelling large datasets is impractical, *e.g.*, long ECG recordings⁷⁹. The core framework of these algorithms involves processing input data, *e.g.*, medical images, by extracting relevant features that are fed into mathematical models and optimisers⁸⁰. These algorithms utilise high-dimensional data to make predictions and infer target labels. Common supervised algorithms include support vector machines (SVM), decision trees, random forests (RF), gradient boosting (GBoost), and logistic regression⁸¹. Combining these models into ensemble classifiers is a widely adopted strategy to enhance performance. Additionally, unsupervised techniques such as principal component analysis (PCA), *k*-means clustering, and hidden Markov models are employed for labelling unlabelled data and reducing the dimensionality of large datasets⁸².

Deep learning is a specialised subset of machine learning that mimics the way humans perceive information and build knowledge hierarchically⁸³. Unlike traditional machine learning, deep learning is considered more advanced because it often eliminates the need for manual feature engineering or extensive domain expertise when performing tasks. Instead, it processes raw input data directly, extracting deeper representations through multiple layers of neural networks⁸⁴. These networks are inspired by the structure of brain neurons, simulating how humans gradually recognise the characteristics of an object or problem. The architecture of a neural network is composed of multiple layers, with its depth referring to the number of these layers⁸⁵. Greater depth allows for increased non-linearity in transforming raw input data, producing complex feature representations. For instance, when working with imaging data, each layer applies mathematical operations, *e.g.*, convolutions, on the features extracted from the previous layer to detect more detailed patterns, including sharper edges or subtle colour gradients, as implemented in convolutional neural networks (CNNs)⁸⁶.

1.3.2 Applications in hypertension assessment

AI algorithms have been widely applied in the assessment of elevated blood pressure across various applications. One major area of application is improving the accuracy of blood pressure measurements, particularly in automatic oscillometric BP monitors⁸⁷. Research in this area has focused on cuff-based approaches, such as the evaluation of auscultatory waveforms, to improve measurement precision⁸⁸. In parallel, cuff-less methods have been investigated using photoplethysmograms (PPG) and electrocardiograms (ECG), with estimations based on pulse transit time (PTT) and pulse arrival time (PAT)^{89,90}. Additionally, imaging-based approaches have been developed, in which AI algorithms are trained to estimate blood pressure levels using features extracted from retinal imaging⁹¹ or ultrasound⁹².

In diagnosis applications, the primary objective is to identify modifiable risk factors to enable early prediction of hypertension⁹³. To achieve this, AI models commonly integrate a wide range of variables, including health, behavioural, socioeconomic, physical activity, and environmental factors, along with comorbid conditions such as diabetes and cholesterol disorders. Various machine learning techniques, such as decision trees⁹⁴, GBoost⁹⁵, and deep neural networks⁹⁶, have demonstrated effectiveness in predicting hypertension. Another application area focuses on combining genetics with risk factors to discover novel genes associated with hypertension and thus, enhance the capability of AI to predict the disease^{97,98}.

Recent studies have also extended AI applications beyond diagnosis, focusing on estimating the risk of future hypertension-related adverse outcomes⁹⁹. Unlike traditional risk stratification methods, machine learning leverages large datasets to provide more accurate prognostic measures¹⁰⁰. Studies have utilised machine learning models trained

on electronic health records, metabolic profiles, blood biomarkers, and other risk factors to estimate the 10-year risk of cardiovascular events^{101,102}. Moreover, these models have demonstrated the capability to identify the most influential features contributing to the prediction of future events. In the management of hypertension, there is growing interest in exploring the value of AI for guiding interventions, including the optimisation of medication plans¹⁰³. Machine learning models have demonstrated the ability to predict patient responses to specific antihypertensive medications and to identify individuals who are more or less likely to benefit from treatment strategies¹⁰⁴. For example, research have shown that the use of beta-blockers, proton pump inhibitors, and statins, when guided by AI predictions, can improve the success rate of hypertension treatment¹⁰⁵.

In investigating the damage caused by hypertension, only a limited number of studies have examined its impact on vital organs¹⁰⁶. These studies often focus on a single organ and rely on relatively small datasets. The applications of AI and medical imaging in the assessment of hypertension-induced organ alterations will be thoroughly discussed in the systematic review and meta-analysis chapter of this thesis (**Chapter 2**). In this thesis, I focus on developing a machine learning approach, the contrastive trajectory inference (cTI), to characterise organ damage resulting from elevated blood pressure using both imaging and non-imaging markers.

1.3.3 Contrastive trajectory inference (cTI)

The cTI method is an unsupervised machine learning algorithm based on the concepts of contrastive learning and trajectory exploration¹⁰⁷. It is often applied to disease staging and phenotyping, particularly with large, multi-dimensional datasets such as genetic data, to study pseudo-temporal patterns¹⁰⁸. In this approach, the method defines two contrastive populations, *e.g.*, background and target, and identifies patterns in the target

group (diseased) in contrast to patterns in the background group (control). To achieve this, the method is built on two main components: the contrastive PCA (cPCA) for dimensionality reduction and the minimum spanning tree (MST) for trajectory inference.

Similar to traditional PCA, the cPCA linearly transforms data into uncorrelated orthogonal principal components that capture maximal variance (**Figure 1.5**). Principal components are new variables formed as linear combinations of the original variables, ensuring they are mutually uncorrelated¹⁰⁹. The process begins by computing the covariance matrix to assess how input variables vary with the mean of each other. Eigen decomposition is then applied to obtain eigenvectors (directions of maximum variance) and eigenvalues (amount of variance explained). Principal components are ranked by their eigenvalues, with the first component typically explaining the greatest variance and being the most informative. However, because PCA constructs a joint covariance matrix from all data, variance associated with background individuals can overlap with that of the target group. To address this, cPCA introduces a contrastive step that downweights the background covariance, emphasising variance unique to the target dataset. Here, cPCA performs a weighted subtraction between the two covariance matrices of the two extremities (background and target) to form the contrastive covariance matrix (cov_c) as,

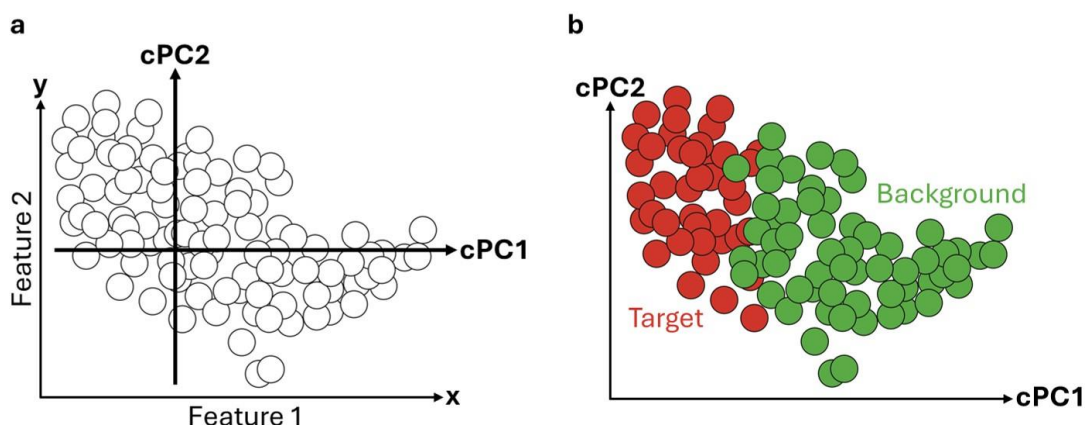


Figure 1.5 | Contrastive principal component analysis (cPCA). **a**, Feature covariance matrices are subtracted between the background (green) and target (red) datasets, then passed to cPCA for dimensionality reduction. **b**, The reduced cPCA space enables better discrimination between the contrastive groups, with contrastive principal components (cPCs) ordered by importance.

$$cov_c = cov_t - \alpha * cov_b$$

where cov_t and cov_b are the covariance matrices of the target and background datasets, respectively, and α is the contrast parameter that quantifies the trade-off between having high target and low background variance, and thus, target variance dominates.

Accordingly, importance of each variable (C_i) to the generation of principal components can be calculated using the eigenvectors (ω_i^2) as follows,

$$C_i = 100 \times \left(\frac{\omega_i^2}{\sum_{k=1}^K \omega_k^2} \right) \times \lambda$$

where i is the selected variable, $\sum_{k=1}^K \omega_k^2$ is the sum of all eigenvectors at every component (k), K is the total number of principal components, and λ is the eigenvalues.

In addition to dimensionality reduction using cPCA, cTI allows for connecting data of all subjects with the aim of forming optimal connections using MST¹¹⁰ (**Figure 1.6**). Here, all subjects are represented as nodes and connected in an un-directed graph such that the summed length of all paths from the root node is minimised. By tracing the pathways from the root node to each endpoint, a pseudo-temporal distance is computed relative to the maximum path length, the farthest point representing the most advanced

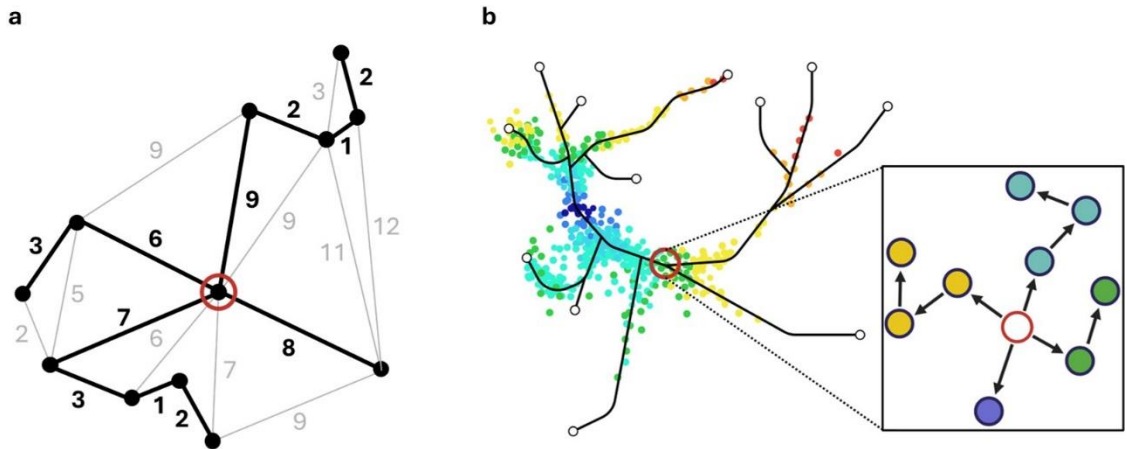


Figure 1.6 | Contrastive trajectory inference (cTI) with minimum spanning tree (MST). **a**, A Euclidean distance matrix is constructed, and all individuals are connected using an MST to minimize the total length of all paths from the root node (red circle). **b**, Tracking these connections from the root node to different endpoints allows the formation of dominant pathways, *i.e.*, trajectories.

disease state. This provides a normalised quantitative measure of disease state, where 0 indicates proximity from the pathology-free state (root) and 1 indicates the most advanced disease state (farthest node).

1.4 Thesis aims and objectives

Currently, there are limited studies investigating the impact of elevated blood pressure on multiple organs simultaneously, despite hypertension being a systemic condition that affects more than a single organ. There is a critical need for comprehensive assessments of hypertension-induced end-organ damage across multiple systems to enhance our understanding of its underlying mechanisms and improve targeted interventions. Therefore, this thesis aims to address this gap by conducting a holistic analysis that integrates multi-organ, multi-modal, large-scale datasets of hypertension-related organ changes, with the goal of discovering new subclinical organ-specific disease phenotypes and producing an AI model capable of assessing the organ damage.

The thesis is structured around four main objectives. The first objective is to review and synthesise current AI applications in characterising hypertension and end-organ damage using medical imaging, addressing the need to understand how imaging and AI are integrated to characterise hypertension. To achieve this, I conducted a comprehensive systematic review to capture all relevant studies within this scope and identified the major AI tasks applied to predict, phenotype, and manage hypertension. The systematic review provides an overview of the state-of-the-art, highlights current gaps, and outlines potential areas for improvement, potentially providing guidance for the current research.

The second objective is to develop a comprehensive AI-assisted model that captures the contrastive spectrum of organ damage between normotensive and hypertensive states,

thereby addressing the existing gap in understanding the multi-organ effects of hypertension. To achieve this, I developed a semi-supervised machine learning model using cTI and trained it on a large cohort derived from the multi-modal UK Biobank dataset. The primary aim was to model global end-organ damage severity and quantify it as an imaging risk score while identifying predominant organ-specific trajectories of damage progression, each spanning from normality to advanced disease states. In this thesis the AI model is referred to as the *HyTwin*.

The third objective is to evaluate the generalisability of the developed *HyTwin* model by testing it on an out-of-domain dataset, thereby addressing the limitation of existing models that often fail to perform across broader and more diverse populations. To achieve this, I got access to a large external dataset, the Atherosclerosis Risk in Communities Study (ARIC) cohort from the United States. Individuals from this dataset were positioned along one or more of the developed disease trajectories, and their organ damage severity was quantified accordingly.

Finally, the fourth objective is to explore the clinical reliability of the developed model by applying it to a closely related domain, the hypertensive disorders of pregnancy (HDP). Here, I aimed to leverage the model of hypertension-induced organ damage to identify novel pregnancy-related phenotypes in relation to blood pressure elevation. To achieve this, I got access to *HyPrevent*, a locally recorded database in lab settings, and obtained an HDP-based data from a clinical trial. These individuals were positioned within the developed disease trajectories, enabling the discovery of unique lifestyle-, demographic-, and pregnancy-related phenotypes based on their mid-life organ damage profiles.

2. SYSTEMATIC REVIEW AND META-ANALYSIS

2.1 Abstract

Background. Current hypertension guidelines focus on blood pressure control, but incorporating end-organ imaging could improve understanding of disease manifestations.

Objectives. In this chapter, I aimed to conduct a systematic review and meta-analysis to evaluate current task-level uses and efficacy of AI and computational approaches in imaging for hypertension identification, phenotyping, and outcome prediction.

Methods. A systematic search was conducted across multiple databases up to May 2025. Retrieved studies were grouped by AI task, and a thematic qualitative analysis per-task was conducted to evaluate organ-specific findings, AI methodologies, and research gaps. For quantitative synthesis, the I^2 statistic derived from Cochran's Q test was used to assess heterogeneity, and forest plots were generated to visualise effect sizes. The review was registered with PROSPERO (CRD42023427430).

Results. The search strategy yielded 45 studies. Thematic analysis categorised the studies into five major tasks, with the majority employing supervised learning for classification processes. Nearly half of the studies focused on the heart. However, paucity of studies performed multi-organ assessment, external validation, and phenotyping or predicting future risk. AI using imaging achieved an overall sensitivity of 0.84 [0.69—0.93] in identifying hypertension from normotension, highest with brain imaging and lowest with retinal imaging. It also reached 0.92 [0.89—0.94] in discriminating hypertension from hypertrophic cardiomyopathy.

Conclusion. Current research focusses primarily on optimising hypertension prediction. While results are promising, datasets remain small with limited external validation.

There remains a need for discovery-oriented research to uncover disease heterogeneity, organ-level phenotypes, and support personalised and targeted interventions.

Publication Status.

The results presented in this chapter are currently under-review.

1. **Alkhodari M**, Sattwika P, Cutler H, Milner G, Kart T, Hadjileontiadis L, Khandoker A, Banerjee A, Leeson P, Lapidaire W. *Artificial intelligence applied to imaging for prediction, phenotyping, and management of hypertension: a systematic review.*

2.2 Introduction

Current diagnostic and therapeutic approaches largely emphasise monitoring and pharmacologically managing blood pressure. However, this focus may overlook early-stage hypertensive end-organ damage, particularly in patients without symptoms^{111,112}. Prolonged exposure to hypertension can cause severe organ damage, markedly elevating the risk of cardiovascular, cerebrovascular, renal, and other organ-related diseases¹¹³. Despite its prevalence, hypertensive end-organ disease remains under-diagnosed¹¹⁴.

Common imaging modalities such as MR imaging, computed tomography (CT), and echocardiography are capable of detecting structural and functional changes in organs. Incorporating multi-modal medical imaging into hypertension care could provide a more thorough understanding of the disease's organ-level effects and progression¹¹⁵. Nevertheless, the use of medical imaging in routine hypertension management remains limited due to factors such as high costs, restricted accessibility, reliance on specialist interpretation, and the predominant clinical emphasis on blood biochemistry¹¹⁶. Consequently, imaging is typically employed to identify secondary causes and to detect organ damage only at more advanced stages.

AI holds significant promise for streamlining image interpretation workflows while improving diagnostic accuracy¹¹⁷. Recent developments in AI have substantially advanced image analysis capabilities, enabling the extraction of complex and often subtle features from organ-specific images^{118,119}. These advancements have the potential to deliver automated, precise, and cost-effective diagnostic and prognostic support for clinicians¹²⁰. In this chapter, a systematic review was conducted to establish the current state-of-the-art for use, effectiveness and methodologies underlying AI and

computational tools applied to imaging to identify, phenotype, and predict hypertension and its outcomes using organ information.

2.3 Methods

2.3.1 Study design

This review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement¹²¹. The review protocol was registered with PROSPERO ID: CRD42023427430 in June 2023. As this review did not involve the recruitment of participants or animals, approval from an institutional review board or ethics committee was not required.

2.3.2 Search strategy and eligibility criteria

The search strategies were developed using multiple electronic databases by combining different key terms (**Table 2.1**) to form search queries. MEDLINE, EMBASE (via Ovid), Scopus, PubMed, IEEE Xplore, CINAHL, and Web of Science were searched from inception to May 2025. Searches were conducted by myself without language restrictions. Reference lists of retrieved articles were manually screened to ensure study eligibility.

Eligibility criteria required studies to: 1) be diagnostic or prognostic in design, 2) involve human adults (>18 years), 3) report on organ structure or function, and 4) apply AI for identifying, phenotyping, or predicting hypertension and its outcomes. Exclusion criteria included: 1) conference abstracts, reviews (narrative or systematic), meta-analyses, editorials, letters, comments, and retracted papers, and 2) research focused on pulmonary

Table 2.1 | Key terms used for the search strategy across different databases.

Search combinations	Search queries
#1	exp hypertension/ OR "high blood pressure".mp OR hypertensi*.tw,kf. OR ((high* OR rais* OR elevat* OR increas*) ADJ3 (systolic OR diastolic OR arterial OR blood) ADJ3 pressure*).tw,kf. NOT pulmonary
#2	exp "artificial intelligence"/ or (("artificial intelligence" or "machine learning" or "deep learning" or "supervised" or "unsupervised" or "reinforcement" or "neural network*" or "naive bayes" or "decision tree" or "random forest" or "support vector machine" or "k-nearest neighbour" or "linear discriminant analysis") adj3 (classifier or regression or classification or clustering)).tw,kf.
#3	exp "computed tomography"/ or exp "magnetic resonance imaging"/ or exp ultrasonography/ or exp "pulse wave analysis"/ or ("CT imaging" or CT or X-ray or "magnetic resonance" or "MR imaging" or MRI or "cardiac MR*" or "cardiovascular MR*" or "diffusion tensor imaging" or "diffusion weight* imaging" or "diffusion imaging" or echocardiograph* or cardioechograph* or echocardiogram or doppler or "speckle tracking" or "pulse* wave velocity" or "cardiac structure*" or "myocardial structure*" or "cardiac function*" or "myocardial function*" or "intima media thickness" or "flow mediated dilatation" or microvascula* or microvessel* or macrovascula* or macrovessel*).tw,kf. or (ultraso* adj3 (brain or heart or cardj* or myocard* or ventric* or kidney* or neph* or renal or vascul* or vessel*)).tw,kf.
#4	#1 AND #2 AND #3

hypertension or hypertensive disorders of pregnancy. No restrictions were placed on imaging modality or organ studied.

2.3.3 Population, intervention, and outcomes

The population included adults diagnosed with elevated blood pressure or hypertension. Interventions involved AI and computational tools, including supervised or unsupervised machine learning, applied to population data for metric extraction, risk prediction, diagnosis, subgroup phenotyping, characterising the disease, or predicting its adverse outcomes. Tools ranged from simple computational models to more advanced deep learning algorithms.

The primary outcome was AI-extracted image features describing structural or functional changes across hypertension stages. Secondary outcomes included AI and computational applications in medical imaging for assessing disease progression, predicting outcomes, and evaluating automated diagnostic performance, either by pre-extracted features (imaging biomarkers) or features extracted directly by AI and computational algorithms.

2.3.4 Selection of studies

All records identified through the database search were first deduplicated using the Systematic Review Accelerator¹²², followed by manual removal of duplicates based on titles and abstracts. The remaining abstracts were screened by three investigators (myself and two other colleagues) using Rayyan software¹²³. Full-text eligibility was assessed to apply inclusion and exclusion criteria. Selection conflicts were reviewed by a fourth investigator (supervisor) and resolved through discussion among all reviewers.

2.3.5 Data extraction

A standardised extraction form was used to collect data from all articles, with extraction performed by a single reviewer (myself) and independently verified by two additional reviewers (two other colleagues) to ensure accuracy and completeness. The extracted data included first author, publication year, country of study conduction, study design, dataset size, age range, female percentage, organ type, imaging modality, hypertension diagnosis, pre-processing steps, type of AI, task of AI, performance metrics, outcome measures, and external validation results. For articles with missing details or ambiguous information, the authors were directly contacted for further clarification.

2.3.6 Risk of bias assessment

The risk of bias assessment was performed to systematically evaluate the reliability and applicability of the methodological quality and potential of studies in the context of AI-driven medical imaging for hypertension. This allows for weighing the studies based on their influence in the interpretation and synthesis of provided evidence¹²⁴. The quality of the included studies was assessed based on two tools: the Quality Assessment of

Diagnostic Accuracy Studies Artificial Intelligence (QUADAS-AI)¹²⁵ for diagnostic studies and the Prediction model Risk of Bias Assessment Tool (PROBAST)¹²⁶ for prognostic studies. Three reviewers (myself and two other colleagues) independently evaluated quality of the articles based on risk of bias, applicability concerns, and overall representation according to the two assessment tools. If there were conflicts between the reviewers, *e.g.*, conflict on bias score, a fourth independent reviewer (supervisor) resolved them after discussion by selecting the best assessment.

The QUADAS-AI criteria included: 1) patient selection, 2) interpretation of index tests, 3) reference standard, and 4) flow/timing, while PROBAST assessed: 1) inclusion/exclusion of participants, 2) predictor selection, 3) outcome standard, and 4) model performance analysis. Each criterion was assigned a score of -1 (high risk), 0 (unclear risk), or 1 (low risk). For the overall evaluation of bias and applicability, if any criterion was rated as high risk, *i.e.*, -1, the study was classified as high risk overall. If only unclear risk, *i.e.*, 0, was assigned alongside low risk, *i.e.*, 1, the study was categorised as unclear risk overall. A study was considered low risk overall only if all criteria were individually rated as low risk (1).

2.3.7 Thematic data synthesis

All identified studies were grouped into categories based on their core objective in applying AI and computational algorithms to medical imaging for hypertension research. I examined the primary tasks performed by AI, defining a task as a separate category if addressed by at least three studies to ensure statistical validity. A thematic analysis was then conducted by reviewing each study's objectives, methods, and findings to identify and group similar AI tasks into thematic categories. This synthesis allowed comparison across studies despite differences in design, imaging modalities, or AI techniques. The

final grouping was reached through consensus among three reviewers (myself and two colleagues), with disagreements resolved through discussions with independent reviewers (supervisors). Analyses were performed separately for each task category.

2.3.8 Meta-analysis

I conducted meta-analysis on studies grouped by thematic categories, primarily focusing on: 1) differentiating hypertension from normotension and 2) distinguishing hypertension from hypertrophic cardiomyopathy (HCM). I evaluated these two categories based on the availability of a sufficient number of studies reporting performance and addressing a common objective. The meta-analysis was conducted using a hierarchical random-effects model to evaluate sensitivity across studies while accounting for variability. Moreover, sensitivity forest plots were analysed to evaluate heterogeneity between studies based on the I^2 statistic from Cochran's Q test¹²⁷, considering heterogeneity as low (25%-49%), moderate (50%-74%), and high (>75%). Statistical significance was tested using the Cochran's Q test and defined as a p value < 0.05 indicating significant heterogeneity. Sensitivity forest plots were generated in R using the "*Meta*" package (version 8.0-1). For the sensitivity calculations, True Positives (TP) were treated as the "Events" and the "Total" was defined as the sum of TP and False Negatives (FN).

Furthermore, to jointly synthesise sensitivity and specificity, we applied a bivariate random-effects model (Reitsma model). A summary receiver operating characteristic (SROC) curve¹²⁸ with a 95% confidence interval was generated, and the area under the curve (AUC) was calculated as an overall performance measure. These analyses were performed using the "*Mada*" package in R. Lastly, funnel plots were examined to evaluate performance symmetry.

2.4 Results

2.4.1 Search results

I identified 1,612 records in the initial database search. After the removal of duplicates, a total of 1,047 records were screened for suitability by evaluating the title and abstract relative to the selection criteria. Here, 965 records were excluded, yielding 82 articles sought for retrieval, all with full-text reports available. Their eligibility was further assessed through deeper analysis of the full-texts, which resulted in excluding 37 reports due to them not being focused on hypertension (n = 6), having non-imaging datasets (n = 11), using non-AI techniques (n = 4), or being either conference abstracts (n = 8), reviews/case reports (n = 6), and retracted papers (n = 2). Finally, a total of 45 articles were selected for further analysis (**Figure 2.1**).

2.4.2 General characteristics of studies

A detailed summary of the characteristics of the 45 included studies is presented in **Figure 2.2**. Most of the selected studies were conducted after 2017 (n = 38; 84.44%) (**Figure 2.2a**). The majority of selected populations were aged above 45 years old (n = 25; 55.56%) (**Figure 2.2b**). Alsharqi *et al.*¹²⁹ conducted a study on young adults (age = 20-35) to generate clinically meaningful scores and phenotypes of cardiac remodelling, whereas Recenti *et al.*¹³⁰ analysed an elderly cohort (age = 68-98) to identify hypertension using muscle radiodensitometry. A total of 12 studies (26.67%) had approximately equal distribution of male and female subjects (**Figure 2.2c**), with only one study, by Pessana *et al.*¹³¹, focused exclusively on male patients to investigate the impact of instantaneous arterial diameter changes on hypertension. Most of the included

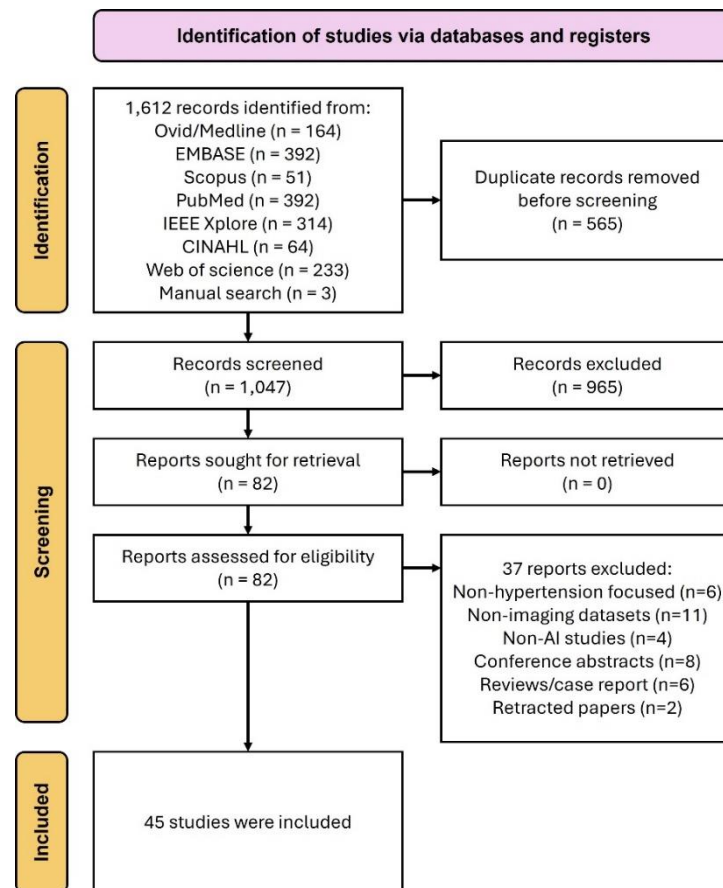


Figure 2.1 | Flow diagram of the study selection process. The process goes through identification of records through databases search followed by title/abstract and full-text screening relative to the inclusion and exclusion criteria. A total of 45 studies out of 1,612 were included in this review for further analysis. Medline: Medical Literature Analysis and Retrieval System Online; EMBASE: Excerpta Medica Database; CINAHL: The Cumulative Index to Nursing and Allied Health Literature.

studies were conducted by corresponding/first authors from China ($n = 11$; 24.44%) and US ($n = 8$; 17.78%) (**Figure 2.2d**).

Across all studies, 22 studied cardiac structure and function (48.89%), followed by 12 focusing on the retina (26.67%) and 6 on the brain (13.33%) (**Figure 2.2e**). Additionally, individual studies each explored carotid vasculature¹³¹, tongue¹³², or conducted a joint analysis of multiple organs¹³³. The imaging modalities varied, with echocardiography being the most common ($n = 14$; 31.11%), followed by retinal imaging ($n = 12$; 26.67%) (**Figure 2.2f**). Additionally, individual studies each employed unique modalities such as tongue imaging¹³², femoral CT¹³⁰, or multi-modality data¹³³. Most studies ($n = 31$; 68.89%) utilised small datasets of no more than 1,000 participants (**Figure 2.2g**). A total

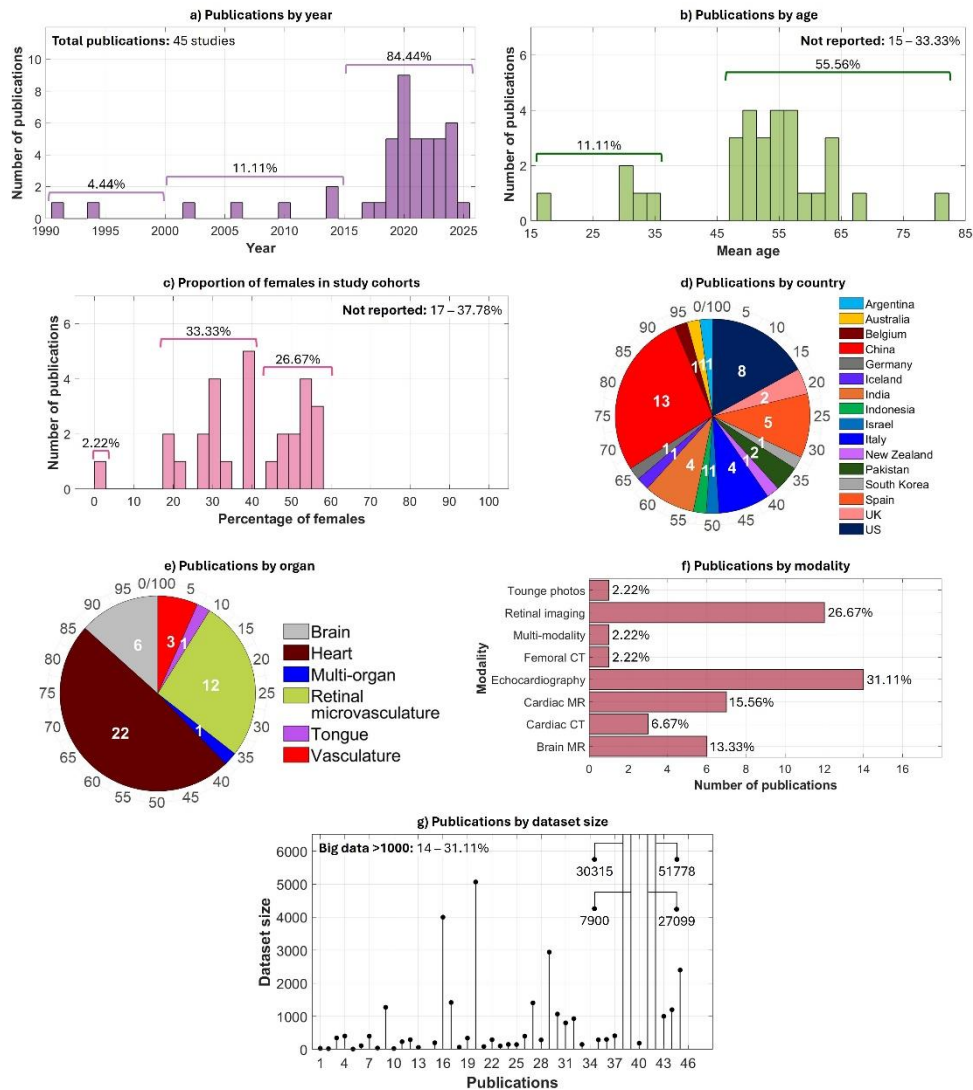


Figure 2.2 | Overall summary of the population included in the selected studies and type of used data. **a)** Publications by year revealed that the range of studies covered was between 1991 and 2025. **b)** Summary of publications by age was performed for studies with available data. **c)** Sex-based analysis was performed as well revealing few studies considering balanced number of females and males. **d)** 16 countries have led the research in these studies by correspondence/first authorship. Publications were further divided by: **e)** targeted organ for investigation and **f)** type of modality. **g)** Various datasets were used and varied between being small or big data. Big data is referred to as having more than 1,000 subjects or samples. CT: computed tomography; MR: magnetic resonance.

of 14 studies (31.11%) incorporated larger cohorts, including data from the UK Biobank^{133,134}, FLEMENGHO¹³⁵, AGES-I/AGES-II¹³⁰, BioHit/Tongji¹³², and local data^{129,136–138}.

In the risk of bias assessment (**Figure 2.3a**), 24 studies (53.33%) were classified as having low risk, 3 studies (6.67%) as unclear risk, and 18 studies (40.00%) as high risk

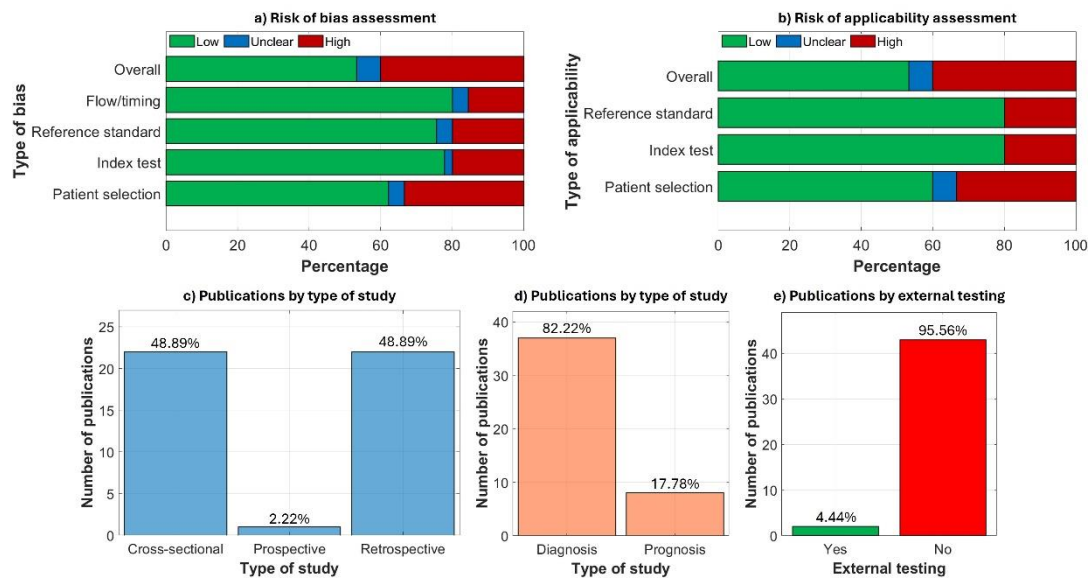


Figure 2.3 | Overall characterisation of included studies based on bias assessment and study type. Summary of studies was based on: the assessment of the risk of **a)** bias and **b)** applicability performed by taking into consideration flow/timing, reference standard, index test, and patient selection. **c)** Type of study design being cross-sectional, prospective, or retrospective, **d)** type of study clinically being diagnostic or prognostic, and **e)** whether an external validation was performed or not.

of bias. Similarly, applicability assessment showed 24 studies (53.33%) with low risk (**Figure 2.3b**).

The overall summary of the type of included studies is presented in **Figure 2.3**. The majority of the included studies were either retrospective ($n = 22$; 48.89%) or cross-sectional ($n = 22$; 48.89%). Only Raghavendra *et al.*¹³⁹ conducted a prospective study (**Figure 2.3c**). Studies were mainly diagnostic ($n = 37$; 82.22%) with a small number of prognostic studies ($n = 8$; 17.78%) (**Figure 2.3e**). Only two studies performed external validation of their findings; Diao *et al.*¹⁴⁰ and Wang *et al.*¹⁴¹ distinguished hypertension from other diseases (**Figure 2.3e**).

The included studies utilised various AI models, primarily for classification tasks (**Figure 2.4a**), for example to distinguish hearts affected by hypertension from their normotensive counterparts using image-related data ($n = 26$; 57.78%). Image segmentation tasks were performed ($n = 7$; 15.56%), particularly in heart- and brain-related research. For example, Wang *et al.*¹⁴² identified adverse blood pressure status by

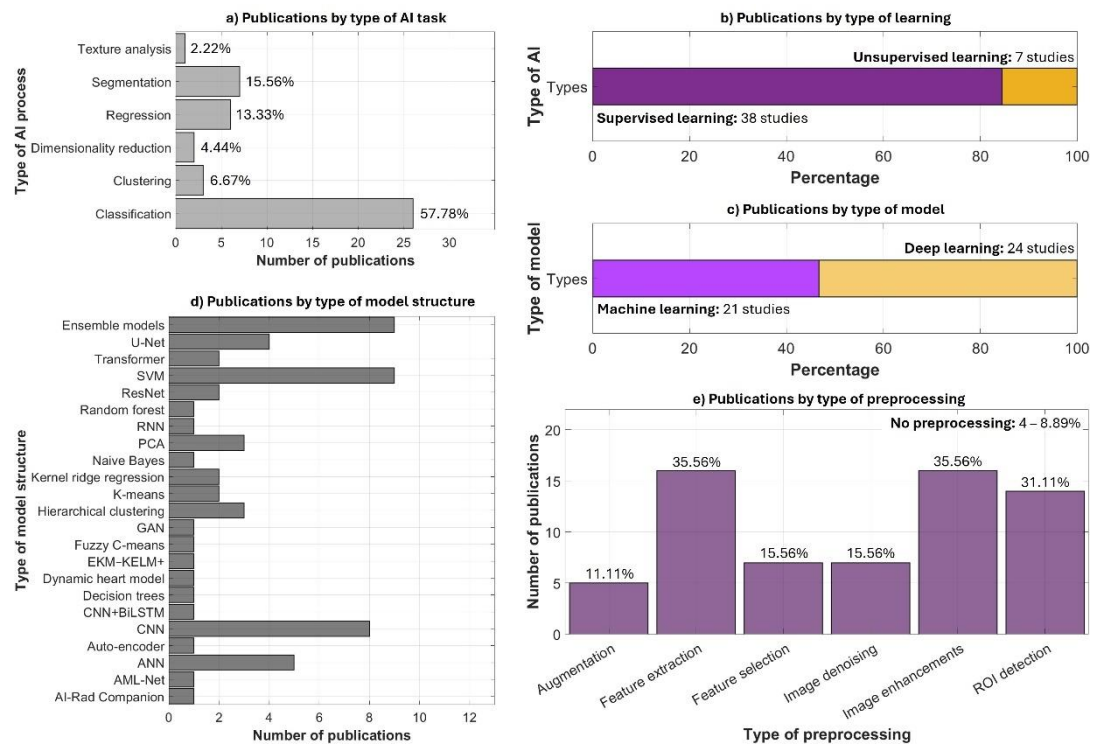


Figure 2.4 | Overall characterisation of included studies based on artificial intelligence approaches. A total of 45 studies were included in this systematic review. Summary of studies was based on: **a)** Studies were divided based on the task of the AI algorithm, **b)** type of learning, and **c)** type of modelling. **d)** A thorough summary of the structure used when building the models is provided alongside **e)** the type of preprocessing procedures performed across studies. Here, the number of studies may be repeated if more than two preprocessing steps were performed. SVM: support vector machine; ResNet: residual network; RNN: recurrent neural network; PCA: principal component analysis; GAN: generative adversarial network; EKM-KELM+: empirical kernel mapping-based kernel extreme learning machine plus; CNN-BiLSTM: convolutional neural network Bi-directional long short-term memory; ANN: artificial neural network; ROI: region-of-interest.

segmenting aortic diameter measurements and Xu *et al.*¹⁴³ analysed myocardial radiomics in left ventricular hypertrophy by segmenting echocardiograms. Other AI applications include regression and clustering. For example, Alkhodari *et al.*¹³³ estimated hypertension progression severity using multi-modality imaging features, while Yang *et al.*¹⁴⁴ identified hypertension subtypes and associated endophenotypic signatures. Additionally, some of the studies focused on dimensionality reduction^{145,146} and texture analysis¹³⁶.

Most studies focused on supervised learning ($n = 38$; 84.44%) as a mechanism to train models in identifying hypertension-related abnormalities (**Figure 2.4b**). Moreover, the studies were nearly evenly split between traditional machine learning-based ($n = 21$; 46.67%) or deep learning-based methods ($n = 24$, 53.33%) (**Figure 2.4c**). Among traditional machine learning-based studies (**Figure 2.4d**), ensemble learning using multiple models were the most common approach ($n = 9$) and support vector machines (SVM) for classification tasks ($n = 9$). In contrast, deep learning-based studies primarily used convolutional neural networks (CNN) to perform feature extraction ($n = 8$), artificial neural networks (ANN) for classification/regression tasks ($n = 5$), and U-Nets for image segmentation ($n = 4$). Regarding preprocessing steps applied to imaging data before training models (**Figure 2.4e**), the most conducted steps were feature extraction ($n = 16$) and image enhancements of contrast or brightness ($n = 16$), followed by region of interest detection ($n = 14$). Other preprocessing techniques included feature selection to reduce complexity while focusing on key details, image denoising to improve assessment quality, and data augmentation to increase size and enhance data variability.

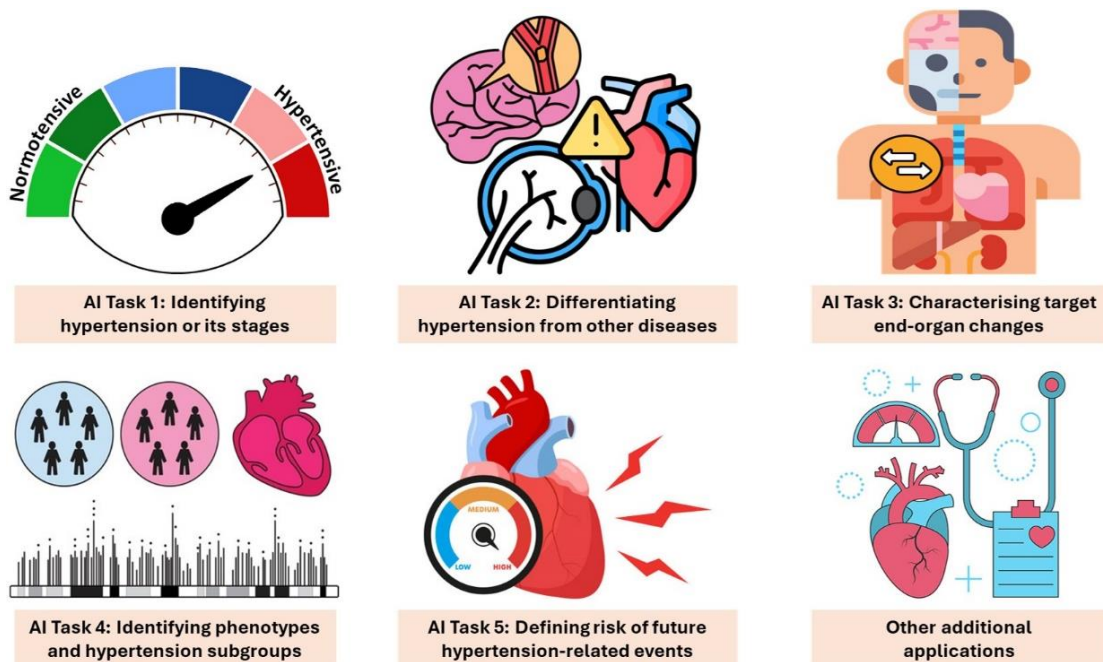


Figure 2.5 | Thematic tasks of included studies. Five major tasks were identified by studies included in this review to understand hypertension and its complications.

2.4.3 Thematic analysis

I categorised studies into five major tasks (**Figure 2.5**) as follows: AI Task 1) identifying hypertension or its stages - 10 studies (22%); AI Task 2) differentiating hypertension from other diseases/conditions - 17 studies (38%); AI Task 3) characterising target end-organ changes associated with hypertension - 11 studies (24%); AI Task 4) identifying subgroups and phenotypes within hypertensive cohorts - 3 studies (7%); and AI Task 5) assessing the risk of future hypertension-related outcomes - 3 studies (7%). A single study focusing on estimating blood pressure levels⁹² was categorised under other applications. A breakdown of all studies per-task is provided in **Table 2.2**.

Table 2.2 | Summary of studies based on major thematic tasks to characterise hypertension.

Study	Year	Target population	Type of modality	Outcome of interest
AI Task 1				
Identifying hypertension or its stages				
Retinal microvasculature				
Morales et al. ¹⁴⁵	2014	- Normal - Hypertensive retinopathy with at least one sign of pathological arteriovenous crossing	Retinal imaging	Discriminate between health fundus and hypertensive
Lo et al. ¹⁴⁷	2021	- Normal - Hypertensive retinopathy with Mitchell-Wong grading system standards		Identify mild hypertension using retinal images
Brain				
Yu et al. ¹⁴⁸	2019	- Hypertensives with four grades based on ACC 2017 guidelines	Brain MR	Discriminate between different hypertension grades by brain structural changes
Kandil et al. ¹⁴⁹	2019	- Normal - Pre-hypertensive with SBP: 120-140 and DBP: 80-90		Identify hypertension based on cerebral vascular changes
Kandil et al. ¹⁵⁰	2020	- Normal - Hypertensive with SBP: >130 and DBP: >80		Identify hypertension by tracking the cerebral vascular alterations
Kandil et al. ¹⁵¹	2020	- Normal - Hypertensive with SBP: >130 and DBP: >80		Identify hypertension using brain vascular features and SBP/DBP groups
Peng et al. ¹⁵²	2020	- Hypertensives with four grades based on ACC 2017 guidelines		Discriminate between different hypertension grades using brain structural changes
Heart				
Cetin et al. ¹⁵³	2019	- Normal - Hypertensive with no cardiovascular diseases	Cardiac MR	Identify hypertension using image radiomics
Wang et al. ¹⁴²	2022	- Hypertensives with four grades based on ESC/ESH 2018 guidelines	Cardiac CT	Identify hypertension using aortic diameter measurements
Tongue				
Karki et al. ¹³²	2020	- Normal - Hypertensive with liver, kidney, and spleen problems	Tongue images	Identify hypertension using tongue images
AI Task 2				
Differentiating hypertension from other diseases				
Retinal microvasculature				
Rose et al. ¹⁵⁴	2023	- Hypertensive retinopathy patients - Patients with heart attack	Retinal imaging	Discriminate between hypertensive retinopathy and heart attack
Koduri et al. ¹⁵⁵	2024	- Hypertensive retinopathy patients		Discriminate between hypertensive retinopathy and diabetic retinopathy

Heart				
Azhari <i>et al.</i> ¹⁵⁶	1991	- HT with LVH and/or WMA - Myocardial infarction - LV aneurysms	Cardiac CT	Discriminate hypertension from different pathological conditions
Neisius <i>et al.</i> ¹⁵⁷	2019	- HT with increased LVWT (>12 mm) - Hypertrophic cardiomyopathy	Cardiac MR	Discriminate between hypertensive heart disease and hypertrophic cardiomyopathy
Cetin <i>et al.</i> ¹³⁴	2020	- Normal - Hypertensives defined by self-reporting conditions	Cardiac MR	Identify hypertension from changes in cardiac structure and tissue texture
Vidal-Sospedra <i>et al.</i> ¹⁵⁸	2020	- Hypertensive heart disease - Hypertrophic cardiomyopathy - Amyloidosis	Cardiac MR	Differentiate between HCM, HIP, and AM using image texture analysis
Yu <i>et al.</i> ¹⁵⁹	2020	- Hypertensive heart with LVH - Hypertrophic cardiomyopathy (LVWT ≥15 mm) - Uremic cardiomyopathy with chronic ESRD	Echocardiography	Discriminate between hypertensive heart disease and other heart diseases
Sabovčik <i>et al.</i> ¹³⁵	2021	- Hypertensive with SBP: >140, DBP: >90 and using antihypertensive drugs - Patients with LVDD and LVH	Echocardiography	Discriminate between LVH and LVDD based echocardiographic ground truth information
Shi <i>et al.</i> ¹⁶⁰	2021	- HT with increased LVWT (>12 mm) - Hypertrophic cardiomyopathy (LVWT ≥15 mm)	Cardiac MR	Discriminate between HCM and HHD using image texture analysis
Barbieri <i>et al.</i> ¹³⁶	2022	- Arterial hypertension with diabetes or other comorbidities	Echocardiography	Determine changes in left atrial and left ventricular in response to hypertension
Hwang <i>et al.</i> ¹⁶¹	2022	- HT with LVH and LVWT >12 mm - Hypertrophic cardiomyopathy (LVWT ≥15 mm) - AM	Echocardiography	Discriminate between hypertensive heart disease and other heart diseases
Xu <i>et al.</i> ¹⁴³	2022	- Hypertensive heart with history of systemic hypertension or aortic valve stenosis	Echocardiography	Discriminate between hypertensive heart disease and normotensives
Zhang <i>et al.</i> ¹⁶²	2023	- HT with history of hypertension - Hypertrophic cardiomyopathy (LVWT ≥15 mm) - Uremic cardiomyopathy with eGFR <15 mL/min/1.73 m ² - AM	Echocardiography	Discriminate between hypertensive heart disease and other heart diseases
Diao <i>et al.</i> ¹⁴⁰	2023	- HT with increased LVWT (>12 mm) - Hypertrophic cardiomyopathy (LVWT ≥15 mm) - AM	Cardiac MR	Discriminate between hypertensive heart disease and other heart diseases
Wang <i>et al.</i> ¹⁶³	2024	- HT with increased LVWT (>12 mm) with no other cardiac diseases - Hypertrophic cardiomyopathy (LVWT ≥15 mm)	Cardiac MR	Discriminate between HHD and HCM using cardiac images
Wang <i>et al.</i> ¹⁴¹	2024	- HT with increased LVWT (>12 mm) with no other cardiac diseases - History of prolonged uncontrolled arterial hypertension	Cardiac MR	Discriminate between HHD and 10 CVDs using cardiac images
Vasculature				
Recenti <i>et al.</i> ¹³⁰	2021	- Normal - Hypertensive and diabetic	Femoral CT	Identify hypertension from muscle radiodensitometry in CT
AI Task 3 Characterising target end-organ changes				
Retinal microvasculature				
Kaupp <i>et al.</i> ¹⁶⁴	1994	- Hypertensive with SBP: >130 and DBP: >80 and minimal retinal vascular alterations.	Retinal imaging	Measurement of morphological properties of retinal vessels
Khitran <i>et al.</i> ¹⁶⁵	2014	- Normal A/V ratio - Hypertensive retinopathy with 0.1-0.5 A/V ratio		Measurement of morphological properties of retinal vessels
Ahmad <i>et al.</i> ¹⁶⁶	2018	- Normal A/V ratio - Hypertensive retinopathy with abnormal A/V ratio		Classification of arteries and veins
Kiruthika <i>et al.</i> ¹⁶⁷	2019	- Normal A/V ratio - Hypertensive retinopathy with abnormal A/V ratio		Classification of arteries and veins
Dai <i>et al.</i> ¹³⁷	2020	- Hypertensive with SBP: >140 and DBP: >90 with history of hypertension or medication with no comorbidities		Identify hypertension using retinal images
Bhimavarapu <i>et al.</i> ¹⁶⁸	2024	- Hypertensive retinopathy patients with no other retinal diseases - Hypertensive retinopathy of stages mild, moderate, severe, and malignant		Discriminate between hypertensive retinopathy at different stages
Triwijoyo <i>et al.</i> ¹⁶⁹	2025	- Diabetic retinopathy labelled into nine classes hypertensive retinopathy based AVR		Discriminate between hypertensive retinopathy at different stages
Heart				
Raghavendra <i>et al.</i> ¹³⁹	2022	- Normal - HT without renal failure, coronary, congenital, or valvular disease	Echocardiography	Predict hypertension using structural alteration
Alsharqi <i>et al.</i> ¹²⁹	2023	- Hypertensive with SBP: >160		Generate a clinically meaningful scores and phenotypes of cardiac remodelling

Vasculature				
Pessana et al. ¹³¹	2010	- Hypertensives without any medications	Carotid ultrasound	Assessment of instantaneous arterial diameter and WAVP in hypertension
Multi-organ, <i>i.e.</i> , heart, brain, liver, kidney, and vasculature				
Alkhodari et al. ¹³³	2023	- Hypertensive with SBP: >160 and DBP: >100 with no comorbidities	Cardiac MR, brain MR, and carotid ultrasound	Generate a unified measure to identify multi-organ phenotypes and trajectories
AI Task 4				
Identifying phenotypes and hypertension subgroups				
Brain				
Yang et al. ¹⁴⁴	2024	- Hypertensive with SBP: >130 and DBP: >80 with history of hypertension or medication	Brain MR	Discovering hypertension subgroups using imaging and genetic data
Heart				
Katz et al. ¹³⁸	2017	- Hypertensive with SBP: >140 and DBP: >90 with no comorbidities	Echocardiography	Identify hypertensive subgroups with varying myocardial substrates
Loncaric et al. ¹⁴⁶	2021	- Hypertensive with history of antihypertensive treatment and no comorbidities		Identify functional phenotypes in velocity profiles of hypertension
AI Task 5				
Defining risk of future hypertension-related events				
Retinal microvasculature				
Squirrell et al. ⁹¹	2024	- No prior ASCVD	Retinal imaging	Predict an individual SBP from retinal image and correlate it with future ASCVD events
Heart				
Giovanna et al. ¹⁷⁰	2002	- Hypertensive with SBP: >140 and DBP: >90	Echocardiography	Cardiovascular risk stratification using microalbuminuria and echocardiography
Viazzi et al. ¹⁷¹	2006	- Hypertensive with history of antihypertensive treatment	Echocardiography	Cardiovascular risk stratification using creatinine and echocardiography
Other additional applications				
Vasculature				
Jana et al. ⁹²	2020	- Normal and low BP - Elevated and HT with stages 1 or 2	Doppler ultrasound	Identify blood pressure levels using cuffless estimation and Doppler ultrasound

The overall thematic analysis is provided in (Figure 2.6). AI Task 4 included populations with the highest range of mean ages (IQR: 56–65), and AI Task 2 had the highest mean age reaching up to 82 years old in Recenti *et al.*¹³⁰. In contrast, AI Task 3 involved cohorts with the lowest mean ages (Q1: 31), with populations reaching as young as 17 years old, as reported by Alsharqi *et al.*¹²⁹. In terms of the distribution of females across the identified thematic tasks, AI Task 4 had the highest with almost balanced datasets (53%, IQR: 49—57%), while AI Task 2 had the lowest, at 32% (IQR: 28—40%), while in AI Task 3, Pessana *et al.*¹³¹ reported no female participants. Thematic analysis of dataset sizes showed that large datasets, *i.e.*, more than 1,000 participants, were most common in AI Task 4 (median: 1,273, IQR: 533—23,055), in Yang *et al.*¹⁴⁴. and AI Task 5 (404, IQR: 360—38,935), in Squirrell *et al.*⁹¹.

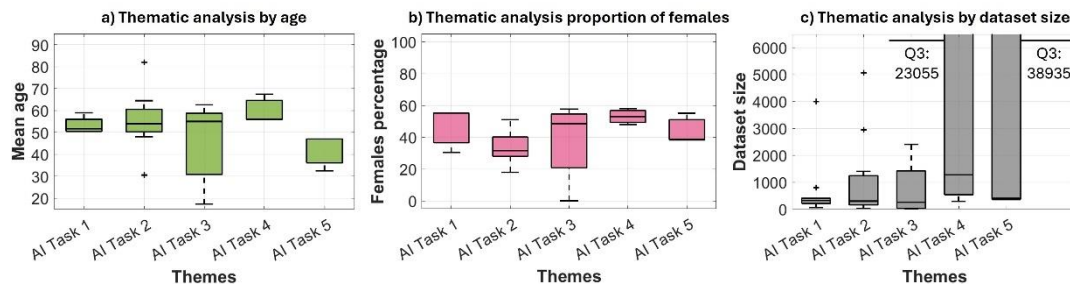


Figure 2.6 | Overall thematic analysis of AI Tasks. Studies were divided based on the five thematic AI tasks and analysed based on: **a)** mean age, **b)** proportion of females, and **c)** dataset size.

2.4.3.1 AI Task 1: Identifying hypertension or its stages

In this task, studies examined multiple organs, with a primary focus on the brain (**Figure 2.7**). Heart studies identified intensity- and texture-based radiomics as impactful metrics for identifying hypertension¹⁵³, outperforming geometric parameters of the myocardium. Furthermore, aortic diameter was higher in hypertensive cases and strongly correlated with the severity of hypertension¹⁴². In brain imaging, cortical thinning, cerebrospinal fluid (CSF) increase, and vascular alterations worsening with higher hypertension grades¹⁴⁸. Moreover, progressive structural changes, mainly reductions in volume, were varied by hypertension severity¹⁵². A single study investigated the use of tongue imaging, reporting a strong association between vascular features in the tongue and concurrent liver and kidney abnormalities¹³².

Most studies employed supervised learning-based models, with the majority relying on deep learning. Overall, they demonstrated a low risk of bias. Models were primarily used for classification, often preceded by segmentation. However, several gaps remain in this body of research. Datasets were relatively small, unsupervised learning approaches were rarely explored, performance was highly influenced by blood pressure thresholds, and the models were purely predictive with limited interpretability.

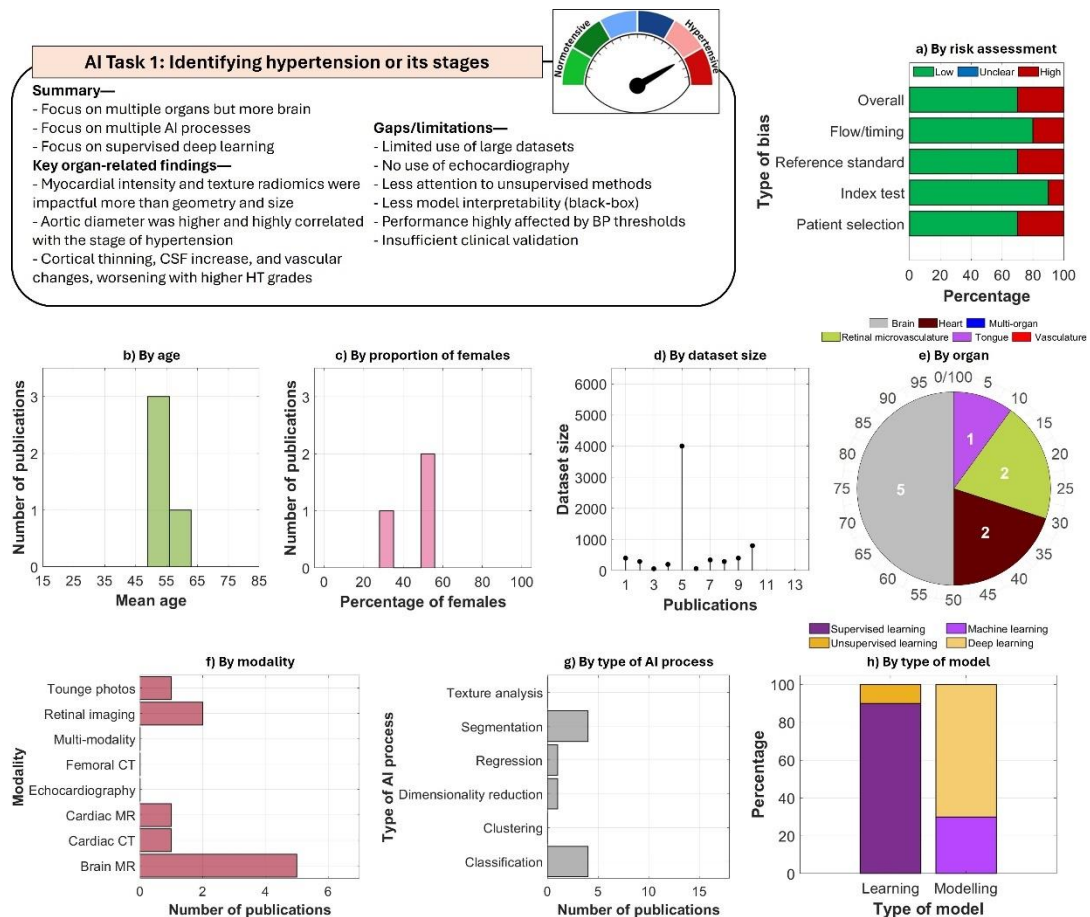


Figure 2.7 | Thematic analysis of the included studies in AI Task 1. Analysis of **a)** risk of bias, **b)** by age, **c)** proportion of females, **d)** size of datasets, **e)** organs being imaged, **f)** imaging modalities, **g)** tasks performed by the AI models, and **h)** types of learning and modelling. The number of publications may vary in plots due to missing information in the original study.

Meta-analysis was conducted on studies that focused on identifying hypertensive patients from normotensives (**Figure 2.8a**). Included studies used data on the heart, brain, and retinal microvasculature. The heart subgroup had a sensitivity of 0.82 [0.79—0.85], the brain showed the highest values at 0.95 [0.83—0.99], and the retinal microvasculature had a single study with the lowest sensitivity of 0.56 [0.40—0.72]. The overall sensitivity was 0.84 [69—93], with overall heterogeneity of 79.2% and no heterogeneity in per-organ analysis. The summary receiver operating characteristics (SROC) curve (**Figure 2.9**) revealed an area under the curve (AUC) of 0.84 [0.79—0.89], while the funnel plot demonstrated symmetry with logit sensitivity of 1.66.

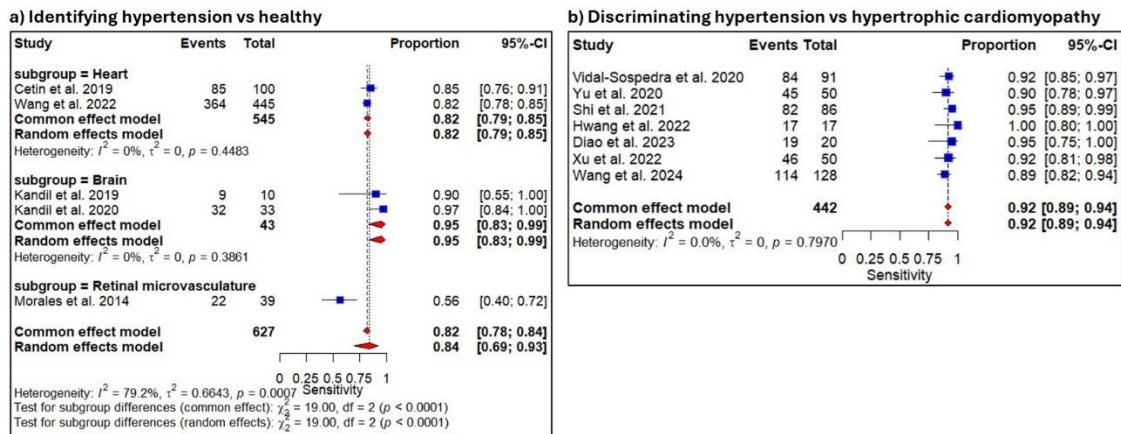


Figure 2.8 | Meta-analysis. a) Identifying hypertension versus healthy conditions in AI Task 1 (5 studies). The analysis was also based on organ subgrouping. **b)** Discriminating hypertension from hypertrophic cardiomyopathy (HCM) in AI Task 2 (7 studies). Only studies that reported performance were included.

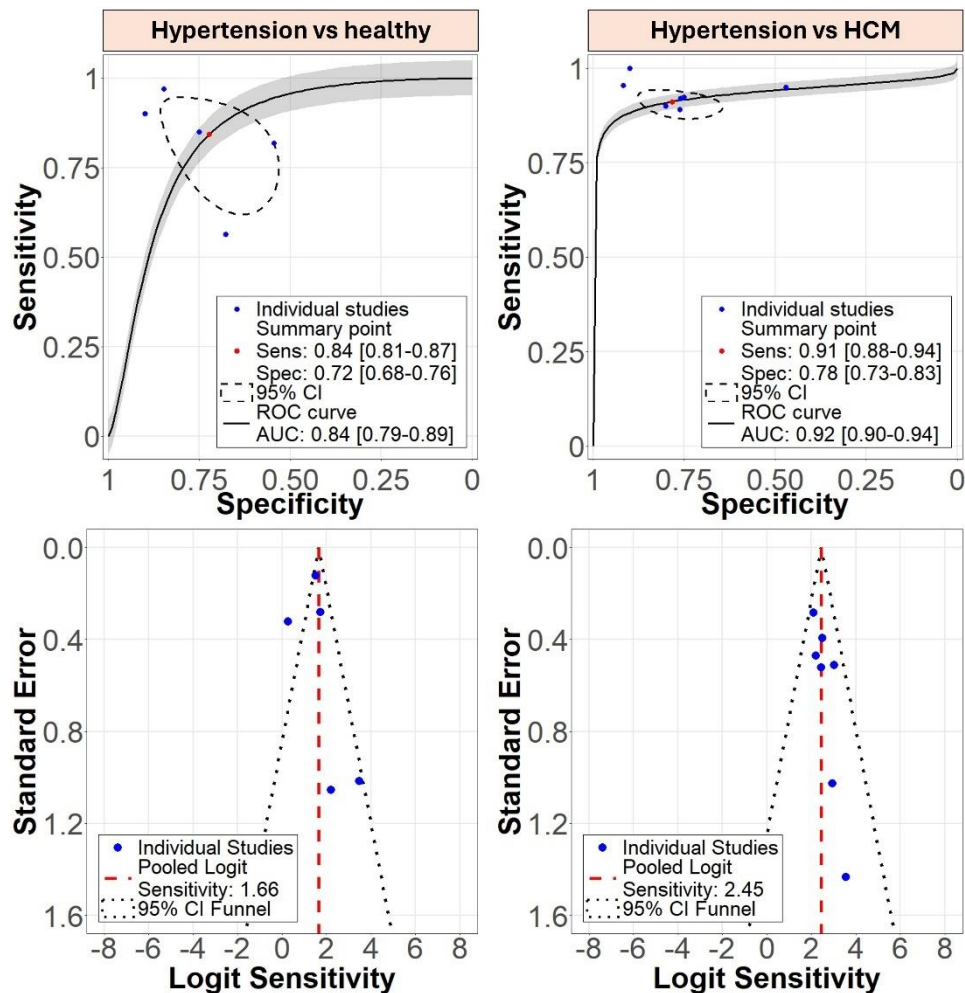


Figure 2.9 | Summary receiver operating characteristics (SROC) curves of the included studies with their corresponding funnel plots. Analysis of studies ($n = 5$) on identifying hypertension from normotension (left). Analysis of studies ($n = 7$) in discriminating between hypertension and hypertrophic cardiomyopathy (HCM) (right). Each SROC was provided with 95% confidence interval, summary point for sensitivity and specificity, and overall area under the ROC curve. Funnel plots were used to check symmetry between studies in their reported performance.

2.4.3.2 AI Task 2: Differentiating hypertension from other diseases

Studies mainly addressed distinguishing hypertensive disease from other cardiovascular conditions, particularly hypertrophic cardiomyopathy (HCM), using cardiac imaging (Figure 2.10). Most focused on the heart using modalities such as cardiac MR, echocardiography, and cardiac CT. Key hypertension markers included reduced IVS thickness, LVPW thickness, and LV mass index (LVMI) compared to HCM^{140,162}. Moreover, radial, circumferential, and longitudinal strain parameters¹⁶⁰, late gadolinium enhancement (LGE) on SAX¹⁴¹, and time/frequency domain contractility features¹⁴³ were effective in distinguishing hypertension. In retinal imaging, severe hypertensive

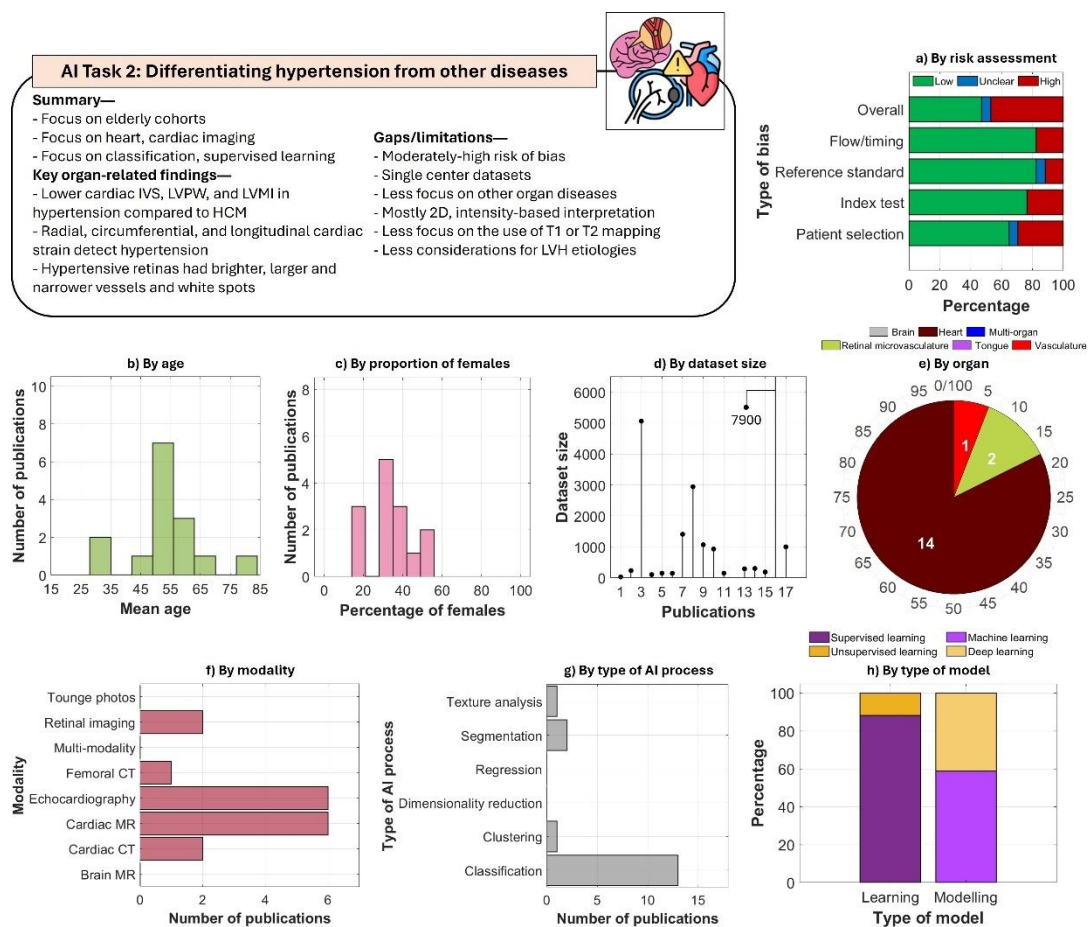


Figure 2.10 | Thematic analysis of the included studies in AI Task 2. Analysis of a) risk of bias, b) by age, c) proportion of females, d) size of datasets, e) organs being imaged, f) imaging modalities, g) tasks performed by the AI models, and h) types of learning and modelling. The number of publications may vary in plots due to missing information in the original study.

retinopathy was characterised by brighter, wider vessels with white spots¹⁵⁴. Additionally, narrower arterioles helped distinguish hypertensive retinopathy from diabetic retinopathy¹⁵⁵.

In this task, supervised classification approaches were more common. Models ranged from traditional machine learning to advanced deep learning architectures, including attention-based transformers¹⁴¹. The overall risk of bias remained moderately high at 47.1%. Several limitations should be noted. First, most studies focused solely on the heart. Moreover, models relied on two-dimensional (2D) imaging with intensity-based interpretations, whereas use of three-dimensional (3D) was lacking. Finally, there was limited focus on T1/T2 mapping and insufficient consideration of LVH etiologies.

Meta-analysis here involved studies that distinguished hypertension from HCM (**Figure 2.8b**). Seven studies were included in the analysis, demonstrating an overall sensitivity of 0.92 [0.89—0.94]. No heterogeneity was found between the studies. AUC was at 0.92 [0.90—0.94] with funnel plot symmetry around logit sensitivity of 2.45 (**Figure 2.9**).

2.4.3.3 AI Task 3: Characterising target end-organ changes

The primary goal of this task was to detect organ alterations using AI and validate its effectiveness in identifying hypertension (**Figure 2.11**). Retinal imaging was the most common imaging studied, with few others on cardiac and vasculature imaging. Heart findings highlighted that aerobic exercise alone had minimal impact on blood pressure¹²⁹. Moreover, geometrical properties derived from the SHELet transform were found key markers of hypertension¹³⁹. In retinal imaging, alterations in hypertensive retinopathy included an arteriovenous ratio (AVR) below 0.533, approximately 25% greater arterial tortuosity compared to normotensives, and significantly reduced arterial diameter¹⁶⁴. Moreover, strong detection performance near the optic nerve and in images

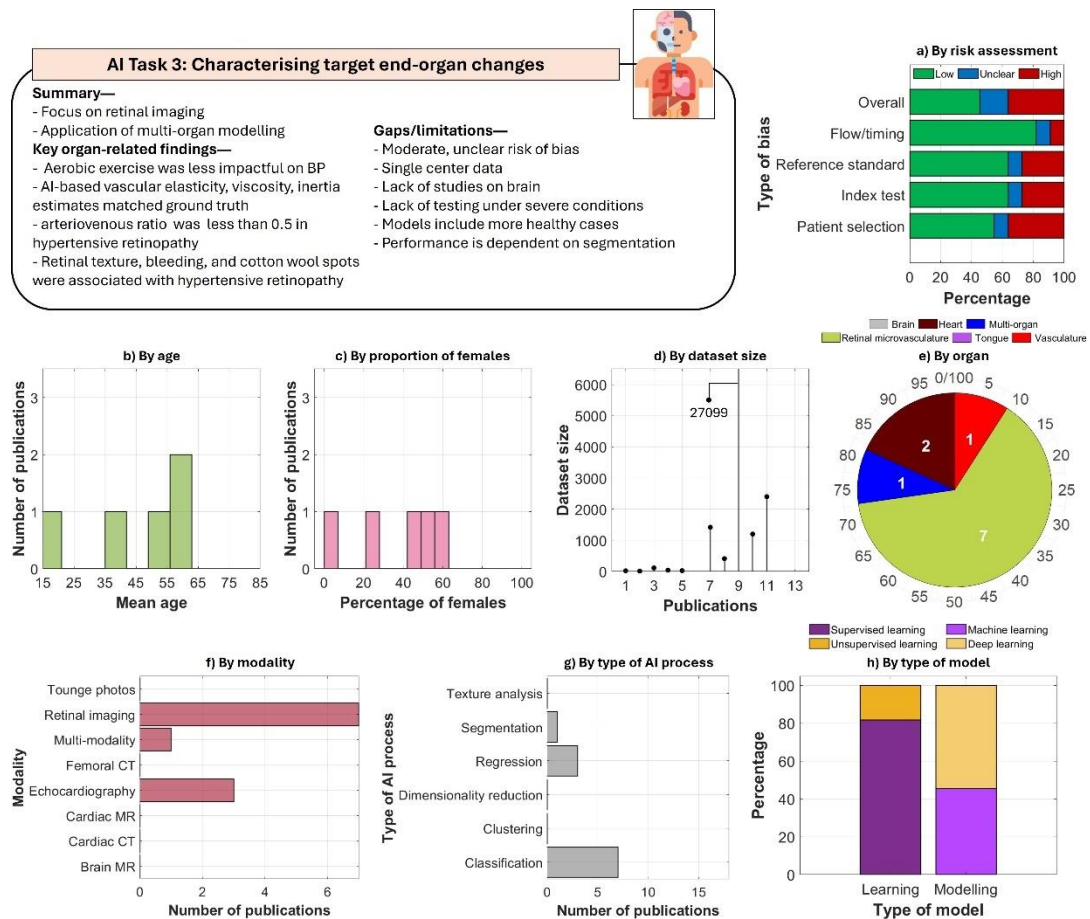


Figure 2.11 | Thematic analysis of the included studies in AI Task 3. Analysis of **a)** risk of bias, **b)** by age, **c)** proportion of females, **d)** size of datasets, **e)** organs being imaged, **f)** imaging modalities, **g)** tasks performed by the AI models, and **h)** types of learning and modelling. The number of publications may vary in plots due to missing information in the original study.

with thin vessels¹⁶⁸, alongside changes in retina texture, blood vessels, hard exudates, bleeding, and cotton wool spots¹⁶⁹. A single study integrated multi-organ information using multi-modality imaging in an AI model¹³³, identifying unique hypertension associations with white matter hyperintensities (WMH), grey matter volume, and intima-media thickness (IMT).

The overall risk of bias in this task was moderate, with 18% of studies classified as having unclear risk. Studies under this task had several limitations. None examined the brain, models were not tested under severe conditions to evaluate feature extraction in complex scenarios, and predictive performance relied heavily on the quality of segmentation and image analysis.

2.4.3.4 AI Task 4: Identifying phenotypes and hypertension subgroups

Studies under this task focused mostly on the heart using echocardiography (**Figure 2.12**). AI-based cardiac analysis identified distinct hypertension-related phenotypes, including one associated with worse outcomes and abnormalities in LV mass, LV end-diastolic dimension (LVED), LV stroke volume (LVS), LV wall thickness (LVWT), and reduced longitudinal strain¹³⁸. Unique phenotypes were also characterised by significant differences in fused E and A waves, septal/lateral mitral annular e', mitral average, and LV function¹⁴⁶. A single study identified five distinct subtypes using brain MR¹⁴⁴, with

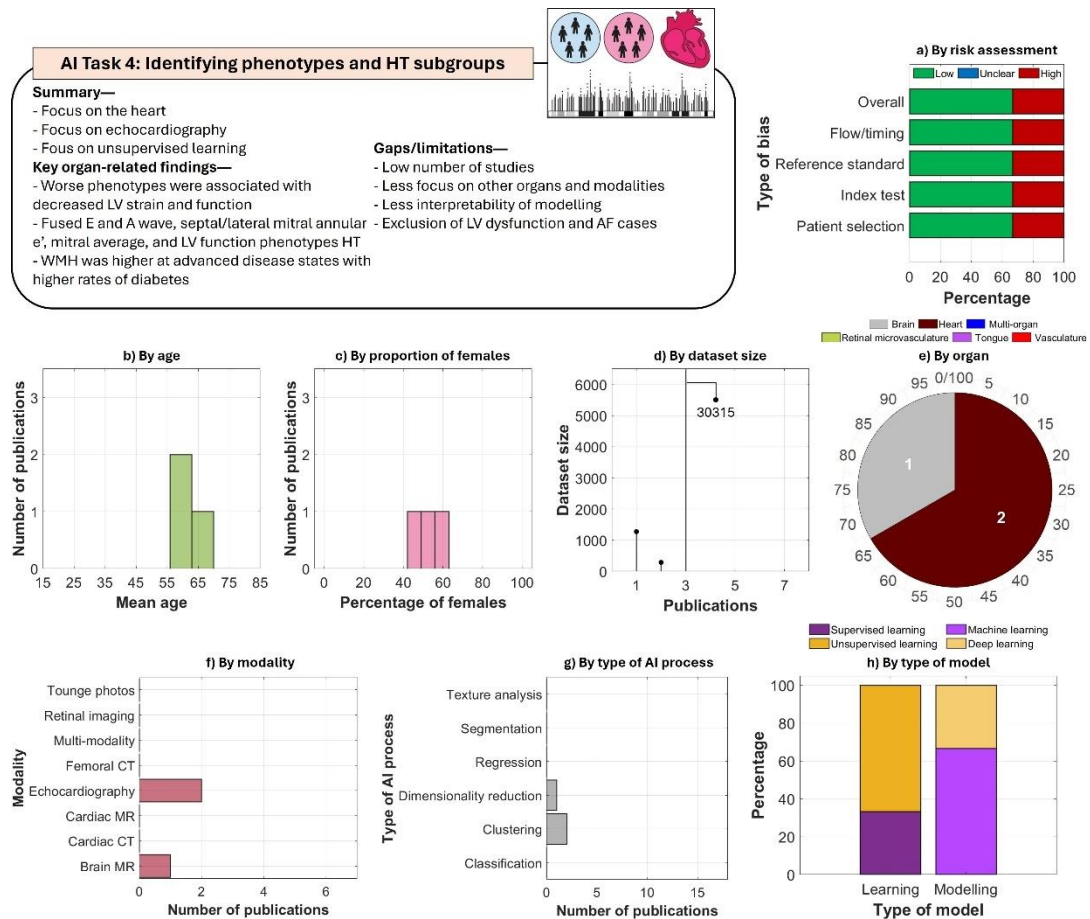


Figure 2.12 | Thematic analysis of the included studies in AI Task 4. Analysis of **a)** risk of bias, **b)** by age, **c)** proportion of females, **d)** size of datasets, **e)** organs being imaged, **f)** imaging modalities, **g)** tasks performed by the AI models, and **h)** types of learning and modelling. The number of publications may vary in plots due to missing information in the original study.

two advanced subtypes showing higher WMH and increased rates of diabetes, while no differences were observed in hyperlipidaemia.

In contrast to other tasks, this task more prominently employed unsupervised machine learning. Notably, this task exhibited low risk of bias. Nonetheless, several gaps remain to be addressed. The number of studies is still limited, with most not exploring phenotyping based on organ characteristics beyond the heart. Only a single study applied deep learning-based analysis¹⁴⁴ that culminated in clustering and phenotype discovery, leaving room to explore more advanced techniques. Additionally, models often excluded severe cases, *e.g.*, LV dysfunction and atrial fibrillation (AF) cases. Finally, given the reliance on machine learning models, interpretability was often limited, with explanations mostly derived from clustering analysis of identified phenotypes.

2.4.3.5 AI Task 5: Defining risk of future hypertension-related events

This task included studies focusing on the heart and the retina, to predict future risk and adverse outcomes (**Figure 2.13**). Heart studies found that higher future cardiovascular risk was associated with reduced creatinine clearance and estimated glomerular filtration rate (eGFR)¹⁷¹. Additionally, microalbuminuria in hypertension correlated with increased LVMI, left ventricular hypertrophy (LVH), and IMT¹⁷⁰. In retinal imaging study, narrower retinal arterioles and wider venules conferred long-term risk of mortality⁹¹.

This task was relatively underexplored. All studies employed supervised deep learning approaches. The included studies demonstrated a low overall risk of bias. Several gaps remain in this task. There is still room for exploration on other organs, especially the brain, for predicting future adverse outcomes. No studies employed processes beyond

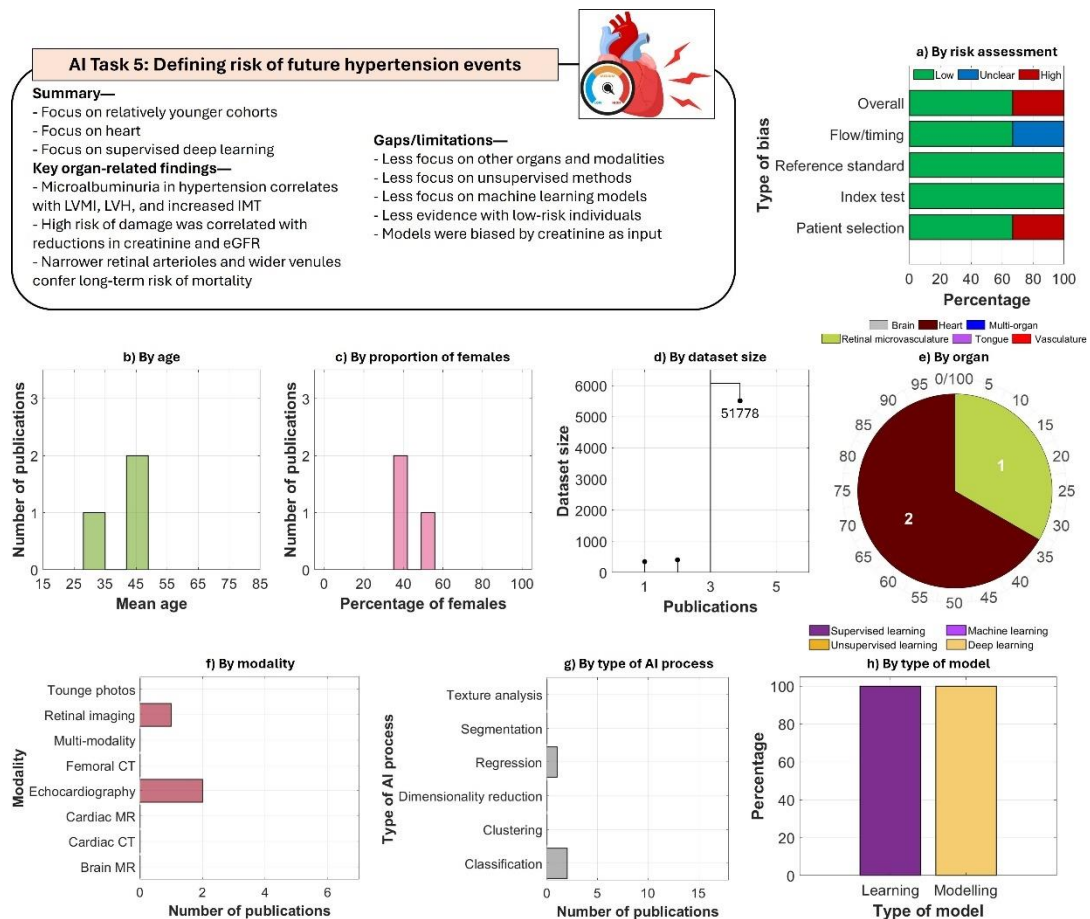


Figure 2.13 | Thematic analysis of the included studies in AI Task 5. Analysis of **a)** risk of bias, **b)** by age, **c)** proportion of females, **d)** size of datasets, **e)** organs being imaged, **f)** imaging modalities, **g)** tasks performed by the AI models, and **h)** types of learning and modelling. The number of publications may vary in plots due to missing information in the original study.

classification, such as segmentation or dimensionality reduction. Moreover, unsupervised machine learning algorithms were not utilised. Finally, the models provided limited evidence for predicting future risk in low-risk individuals, as they often relied on already severe cases that are easier to identify.

2.5 Discussion and conclusion

In this chapter, a systematic review of 45 studies was conducted to examine the integration of advanced AI and computational algorithms with multi-modal imaging to enhance hypertension identification, phenotyping, and outcome prediction. The emergence of several thematic areas across these studies underscores the broad

applicability of AI and computational algorithms in this context, ranging from hypertension diagnosis to enabling more personalised management approaches that move beyond blood pressure-based classifications.

Traditionally, hypertension has been identified through cuff-based blood pressure measurements; however, this approach could be supplemented by AI-driven image analysis. In the context of AI Tasks 1 and 2, the detection of hypertension based on end-organ changes, which are comparatively more stable than blood pressure readings, could signpost clinicians towards possible cases of masked hypertension that would otherwise go unrecognised⁹². Beyond aiding hypertension diagnosis, AI has also shown the capability to guide clinical diagnosis, particularly in distinguishing hypertensive end-organ phenotypes from other closely related cardiovascular diseases. These diseases often present with overlapping features, posing a diagnostic challenge for clinicians.

AI models trained on imaging data have demonstrated strong performance in identifying hypertension-induced organ changes, as outlined in AI Task 3, and effectively identifying hypertension. These models could detect earlier manifestations of organ pathology, necessitating interventions beyond traditional antihypertensive medications that require more aggressive management strategies at an earlier stage. However, the current research is still limited in characterising cerebral alterations with AI despite that brain is well-known to exhibit cerebral changes, including increased white matter hyperintensities and altered cerebral perfusion¹⁴⁴.

A natural next step is the integration of AI into clinical risk stratification frameworks. Existing risk scoring tools, such as QRISK3²⁵, primarily rely on blood pressure, demographics, and laboratory data to estimate future risk of cardiovascular events. However, these scoring systems do not directly take into account severity of target organ

damage, which likely contributes significantly to individual risk profiles. Rather than adopting a one-size-fits-all approach, precise phenotyping enables the identification of hypertension subgroups based on organ-specific alterations, paving the way for more personalised and effective management strategies. Current national guidelines for hypertension management, *e.g.*, NICE, recommend control with blood pressure targets determined by underlying comorbidities, age and estimated cardiovascular risk. Whilst practical, such approaches do not directly account for underlying disease phenotypes. It is plausible that more nuanced approaches to management could be developed with improved risk prediction upon understanding organ-specific phenotypes.

No studies, however, sought to specify the duration of exposure to hypertension, likely because it is challenging to determine with less availability of longitudinal datasets and because early-stage hypertension is often asymptomatic and therefore it is difficult to establish when it started. Accordingly, none of the studies attempted to describe the changes in organ features occurring over time with repeated imaging. This omission is significant, as individuals with substantial end-organ changes despite a relatively short duration of hypertension may represent an accelerated phenotype¹⁷², highlighting the need for further investigation into the temporal progression of hypertensive damage.

2.5.1 Study limitations

Our review has limitations. First, despite the overall number of studies included in this systematic review, the number of studies within each subcategory, *i.e.*, task-based or organ-based within tasks, was relatively low, primarily due to the limited application of AI and medical imaging in hypertension care. Second, although the review aimed to cover all organs across various imaging modalities, the current literature has yet to comprehensively explore AI-assisted imaging of certain organs such as kidneys,

abdomen, or lungs. Third, most included studies were retrospective or cross-sectional, limiting the ability to assess longitudinal disease progression and AI's role in long-term hypertension management. Last, most of the included studies suffered from lack of external validation, which increased bias towards their models' training strategies.

2.5.2 Conclusions

This systematic review highlights a nascent potential for AI-driven computational imaging techniques in the early detection, risk stratification, disease phenotyping, and personalised management of hypertension, which goes beyond traditional diagnostic approaches. There is a long history of hypertension research, using traditional imaging methods and analysis, that has demonstrated hypertension can impact organ structure and function. The limited body of research applying AI and computational techniques to further develop this understanding is therefore surprising and suggests the potential use of AI for personalised hypertension detection and management lags behind other clinical specialities. This review has identified opportunities for future research to explore the impact of hypertension across multiple body organs. In addition, research could be directed towards identifying different subtypes that may benefit from targeted management or drug development.

3. *HYTWIN* DEVELOPMENT AND MODELLING

3.1 Abstract

Background. Hypertension affects the heart, brain, kidneys, and other vital organs. However, current clinical management primarily focuses on controlling blood pressure alone, often overlooking early and emerging signs of end-organ damage.

Objectives. To develop a machine learning model capable of revealing the predominant phenotypes of hypertension-related end-organ changes and to characterise their progression from healthy to diseased.

Methods. Imaging and non-imaging data from the UK Biobank, including 27,099 individuals, was obtained, curated, and pre-processed. A machine learning approach based on the contrastive trajectory inference (cTI) algorithm was used to develop the model, named the *HyTwin*. The performance of the model was internally validated for stability using a 10-fold cross-validation scheme and for clinical relevance using time-to-event Kaplan-Meier analysis.

Results. Our analysis revealed six distinct hypertensive disease phenotypes, namely the *HyperTrajectories*, each characterised by specific alterations related to cardiac-, lipoprotein-, atherothrombosis-, brain-, mortality-, and liver-features, respectively. The *HyTwin* model also positioned individuals along one or more of these *HyperTrajectories* and quantified the severity of their organ damage, namely the *HyperScore*. The *HyperTrajectories* were associated with biomarkers and clinical outcomes aligned with the identified predominant characteristics, supporting their clinical relevance. Internal cross-validation showed that both *HyperScore* and *HyperTrajectory* were robust to variations in training datasets.

Conclusion. The findings establish a new framework for hypertension assessment and monitoring, moving beyond blood pressure alone to enable earlier and phenotype-specific interventions.

Publication Status.

The results presented in this chapter have been presented and published in:

1. **Alkhodari M**, Lapidaire W, Xiong Z, Hadjileontiadis L, Khandoker A, Banerjee A, Leeson P. *HyperScore: a unified measure to model hypertension progression using multi-modality measurements and semi-supervised learning*. 17th IEEE international conference on bioinformatics and biomedicine (BIBM). Istanbul, Turkey. 2023.

More results presented in this chapter are currently under-review.

2. **Alkhodari M**, Lapidaire W, Kart T, Xiong Z, Krasner S, Fletcher A, Savage N, Baumeister T, Ohuma E, Namburete A, Lamata P, Iturria-Medina Y, Chappel T, Thilaganathan B, Fraser A, Mackillop L, McManus R, Ntusi N, Hadjileontiadis L, Khandoker A, Lewandowski A, Banerjee A, Leeson P. *Contrastive machine learning to quantify hypertensive multi-organ damage and identify new disease phenotypes: a multi-national multi-modal study*.

3.2 Introduction

Hypertension is a leading risk factor for cardiovascular disease-related deaths, which account for nearly 50% worldwide¹⁷³. Elevated blood pressure causes structural and functional changes in the heart, vasculature, and brain that predispose to vascular events and death²⁹. Notably, these alterations start to be observed in some individuals even with mild blood pressure elevations¹¹¹, while some will not develop end-organ changes despite prolonged exposure to hypertension³². Therefore, treatment based on blood pressure level alone may under- or over-treat at an individual patient level. Including extent and type of end-organ damage in clinical decisions, such as blood pressure treatment thresholds or medication selection, would require a holistic understanding of how organ remodelling emerges and changes during the life course of hypertensive disease.

Implementation of end-organ damage into clinical decision making has so far been limited by lack of a holistic understanding of how organ remodelling emerges, interacts and changes during the life course of hypertensive disease. Such information would allow severity of global end-organ disease state to be estimated for an individual and identify unique features in their pattern of multi-organ remodelling that may benefit from personalised treatment. Traditionally, large, diverse, longitudinal studies of multi-modal imaging data would be required to develop these models while also considering the most recent techniques in data analysis, learning, and integration.

In this chapter, I hypothesise that applying a semi-supervised contrastive machine learning methodology to a sufficiently large cross-sectional imaging dataset can address existing challenges and, for the first time, characterise the unique patterns of multi-organ change associated with hypertension. The resulting model, named *HyTwin*, is

hypothesised to be able of estimating an individual's position relative to a healthy state, quantified as a measure of end-organ disease severity, named the *HyperScore*. Additionally, the individual's position relative to identified clusters, named the *HyperTrajectory*, can be used to determine their similarity to distinct patterns of end-organ change.

3.3 Methods

3.3.1 Dataset for model development

For model development, I used data from the Imaging Enhancement Study (2015-ongoing) of UK Biobank¹⁷⁴, an open-access resource that includes a large prospective cohort (age: 40-69 years old). The imaging substudy aims to collect multi-organ and multi-modality (brain, heart, and abdomen) scans of 100,000 participants¹⁷⁵. The study complies with the Declaration of Helsinki and was approved ethically by the Northwest Multi-Center Research Ethics Committee (MREC) as a Research Tissue Bank (RTB) approval (REC reference: 11/NW/0382) and received approval extension on 2021 (REC reference: 21/NW/0157). A written consent form was obtained from every participant prior to conducting the imaging study.

The analysis included the currently available data of 46,102 participants. Each participant had 2,442 clinical measurements recorded using multiple tools including clinic resting blood pressure, brain MR, cardiac MR, abdominal MR, body composition, arterial stiffness, carotid ultrasound, spirometry, blood biochemistry, ECG, and baseline/demographic characteristics.

3.3.2 Development of end-organ model

3.3.2.1 Pre-processing and data preparation

First, all participants with missing blood pressure measurements or more than 5% of missing variables were removed ($n = 18,038$). In addition, participants who have suffered from prior heart attack, angina, or stroke ($n = 965$) were also removed. Variables with more than 50% missing values were removed from the dataset ($n = 1,388$). The final dataset included 27,099 participants with 1,054 multi-modality measurements. Any values above 5 standard deviations (SD) above or below the mean, *i.e.*, outliers, were removed and replaced as missing. The median of the selected variable then imputed all remaining missing values.

3.3.2.2 Feature selection to reduce redundancy

Feature selection was performed to reduce redundancy and enhance variability across each modality. The selection was based on reducing the overall covariance between variables where I excluded one of the two variables with covariance above 95%. This has a major impact on modalities with many variables, *i.e.*, brain MR and body composition, which required additional selection through contrastive covariance analysis. A contrastive covariance matrix was computed as the subtraction between the normotensive and hypertensive covariance matrices so that variables greater than twice of the standard deviation were excluded. Accordingly, variables were reduced from 1,054 to 566 collecting most of the complex relationships at every modality; it contained 449 brain MR, 13 cardiac MR, 12 carotid ultrasound, 24 body composition, 58 blood biochemistry, 3 ECG, 5 spirometry, and 2 demographic characteristics.

3.3.2.3 Adjusting covariates and harmonisation

Covariates were adjusted to reduce the effects of confounding factors, *i.e.*, age and sex, that were not controlled during the study. In addition, I used the recent ComBat tool with empirical Bayes shrinkage approach¹⁷⁶ to perform data harmonisation across the three imaging centres where the imaging substudy of the UK Biobank was conducted. Harmonisation was important to reduce the variance caused by non-biological sources in multi-centre studies.

3.3.2.4 Definition of groups for contrastive analysis

The proposed study requires contrastive modelling using participants considered the ‘healthiest’, *i.e.*, having similar measurement values representing the healthy state and comparing these to participants likely to be representative of advanced disease state. In this analysis the ‘healthy’ state is individuals who would not be expected to have evidence of hypertensive end-organ damage, and I refer to these as normotensive (N). I defined this N group as participants having systolic and diastolic blood pressure below 120/80 mmHg during their first imaging visit. In addition, they had never received any blood pressure medications and had never been diagnosed with hypertension. In addition, as several of the N group, despite having normal blood pressure levels may still have exhibit altered multi-organ clinical measurements due to other factors such as athletic performance or inherited abnormalities, I identified any outlier participants in this group (the N+ group), using an isolation forest algorithm¹⁷⁷. A participant was assigned as N+ if the outlier probability given by the isolation forest was more than the third quartile of all probabilities (≥ 0.43). This reduced the N group by 344 individuals. N and N+ groups were significantly different in key organ measurements.

For the contrastive analysis, I then identified the group most likely to have evidence of end-organ disease and refer to them as the hypertensive end-organ group (H). These participants had either a systolic or diastolic blood pressure of more than 160 or 100 mmHg, respectively, in line with stage 2 hypertension as defined by all three hypertension management guidelines. All remaining participants were assigned to an intermediate elevated group (E) and were not used for the contrastive principal component analysis.

3.3.2.5 Dimensionality reduction and cPCA

To capture the most informative components in the data and reduce dataset dimensionality, I used the cPCA algorithm¹⁷⁸ described in **Chapter 1**. In this study, a logarithmic range of 50 weighting factors from $\alpha = 10^{-2}$ to 10^1 was applied to the weighted subtraction, and the optimal factor was determined iteratively based on an evaluation of the ability of the model to differentiate group H from group N (**Figure 3.1**). This performance testing allowed us to identify settings that allowed the highest discrimination between the two groups, which had been predefined to represent the two most different end-organ disease states. On each iteration, regular PCA was performed and the optimal α was determined when there was a high clustering tendency between N and H groups (using spectral and k -means clustering). The eigenvectors produced from the optimal α were the final reduced contrastive principal components (cPCs) for the dataset.

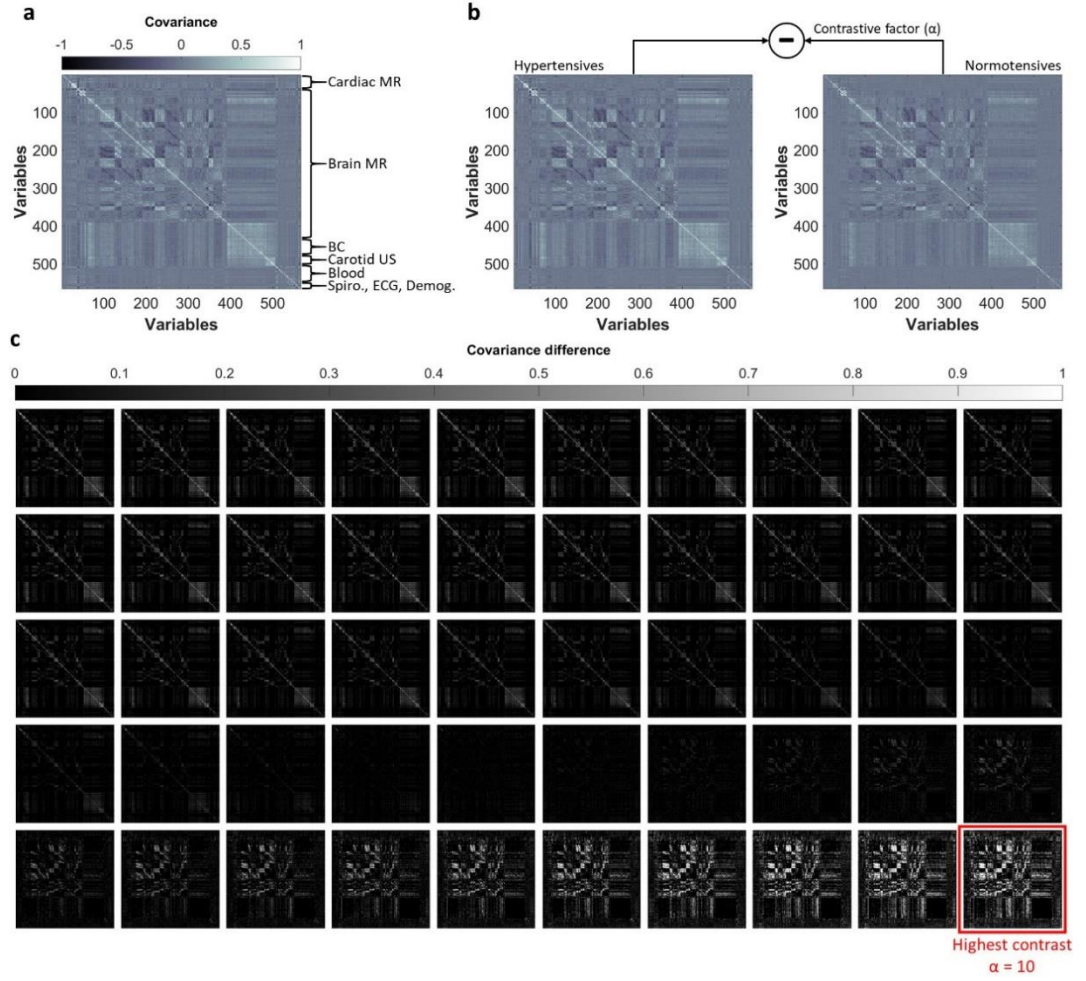


Figure 3.1 | Preparation and selection of the optimal contrastive covariance matrix for the normotensive and hypertensive participants in the overall dataset. a, Covariance representation of the combined set with 566 variables ordered by modality. **b,** Subtraction between the covariance matrices of the hypertensives (left) and normotensives (right). The normotensives covariance matrix is iteratively multiplied by a contrastive factor (α) prior to the subtraction. **c,** Generating 50 contrasts of the resulting subtracted covariance and selecting the optimal factor with the highly contrastive covariance for the contrastive principal components analysis (cPCA).

The most important variables were determined using a cut-off threshold of the expected contribution value ($C_{expected}$) applied to the variable importance, C_i , as follows,

$$C_{expected} = 100 \times \left(\frac{\sum_{k=1}^K \lambda_k}{N} \right)$$

where K is the total number of cPCs, λ is the eigenvalues, and N is the total number of variables.

3.3.2.6 Contrastive trajectory inference (cTI)

Upon obtaining the reduced dimensional cPCA space, I performed contrastive trajectory inference (cTI)¹⁰⁷ by first calculating the Euclidean distance matrix of the cPCA components among all participants, from groups H, E and N. Then, I selected the healthiest participant, *i.e.*, the closest to all other N participants in the distance, as the root node. The root node was then used to create an MST¹¹⁰ un-directed graph (**Figure 3.2a**). The distance of each participant from the root node was generated within a normalised (range: 0—1) relative to the maximum value (the farthest distance from the root node) that should represent the extreme end-organ condition. This value thereby provides a quantitative measure of end-organ disease state, *i.e.*, the proximity of each participant from the pathology-free state relative to the most advanced disease state.

As participants were not distributed homogeneously through the cPCA space, I also investigated whether there were dominant ‘clusters’ of participants tracking through the space from the root node to various advanced disease states, *i.e.*, ‘trajectories’ of end-organ change within the space. To do this, I identified dominant contrasted pseudo-temporal connections in the MST. I used Laplacian transformation¹⁷⁹ to create an

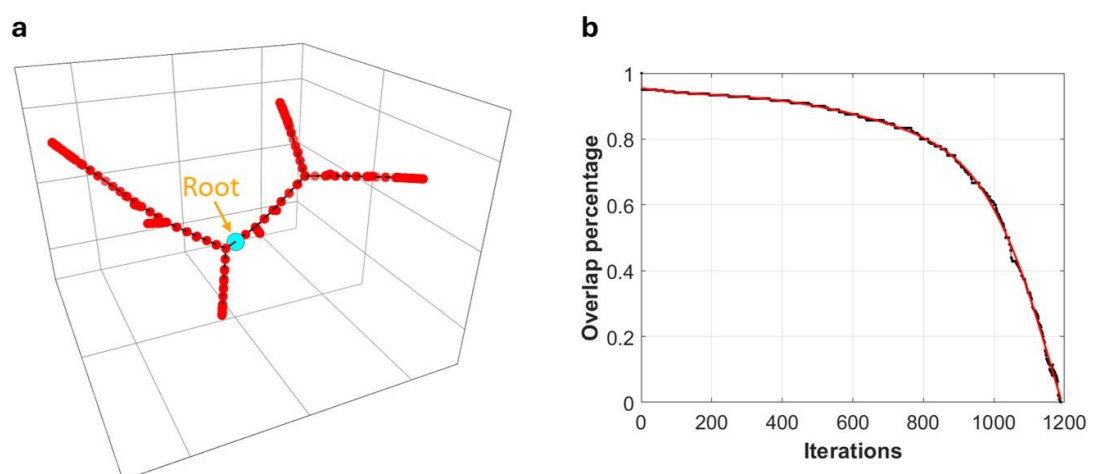


Figure 3.2 | Contrastive trajectory inference (cTI) and construction of the minimum spanning tree (MST). **a**, cTI defines an MST un-directed graph by connecting all individuals in a way that minimises the total distance from the root node. **b**, Merging of overlapping trajectories reduces redundancy and identifies “clusters” in the dataset, which are then used to define major phenotypes.

embedding eigenspace with no dimensional meaning for the axes (**Figure 3.2a**). I then defined major pathways in the constructed graph (**Figure 3.2b**) by following three major steps. First, minor pathways with more than 95% overlap with any other ‘trajectory’ were combined. Then, the objective merging of highly overlapping ‘trajectories’ was performed iteratively, where on each iteration, the highest overlapping two ‘trajectories’ were determined and combined as one. This was performed until 25 ‘trajectories’ remained. As some of these ‘trajectories’ contained a very small number of participants, further combinations were then performed of highly overlapping ‘trajectories’, until each major ‘trajectory’ included at least 10% of the participants; a threshold which I chose a priori to avoid bias in trajectory definition. After this step, six ‘trajectories’ were definable.

3.3.3 Generation of *HyperScore* and *HyperTrajectory* predictions

3.3.3.1 Definition of *HyperScore* and *HyperTrajectory*

For ease of description, I refer to the distance of each participant from the root node, *i.e.*, the quantitative measure of end-organ disease state as their, *HyperScore*. Similarly, I also describe the ‘trajectory’ that each participant was allocated to as their, *HyperTrajectory*. To simplify group comparison, I also allocated participants to three *HyperScore* stages based on the distribution of scores within the H group, being the group defined as likely to have end-organ damage. The first ‘minimal end-organ disease’ stage was thereby defined by a *HyperScore* between 0 and the first quartile (Q1) of participants within the H group. This threshold was also consistent with the threshold identified in a Youden index analysis¹⁸⁰ as the value that most effectively discriminated the N from the H

groups. The second stage, ‘medium end-organ disease’ stage, was the range of scores between Q1 and the third quartile (Q3) of the H group. Lastly, the third stage, ‘high end-organ disease’ stage, is the range of scores between Q3 and the maximum of all scores.

3.3.3.2 *HyperScore* and *HyperTrajectory* unseen predictions

In addition, I developed methodology to assign a *HyperScore*, and identify the most likely *HyperTrajectory*, for participants who were not included in the development of the baseline model. To estimate a *HyperScore* for hidden participants, *e.g.*, those in any subsequent ‘testing dataset’ or new patients in any clinical application, I used a *k*-nearest neighbour (*k*-NN) approach¹⁸¹. Here, I first transformed hidden data using the pre-existing cPCA space to form their principal components. Then, a *k*-NN of 25 neighbours was used to find the most similar participants in the original dataset to the hidden data. A *HyperScore* was assigned accordingly as the mean of the closest 25 participant scores. The selection of *k* to be 25 was based on quantitative analysis to determine the optimal number of close participants to the testing individual while ensuring less variability between their cPCs. Similarly, by using a *k*-NN approach I was able to determine the likelihood for a participant to be on any *HyperTrajectory* based on the allocation of each nearest neighbour to different trajectories. As it was possible that not all 25 nearest neighbours were allocated to the same *HyperTrajectory*, particularly if participants were positioned nearer to the root node, where there was less distance between trajectories, participants could be provided with a likelihood, or probability, for allocation to each trajectory.

3.3.4 Internal model validation

3.3.4.1 Stability of model generation of *HyperScores*

To understand the stability of the modelling approach as a quantitative measure of end-organ changes, I performed internal validation by building models based on smaller subsets of the full dataset. I then used this model to predict *HyperScores* for participants in the held-out data. The re-generated *HyperScore* was compared with the score generated originally when modelling with the full dataset, and size of differences between scores evaluated using root mean square error (RMSE), with an expectation that there would be <10% variation between scores. I did this in two ways. Firstly, using a stratified 10-fold cross-validation scheme. A random 10% of the input data was removed at each iteration, and the modelling was performed to generate a new set of *HyperScore* for the remaining participants. Secondly, I performed this testing step by building two separate models based on half the available data. I did this by splitting the data randomly into two datasets representing 50% of the participating cohort and also by splitting the dataset into two separate models based on the time of the imaging visit, namely from 2015 to 2018 and from 2018 to 2021. Normalisation and data imputation were performed on each subset separately prior to modelling.

As an additional exploration of the stability of the modelling in the absence of ground-truth, I also followed a confidence-based performance estimation (CBPE) approach¹⁸². Here, performance was determined by means of the overall probability towards a certain condition, *i.e.*, presence of hypertensive end-organ changes. This test is based on where an individual was positioned relative to the optimal threshold previously identified for differentiation of *HyperScore* range between ‘no end-organ disease’, or group N, and ‘presence of end-organ disease’, or group H. It allows for determining an optimal

threshold to group participants based on end-organ damage, independently of blood pressure grouping. The generated *HyperScores* were used as the predicted probability. A confusion matrix was created to calculate performance metrics based on the sum of calibrated probabilities.

I further evaluated stability by studying ‘data drift’ when continuously adding new data to the trained model. To evaluate this, I first held out 10% of the data as a test set, trained the model on 50% of the remaining data, and kept adding 10% iteratively (5 iterations). At each iteration, the RMSE was calculated between the first modelled *HyperScores* of the test set and the new scores generated after the addition of new data. The same approach was evaluated for the CBPE metrics (optimal threshold and AUC), where I started with 20% of the data, then continuously added 20%, and observed the changes.

3.3.4.2 Stability of allocation of *HyperTrajectory* probability

To understand the stability of the modelling as an approach to identify clusters of individuals with similar patterns of end-organ changes, I performed internal validation with stratified 10-fold cross-validation. As the trajectories rely on identification of paths between the root node and most extreme end-organ disease state, I ensured the training set always included the 5% of individuals closest to the root node, *i.e.*, healthiest participant in the dataset, and the 5% of participants representing the most severe form of end-organ changes within the full dataset. I then used the MST graph created with each fold of data and regenerated six major pathways within that data. Accordingly, each individual within the held-out data was then attributed a probability for location on each trajectory, their *HyperTrajectory* probability. This allocation was compared to their original attribution within the full model.

As each individual is attributed probabilities across the six trajectories, I firstly evaluated similarity simply based on prediction of the top-2 most likely trajectories, *i.e.*, a prediction is considered correct if it was predicted among the highest two predicted trajectories. A confusion matrix was constructed, and the performance metrics were calculated relative to the ground truth, *i.e.*, the original model. Additionally, to take account of the full distribution of predictions for each individual, I also evaluated performance using Jensen-Shannon divergence (JS)¹⁸³ which compares two probability distributions to measure their similarity. The divergence is calculated between two probability distributions (P) and (Q) as follows,

$$JS(P||Q) = \frac{KL(P||M) + KL(Q||M)}{2}$$

where KL is Kullback-Leibler divergence¹⁸⁴ and $M = (P + Q)/2$ is the matrix distribution.

JS is bounded by 1 for two probability distributions, given that the base 2 logarithm is used. In general, the lower the JS value, *i.e.*, close to zero, indicates greater similarity between the two distributions, whereas the higher JS value, *i.e.*, close to 0.7, indicates dissimilarity between the two distributions. Here, I compared between the distribution of probabilities for the original *HyperTrajectory* and the per-fold trajectories for the three *HyperScore* stages defined earlier.

3.3.5 Clinical relevance of *HyperScore*

To test whether the *HyperScore* had biological and clinical relevance, I first visually assessed how individual clinical phenotypes known to vary with exposure to hypertension, *e.g.*, white matter hyperintensities and left ventricular volume, varied across the *HyperScore* range to sense check direction and shape of association.

I then hypothesised that individuals with more advanced end-organ changes should be closer to having a hypertensive-related event, *e.g.*, stroke, myocardial infarction or cardiovascular death, than individuals who may have similar blood pressure but have less evidence of end-organ disease. To test this hypothesis, I used prospective follow-up data collected on individuals who had participated in the UK Biobank Imaging Enhancement and studied time to event within each *HyperScore* range, *i.e.*, low, medium and high using Kaplan-Meier¹⁸⁵ and log-rank (Mantel-Cox) testing with a significance threshold of 0.05. The diagnosis was derived from the International Classification of Disease Version 10 (ICD-10) reports derived from Hospital Episode Statistics (HES) data and analysed as an event-free probability.

Additionally, to evaluate the independence from blood pressure level, I studied the time-to-event with the study group stratified by blood pressure level in two ways. Firstly, in four groups, with participants grouped as normotensive and with a low *HyperScore* (NL), normotensive and medium- or high-*HyperScore* (NMH), elevated or hypertensive and low-*HyperScore* (EHL), and elevated or hypertensive and a medium- or high-*HyperScore* (EHMH). Secondly, in nine groups, participants were grouped as normotensive and low-*HyperScore* (NL), normotensive and medium-*HyperScore* (NM), normotensive and high-*HyperScore* (NH), elevated and low-*HyperScore* (EL), elevated and medium-*HyperScore* (EM), elevated and high-*HyperScore* (EH), hypertensive and low-*HyperScore* (HL), hypertensive and medium-*HyperScore* (HM), and hypertensive and high-*HyperScore* (HH).

As a final evaluation of clinical relevance, I compared how effectively *HyperScore* predicts time-to-event, compared to widely available ‘gold standard’ risk models, optimised to predict future events, *e.g.*, QRISK3, ASCVD, Framingham Risk Score, MESA¹⁸⁶, and imaging measures proposed as predictive risk markers, *i.e.*, carotid IMT.

3.3.6 Clinical interpretation of *HyperTrajectory* allocation

To study whether the *HyperTrajectory* provided additional biological and clinical relevance beyond that obtained from the *HyperScore* alone, I examined demographic information, disease outcomes, and imaging phenotypes across the six *HyperTrajectories*. I summarised the trajectories by identifying distinct patterns evident across them. Here, I allocated to each *HyperTrajectory* the predominant characteristic evident, *i.e.*, the major end-organ-related attribute that had the most marked differences at the highest 5% *HyperScores* within that *HyperTrajectory*, which represents the advanced disease state. The criteria used to determine the predominant characteristic were as follows,

- 1) ***Diseases distribution:*** I used ICD-10 codes to determine the proportions of diseases such as cardiac, brain, vasculature, pulmonary, liver, kidney, and metabolic diseases, alongside circulatory-related mortality, across the *HyperTrajectories*.
- 2) ***Time-to-event:*** I looked for event-free probabilities related to the aforementioned diseases affecting multiple organs within each *HyperTrajectory*.
- 3) ***Phenotypical changes:*** I identified significantly different imaging variables using a multivariate linear mixed effects model adjusted for age and sex. These phenotypes were then examined to validate outcome-related findings and to determine potential associations with the targeted organ-related changes.

3.3.7 Statistical and regression analysis

If not otherwise indicated, all statistical tests in this study were carried out using the analysis of variance (ANOVA) with a significance level of 0.05. To evaluate phenotypic patterns, I performed regression analysis using locally estimated scatterplot smoothing (LOESS) fitting and compared it with linear fitting. Error percentage was provided as the overall ratio of RMSE between data points and the fitted line. Linearity analysis was performed by evaluating the coefficient of determination (r^2) provided by the fitted line to the data points. Moreover, the p -value shows the significance of this fitting relation. I performed statistical analysis over the fitted line coefficients to compare models during the validation of phenotypic patterns. The statistical analysis of variables predicting trajectories was performed using a multivariate linear mixed effects model (LME) adjusted for age and sex.

3.4 Results

3.4.1 Dataset for model development

Following curation and pre-processing steps, 27,099 individuals from the imaging sub-study of UK Biobank were included in the final model development dataset. Individuals with known cardiovascular or cerebrovascular events prior to the imaging visit were excluded. The dataset included 14,467 females (53%) and 12,632 males (47%), with a mean age of 63.3 ± 7.5 years, and the majority were from a white ethnic background (26,254 – 97%). Of these participants, 3,450 (13%) had pre-existing hypertension prior to their imaging visit, 5,823 (21%) participants were taking anti-hypertensive agents,

and 5,563 (21%) had cholesterol-lowering medications. Detailed demographic characteristics are shown in **Table 3.1**.

Table 3.1 | Baseline characteristics of the UK Biobank cohort included in this study.

Variable	Overall (N = 27,099)				
	Normotensive (N = 1,386 – 5.12%)	Elevated (N = 21,544 – 79.50%)	Hypertensive (N = 4,169 – 15.38%)	Combined	p-value
Age, years	58.19±6.59	62.91±7.39	66.84±6.72	63.27±7.48	<0.001^{*o+}
BMI, kg/m²	23.92±3.43	26.26±3.95	27.06±4.00	26.26±3.98	<0.001^{*o+}
Cholesterol level, mmol/L	5.43±1.00	5.74±1.06	5.85±1.11	5.74±1.07	<0.001^{*o+}
Diabetic, n (%)	11 (0.79%)	1,042 (4.84%)	229 (5.49%)	1,282 (4.73%)	<0.001^{*o}
Hypertension medication, n (%) ----- a. Cholesterol medication b. Other c. Unreported	0 (0.00%) ----- a. 62 (4.47%) b. 133 (9.60%) c. 1,198 (86.44%)	4,392 (20.39%) ----- a. 4,356 (20.22%) b. 1,053 (4.89%) c. 14,302 (66.39%)	1,431 (34.32%) ----- a. 1,145 (27.46%) b. 138 (3.31%) c. 2,217 (53.18%)	5,823 (21.49%) ----- a. 5,563 (20.53%) b. 1,324 (4.89%) c. 17,717 (65.38%)	<0.001^{*o+}
White ethnicity, n (%) ----- a. Asian or Asian British b. Black or black British c. Chinese d. Other mixed e. Unreported	1,327 (95.74%) ----- a. 17 (1.23%) b. 8 (0.58%) c. 6 (0.43%) d. 10 (0.72%) e. 18 (1.30%)	20,861 (96.83%) ----- a. 210 (0.97%) b. 119 (0.55%) c. 63 (0.29%) d. 100 (0.46%) e. 191 (0.89%)	4,066 (97.53%) ----- a. 39 (0.94%) b. 23 (0.55%) c. 11 (0.26%) d. 8 (0.19%) e. 22 (0.53%)	26,254 (96.88%) ----- a. 266 (0.98%) b. 150 (0.55%) c. 80 (0.30%) d. 118 (0.44%) e. 231 (0.85%)	<0.001^{*o+}
Currently smoker, n (%) ----- a. Previous b. Never c. Unreported	55 (3.97%) ----- a. 416 (30.01%) b. 906 (65.37%) c. 9 (0.65%)	767 (3.56%) ----- a. 7,001 (32.50%) b. 13,628 (63.26%) c. 148 (0.69%)	95 (2.28%) ----- a. 7,001 (32.50%) b. 2,499 (59.94%) c. 47 (1.13%)	917 (3.38%) ----- a. 1,528 (36.65%) b. 17,033 (62.85%) c. 204 (0.75%)	<0.001^{*o+}
Heavy alcohol drinker, n (%) ----- a. Moderate alcohol drinker b. Unreported	926 (66.81%) ----- a. 454 (32.76%) b. 6 (0.43%)	15,421 (71.58%) ----- a. 6,021 (27.95%) b. 102 (0.47%)	3,034 (72.78%) ----- a. 1,103 (26.46%) b. 32 (0.77%)	19,381 (71.52%) ----- a. 7,578 (27.96%) b. 140 (0.52%)	<0.001^{*o}

Continuous variables are represented as Mean±SD. Categorical variables are represented as n (%). Bold p-values show statistically significant differences ($p < 0.05$) between all three blood pressure groups using one-way analysis of variance (ANOVA) for continuous variables. Categorical variables were statistically analysed by fitting a generalised linear model. BMI = Body mass index; *: significance between normotensive and elevated; o: significance between normotensive and hypertensive; +: significance between elevated and hypertensive.

3.4.2 Development of end-organ model

The development pipeline of the *HyTwin* model is shown in **Figure 3.3**. Participants underwent brain MR, cardiac MR, whole body MR, carotid ultrasound, dual-energy X-ray absorptiometry (DEXA), electrocardiography (ECG), body composition, spirometry, and blood biochemistry. Following feature selection to remove redundancy, 566 variables were taken forward into the contrastive modelling. A total of 1,386 participants (5.12%) had no end-organ damage and are referred to as the “N” group; 4,169 (15.38%) were deemed likely to have significant hypertensive end-organ changes, *i.e.*, the “H” group. This left 21,544 individuals (79.50%) grouped within the intermediate “E” group.

When comparing the N and H groups, 189 variables were associated with group membership at $p < 0.05$, where the top three major contributions came from the brain MR biomarkers (66.9% of contribution), body composition measures (9.8% of contribution), and carotid ultrasound measures (8.6% of contribution). The top 15 measurements that contributed the most to the modelling included white matter hyperintensities (7.29% of contribution), trunk fat mass (1.67%), and mean carotid intima-media thickness (IMT – 1.37%). All imaging and non-imaging measurements were then analysed using the cPCA approach, which highlighted the difference between the two groups by generating a higher contrast between their covariance matrices through a weighted subtraction of the “N” group from the “H” group. As a result of this analysis, six predominant patterns of disease trajectory (*HyperTrajectories 1 to 6*) were discovered with associated severity of end-organ change (*HyperScore from 0 to 1*).

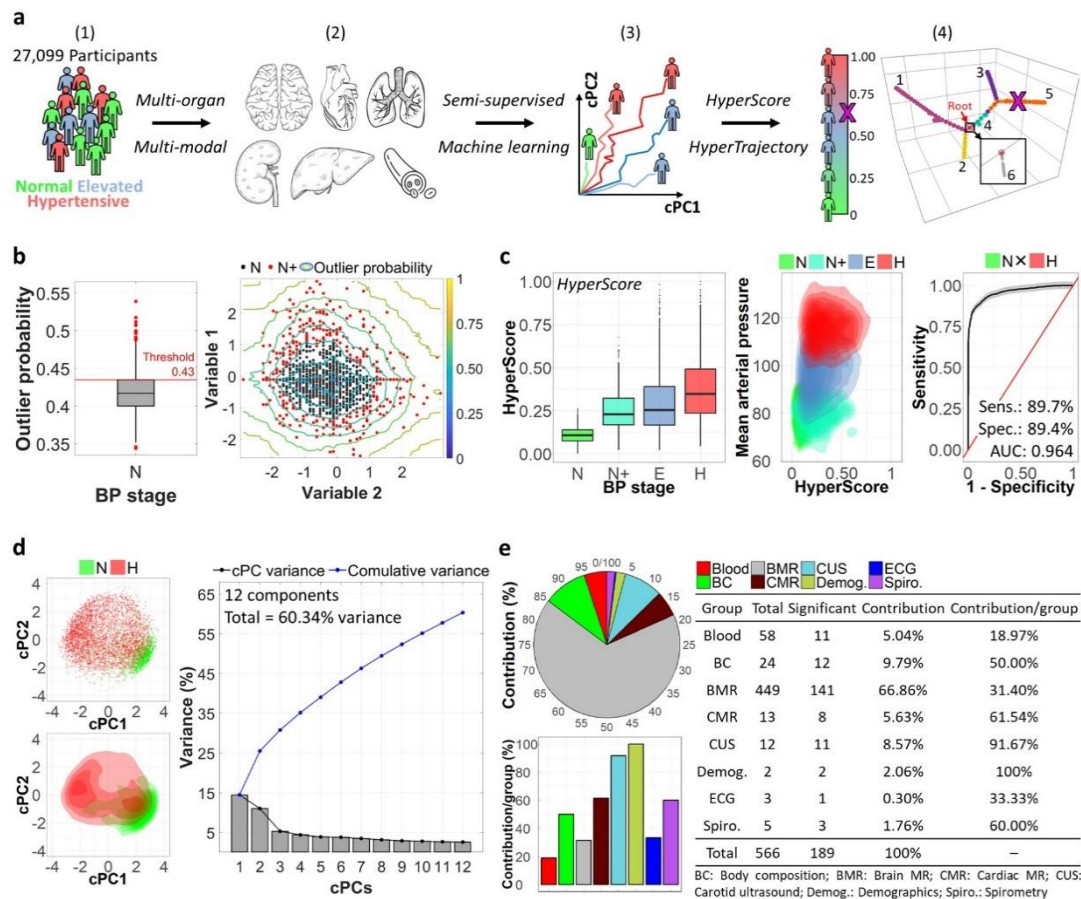


Figure 3.3 | Machine learning-based multi-organ and multi-modality modelling of the end-organ changes associated with hypertension. **a**, Overview of the modelling approach that starts by (1) enrolment of participants with different blood pressure stages at the UK Biobank imaging study (2015-ongoing) to (2) obtain their variables of multiple organs using different modalities. Then, (3) data passes through a semi-supervised contrastive trajectory inference (cTI) approach using contrastive principal component analysis (cPCA) and minimum spanning tree (MST) which (4) aid in assigning a numerical value estimating the disease-state risk for any individual, *HyperScore*, and locating the individual on a constructed disease pathways map, *HyperTrajectory*. **b**, Process of identifying outliers (N+) in the normal (N) group who are above the third quartile of probabilities (≥ 0.43) assigned using isolation forest and clinical variables. **c**, Estimating *HyperScore* and grouping patients relative to their blood pressure level and stage, *i.e.*, N, N+, intermediate elevated (E), and hypertensive (H). The discrimination between N and H showed an excellent (96.4%) area under the receiver operating characteristics (AUC) curve with a sensitivity and specificity of 89.7% and 89.4%, respectively. **d**, Interpretation of cPCA modelling for the N and H groups in the cPC1 and cPC2 space with the total variance (60.34%) explained by the selected 12 components. **e**, Clinical variables' importance relative to the modelling grouped by modality.

3.4.3 Internal validation

3.4.3.1 Stability of *HyperScore* estimation

To evaluate the stability of *HyperScore*, a stratified 10-fold cross-validation was used and showed no differences between scores across folds based on a priori acceptance criteria, with a mean root mean square error (MRMSE) of 0.11 ± 0.01 . Moreover, the predictability of testing sets at each fold had a MRMSE of 0.10 ± 0.08 . In both analyses, there was an association between the absolute level of *HyperScore* for an individual and the variance in their *HyperScore* across folds (**Figure 3.4a**), with higher *HyperScores* having wider variance.

I then compared *HyperScores* assigned to individuals in a random 50% of the data, based on a model built from the other 50% of the dataset ($n = 13,055$) (Model A), to *HyperScores* attributed to an individual when that set of data was used to develop the

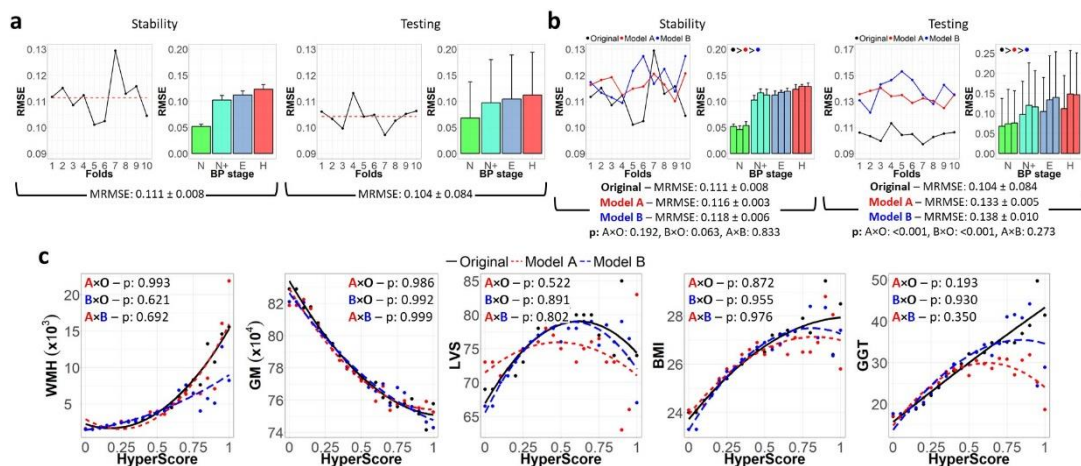


Figure 3.4 | Internal validation of the modelling approach and *HyperScore* estimation performance. **a**, Evaluating model stability and testing in predicting *HyperScore* when changing data used to generate the model. A stratified 10-fold cross-validation scheme was followed, and difference in *HyperScore* between the full dataset and per-fold subsets was calculated using the mean root mean square error (MRMSE) to find stability. Model error in predicting *HyperScore* for testing subsets at each fold was calculated to find testing performance as well. **b**, Comparing stability and testing performance when generating two additional models (A and B) using 50% ($n = 13,550$) of the overall participants cohort in each. **c**, Calculating the statistical difference between the three models (original, A, and B) for several phenotypic patterns. Non-significance indicates similar phenotypic patterns across the models. Points within plots represent median value of participants at each 0.025 *HyperScore* step.

model. I then tried this again, building the model on the alternative 50% and attributing scores to the original 50% (Model B). MRMSE values (**Figure 3.4b**) did not significantly differ between the modelled and predicted *HyperScores* for individuals in both approaches (A: 0.13 ± 0.01 , B: 0.14 ± 0.01). I repeated this stability testing, splitting the dataset into two halves based on the time of the imaging visit, *i.e.*, set 1 (2015 to 2018) and set 2 (2018 to 2021). Again, both models had low MRMSE values of 0.16 ± 0.08 and 0.17 ± 0.07 , respectively, whether compared to *HyperScores* generated in the combined model or when using set 1 to build the model and set 2 as the predicting dataset. In this case, the MRMSE value was 0.14 ± 0.03 and the opposite setting resulted in a MRMSE of 0.16 ± 0.05 .

Confidence-based performance estimation (CBPE) analysis, as an alternative recognised approach to consider performance in the absence of ground truth for end-organ damage, considered the ability of generated *HyperScores* to differentiate between those with no organ damage (N) and advanced organ disease (H). This test showed an optimal *HyperScore* threshold of $T = 0.28$ to differentiate groups, having the highest sensitivity

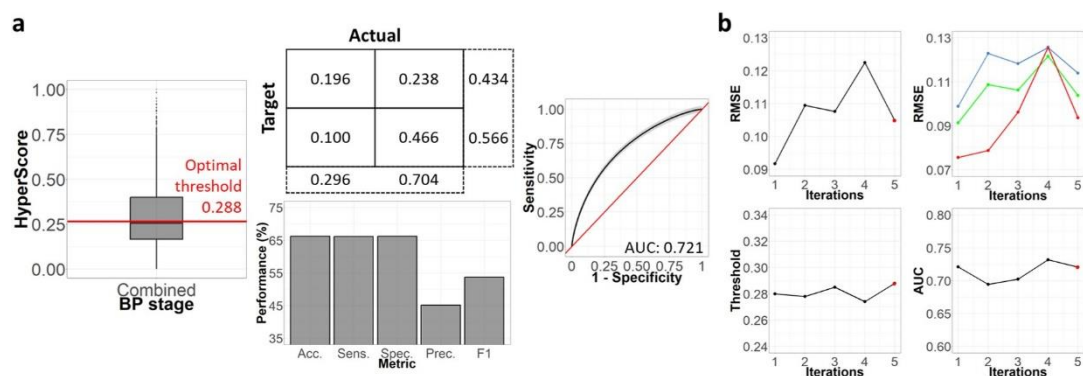


Figure 3.5 | Additional internal validation of the *HyperScore* using confidence-based performance estimation (CBPE) analysis. a, Validation in absence of ground truth using CBPE. *HyperScore* values were used as probabilities towards being hypertensive (0 to 1 scale), and the optimal threshold was calculated to ensure maximum area under the receiver operating characteristics (ROC) curve. The confusion matrix was formed by summation of true/false and positive/negative probabilities relative to the selected threshold. **b**, Data drifting analysis through *HyperScore* prediction error on a testing set after continuously adding data to the trained model on each iteration (top row). The same analysis was applied on the confidence-based metrics (threshold and area under the curve – AUC) when iteratively adding data to the training set.

and specificity levels (66.2% each) with true positive and true negative ratios of 0.196 and 0.466, respectively, and an overall AUC of 0.72 (**Figure 3.5a**). This information was then used in the data drifting analysis to understand stability. The drifting effect was minimal when considering the impact on the optimal threshold for differentiation of the N and H groups (± 0.007) and AUC (± 0.023) values in the CBPE during the continuous addition of data to the model (**Figure 3.5b**). Finally, evaluation of MRMSE demonstrated the measure remained low at 0.11 ± 0.11 , when iteratively adding data to the training dataset and comparing to the test dataset in the data drifting analysis, consistent with a stable model.

3.4.3.2 Stability of *HyperTrajectory* allocation

The stability of the re-generation of a similar trajectory map based on 10-fold cross validation with stability considered on a per-trajectory basis demonstrated the highest stability of regeneration for trajectories 1 and 5, with sensitivity of $64 \pm 5.8\%$ and $74 \pm 6.8\%$, respectively (**Figure 3.6a**). The highest precision values were achieved by trajectories 1 and 2, reaching up to $68.96 \pm 3.85\%$ and $61.98 \pm 13.80\%$, respectively. On the other hand, trajectory 6 could not be consistently reproduced on cross fold testing.

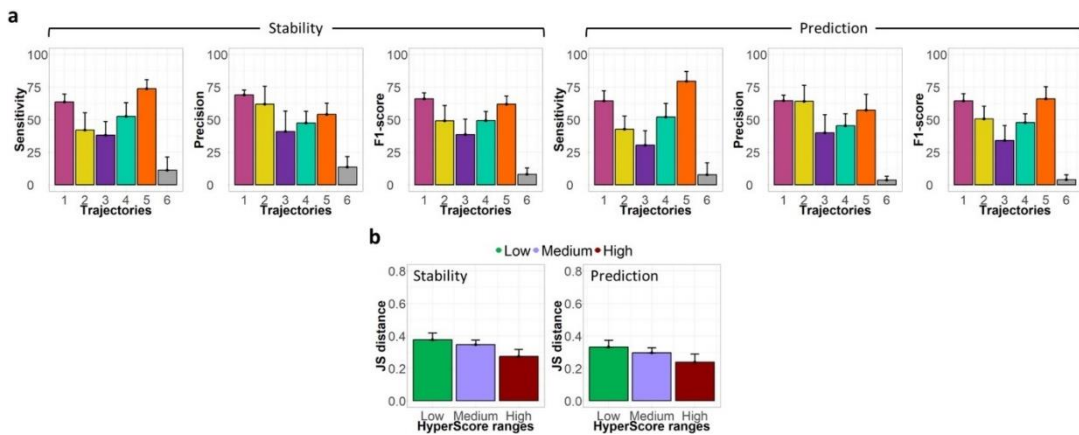


Figure 3.6 | Internal validation performance of the HyperTrajectory. a, Per-trajectory average sensitivity, precision, and F1-score for the stability and prediction with the standard deviation (SD) error bar. The ground-truth is selected as the HyperTrajectory map of the original full model. **b**, Jensen-Shannon distance for stability and prediction when categorising participants based on *HyperScore*, i.e., low-, medium-, and high-risk.

This trajectory had the smallest number of participants with a low average *HyperScore*, suggesting minimal variation in end-organ change. Similarly, the prediction of the testing set across the folds using the regenerated *HyperTrajectory* at each fold demonstrated that trajectories 1 and 5 had the highest sensitivity measures at $64\pm 7.6\%$ and $79\pm 7.2\%$, respectively, whereas trajectory 6 again could not be identified reproducibly. Trajectories 1 and 2 had the highest precision values of $64.55\pm 4.24\%$ and $64.15\pm 12.39\%$, respectively.

The evaluation of performance using the JS distance metric to compare distribution of probabilities for allocation of participants into groups based on *HyperScore* ranges, *i.e.*, low-, medium-, and high-*HyperScore* (**Figure 3.6b**) showed the high-risk group always had the lowest average distance across folds compared to the full model, with 0.27 ± 0.04 and 0.24 ± 0.05 for the stability and prediction tests, respectively. On the other hand, the highest average distance values were found for the low-*HyperScore* group for stability and prediction with 0.38 ± 0.04 and 0.33 ± 0.04 , respectively.

3.4.4 Clinical relevance of *HyperScore*

3.4.4.1 Association with demographics and organ phenotypes

Associations of *HyperScore* with pre-specified individual variables previously reported to characterise end-organ damage in hypertension demonstrated the model described expected phenotypical variations (**Figure 3.7**). Generally, males had higher estimated *HyperScore* than females ($p < 0.001$). Similarly, in regression analysis, a higher *HyperScore* was associated with more white matter hyperintensities ($r^2: 0.61, p < 0.001$) and less grey matter volume ($r^2: 0.81, p < 0.001$). Similarly, a positive association was observed with body mass index (BMI — $r^2: 0.62, p < 0.001$), mean carotid IMT ($r^2: 0.69, p < 0.001$), ventricular rate ($r^2: 0.42, p < 0.001$), gamma-glutamyl transferase (GGT — $r^2:$

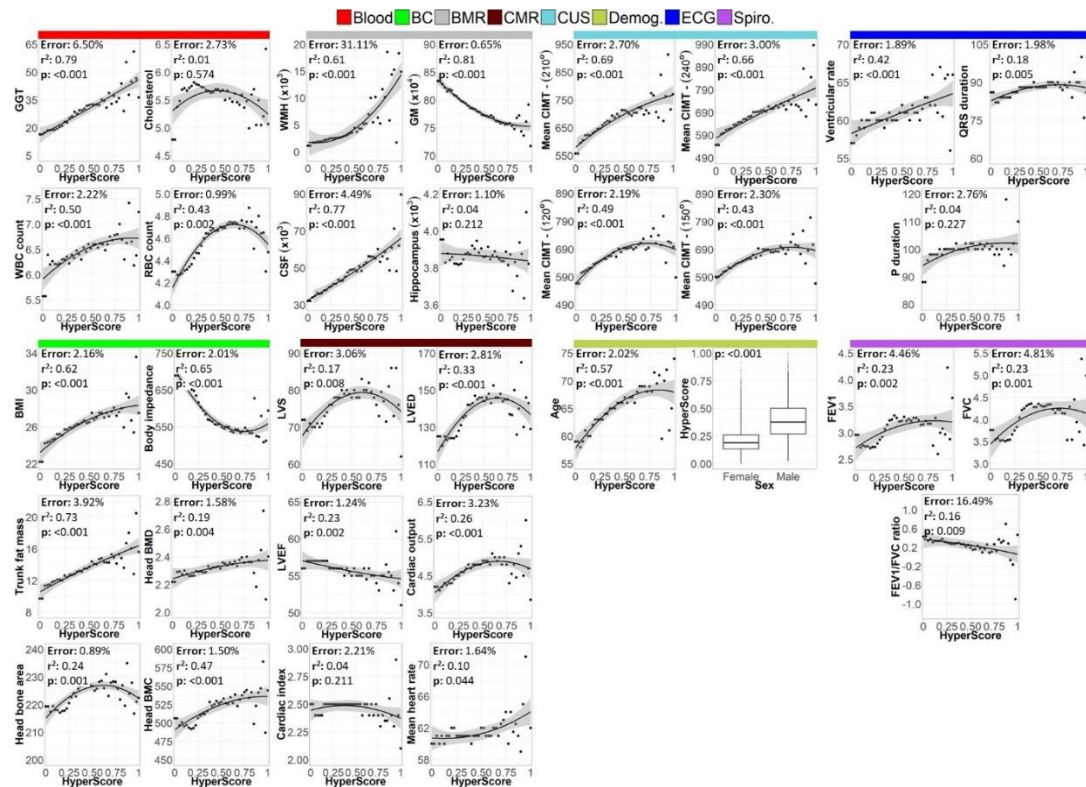


Figure 3.7 | Association between end-organ phenotypes and hypertension progression. Phenotypic patterns with respect to *HyperScore* showing examples from every modality. Percentage error was calculated between the fitted line using locally estimated scatterplot smoothing (LOESS) and points, whereas the r^2 and p-value were obtained by fitting a linear model. Low r^2 and non-significant p-value indicate that there is no linear relation. Points within plots represent median value of participants at each 0.025 *HyperScore* step.

0.79, $p < 0.001$), and age (r^2 : 0.57, $p < 0.001$); whereas a negative association was observed for body impedance (r^2 : 0.65, $p < 0.001$). A non-linear relationship was observed with both left ventricular stroke (LVS) volume and left ventricular end-diastolic (LVED) volume with an initial positive relationship, with an inflection point and a switch to an inverse relationship at high *HyperScore* ranges consistent with expected pathological change.

3.4.4.2 Time-to-event analysis

I identified a total of 1,127 participants (Female: 375 – 33.27%, Male: 752 – 66.73%) who had suffered from adverse events related to circulatory system diseases, *i.e.*, cardiac, cerebral, or vascular disorders ($n = 1,021$), or died due to circulatory system diseases (n

= 106), since they had attended their first imaging visit (**Figure 3.8a**). Time-to-event analysis and cumulative hazard estimation performed using the Kaplan-Meier method with participants grouped by *HyperScore* stage as low, medium or high *HyperScore* (**Figure 3.8b**) confirmed a higher *HyperScore* associated with a shorter time to first event, *e.g.*, myocardial infarction or stroke. When compared to a similar analysis with the study group separated based on their distribution of blood pressure into low, intermediate, or high blood pressure, while there were differences in survival odds between the high-*HyperScore* groups and both the low ($p: <0.001$, c^2 : 249.37) and medium ($p: 0.001$, c^2 : 56.90) *HyperScore* groups, there were only differences in time-to-event between the low and high blood pressure groups. In terms of the cumulative hazard, the high-*HyperScore* and medium-*HyperScore* groups had hazard ratio (HR) values of 3.35 [2.76—4.07] and 2.11 [1.84—2.42], respectively.

When I divided participants into four groups taking into account both their *HyperScore* (low or medium/high) and blood pressure level (normal or elevated) (**Figure 3.8c**) I found that for those with elevated or hypertensive blood pressures, *HyperScore* contributed additional stratification with a significant difference in odds of time-to-event ($p: <0.001$) observed between participants with low *HyperScore* compared to participants with medium-high *HyperScore* at each blood pressure stage. Repeating this analysis, but with participants divided into nine groups, *i.e.*, *HyperScore* and blood pressure divided into three groups each based on level, confirmed these findings with participants with medium and high *HyperScore* having progressively shorter time-to-event compared to those with low *HyperScore* at each blood pressure stage.

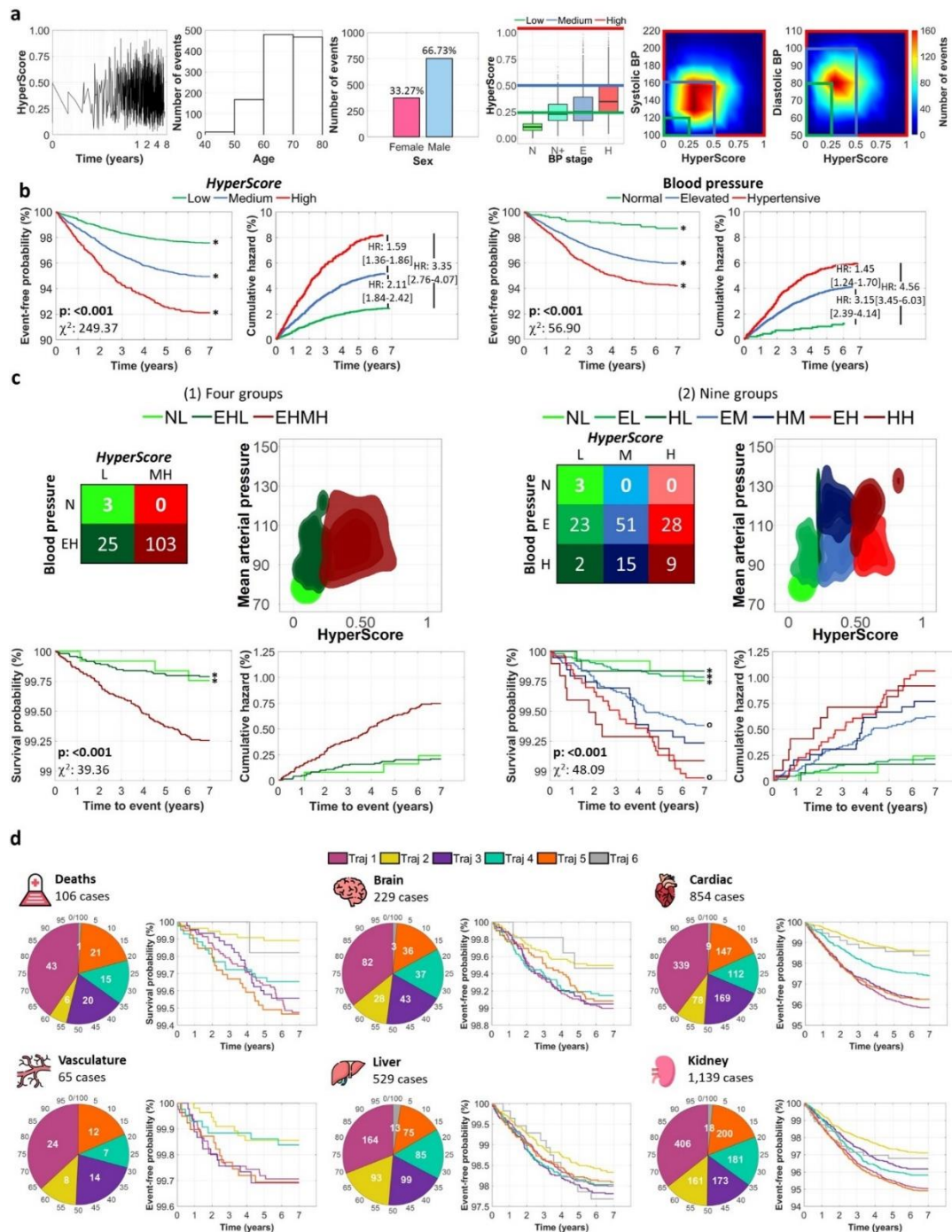


Figure 3.8 | Prediction of adverse outcomes and analysis of survival. a, Overall representation of the 1,127 participants who suffered from events (circulatory deaths, international classification of disease version 10 (ICD-10) circulatory diseases) with stratified analysis based on *HyperScore* stages. **b**, Analysis of survival probability and cumulative hazard for these participants when grouped based on *HyperScore* stages (low, medium, and high) and blood pressure stages (normal, elevated, and hypertensive). * indicates significance against all other curves. **c**, Analysis when grouping participants into (1) four and (2) nine groups based on combined blood pressure and *HyperScore* stages as normal (N) or low (L), elevated (E) or medium (M), and hypertensive (H) or high (H). **d**, Analysis of time-to-event in *HyperTrajectory* with the total number of cases in every trajectory.

3.4.4.3 Comparison to conventional risk prediction clinical tools

Comparison of our imaging-based metric of current end-organ disease state, *HyperScore*, as a predictor of time-to-event, compared to clinical scoring systems, *i.e.*, QRISK3, ACC/AHA, Framingham, MESA and the single imaging-based measure of vascular disease state, carotid IMT, confirmed *HyperScore* had comparable predictive ability for future events, in particular, circulatory deaths across all six scores (**Figure 3.9**). In addition, *HyperScore* effectively predicted time-to-event for a range of other

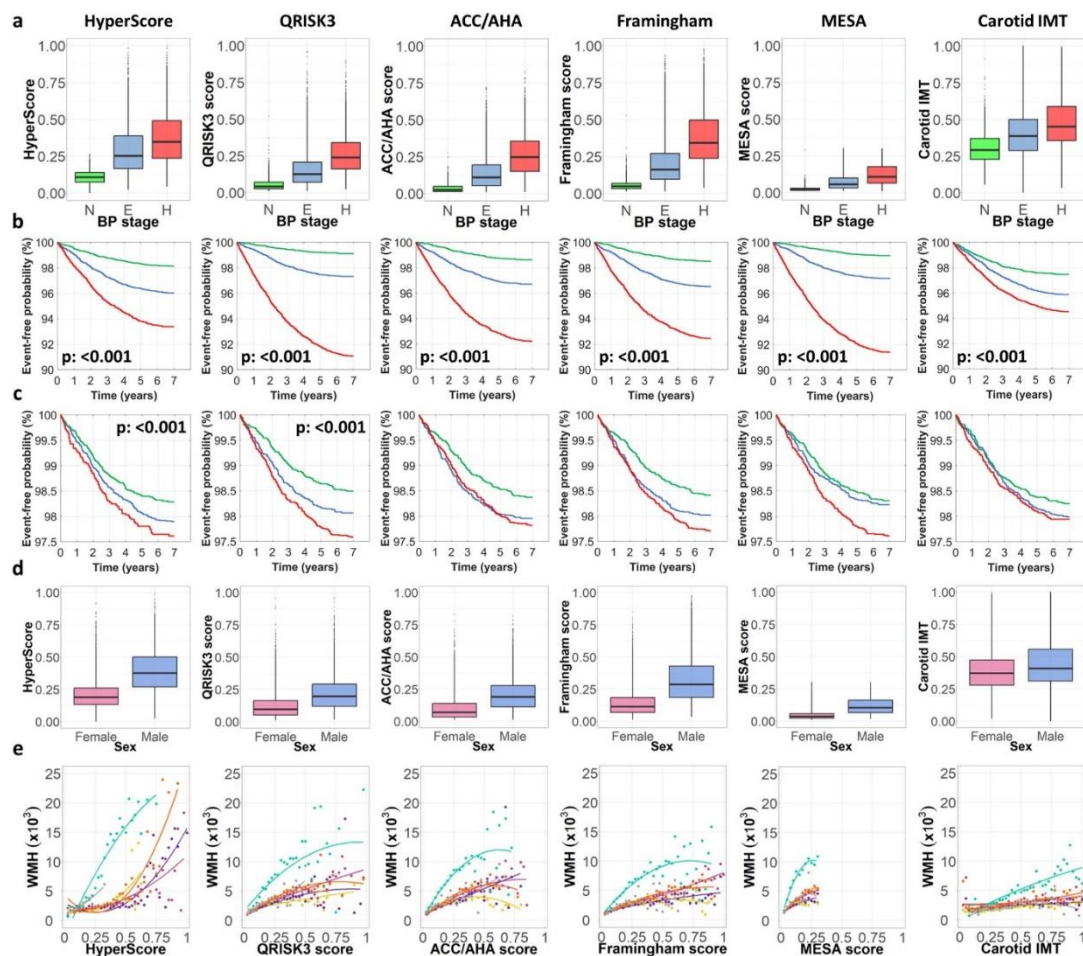


Figure 3.9 | Comparison of *HyperScore* and *HyperTrajectory* relative to conventional and clinically valid scores. **a**, Scores grouped based on blood pressure stages. **b**, Survival curves of circulatory deaths and diseases when grouping scores into three quantiles based on distribution of scores. **c**, Survival curves for a selected disease, namely liver disorders. **d**, Sex-based analysis of the generated scores with respect to male and female individuals. **e**, A selected example of the phenotypical pattern of progression for white matter hyperintensities with scores at different *HyperTrajectories*. Points within plots represent median value of participants at each 0.025 *HyperScore* step.

individual clinical outcomes relevant to hypertension, whereas performance of established risk scores and carotid IMT was less consistent across conditions.

3.4.5 Clinical relevance of *HyperTrajectory*

3.4.5.1 Variation in end-organ phenotypes by trajectory

Individuals within the six trajectories had similar mean ages (**Figure 3.10**). Three trajectories, namely trajectories 2, 4, and 6, were predominantly female (>79%), while the other three trajectories were comprised of relatively more males. There were no significant differences in mean systolic, diastolic, pulse or mean arterial pressure across the six trajectories.

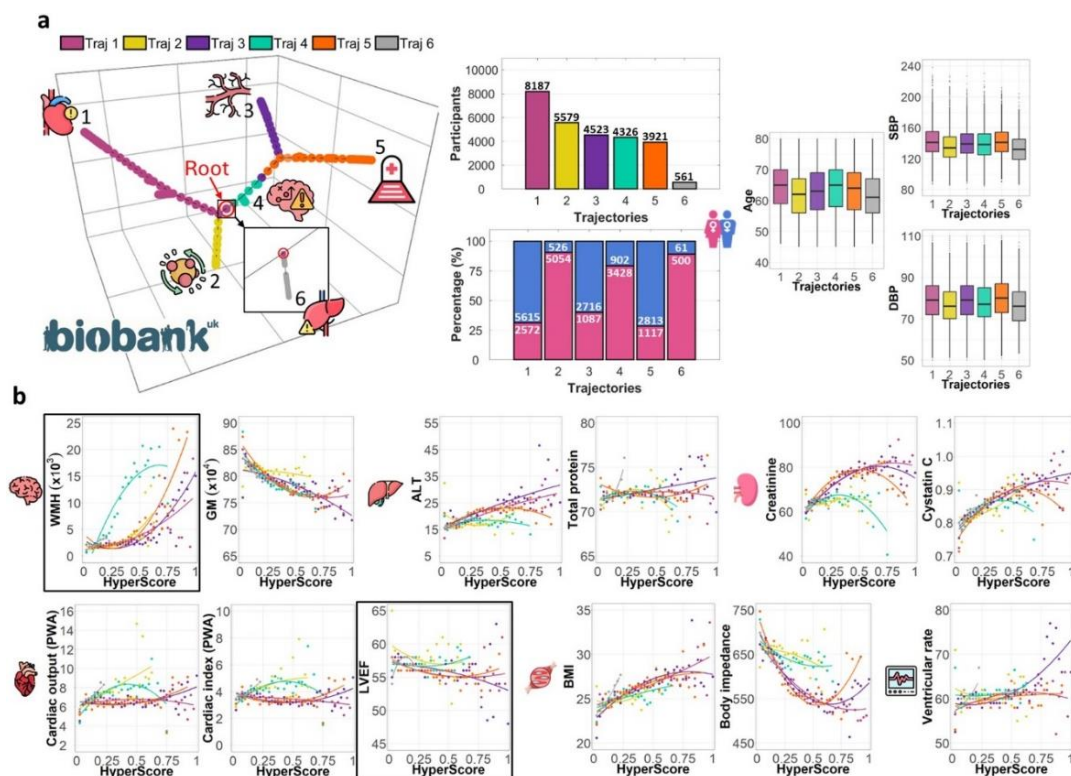


Figure 3.10 | Characterisation of the *HyperTrajectory* in the UK Biobank. a, The trajectory map was created using the minimum spanning tree (MST) and visualised in the embedded Laplacian Eigenspace. Six major trajectories were formed following a systematic approach and their demographics were analysed including sex, age, and blood pressure. Each trajectory was characterised by the increasingly occurring outcomes and organ-specific survival rate. **b**, Progression phenotypes in the *HyperTrajectory*. Each progression phenotype illustrates the change in the per-organ function with respect to *HyperScore*, i.e., throughout the trajectory path. Points within plots represent median value of participants at each 0.025 *HyperScore* step.

Visual inspection of the relative position and branching of *HyperTrajectories*, displayed as a three-dimensional representation of the six trajectories, identified that trajectories 3 and 5 share a common pathway at lower levels of *HyperScore* and then separate as score increases. As expected, visual comparison of individual imaging biomarkers showed similar levels at the start of the *HyperTrajectories* with divergence of phenotypes between *HyperTrajectories* with increasing distance from the root node. This was particularly evident for white matter hyperintensities, that demonstrated a gradual increase in the male-dominant trajectories, 1, 3, and 5, and a much more accelerated increase in female-dominant trajectories 4 and 6.

A multi-variate linear mixed-effects model adjusted for age and sex identified imaging biomarkers that differed significantly by trajectories. Trajectory 1 had the lowest cardiac index and cardiac output during pulse wave analysis (PWA), reduced left ventricular ejection fraction on imaging (LVEF), and the highest levels of creatinine, cystatin C, glucose, and glycated haemoglobin (HbA1c). Participants in trajectory 2 had relatively normal imaging-based health profiles but high levels of low-density lipoprotein (LDL), high-density lipoprotein (HDL), and total cholesterol. Trajectory 3 had an inflammatory profile of elevated white blood cells and reticulocytes. Trajectory 4 had high white matter hyperintensities and low grey matter volume. Trajectory 5 had low cardiac index, high creatinine, and high cystatin C. Trajectory 6 had a small size, so I did not perform detailed analysis but did identify higher levels of body fat, alcohol intake and cholesterol.

3.4.5.2 Variation in events and time-to-event across

HyperTrajectories

Analysis of future outcomes for individuals within each *HyperTrajectory* (**Figure 3.8d and Table 3.2**) showed striking similarities between disease outcomes and phenotype

HyperTrajectory				
Trajectory 1	Trajectory 2	Trajectory 3	Trajectory 4	Trajectory 5
 Cardiac	 Lipoprotein	 Atherothrombotic	 Brain	 Mortality
Trajectory	Advanced disease state – top 5%			
	Total participants	Predominant characteristic	Major outcomes	Key phenotype
1	N = 410	Cardiac	- Lowest cardiac event-free probability - Second-lowest renal event-free probability	- Lowest COp/Clp - Lowest LVEF - Highest creatinine/cystatin - Highest BMI - Highest triglycerides - Most diabetic - Highest HbA1c/glucose
2	N = 279	Lipoprotein	- Highest proportion of metabolic outcomes (disorders of lipoprotein—lipidaemias)	- Highest LDL/HDL - Highest cholesterol - Lowest BMI/BFM
3	N = 227	Atherothrombotic	- Lowest vasculature event-free probability (thoracic aortic aneurysm/atherosclerosis) - Lowest pulmonary event-free probability (pulmonary embolism) - Second-lowest brain event-free probability	- Highest DBP - Highest hypertension medication - Highest cholesterol medication - High WMH - Low grey matter
4	N = 217	Brain	- Highest proportion of brain outcomes (cerebral infarction/aneurysm)	- Highest WMH
5	N = 197	Mortality	- Lowest circulatory death survival rate - Lowest renal event-free probability - Highest proportion of renal outcomes - Highest proportion of cardiac outcomes	- Highest SBP - Lowest CI - Second-lowest COp/Clp - Second-lowest LVEF - High creatinine/cystatin - High WMH - Lowest grey matter

Figure 3.11 | Overall summary of the formed *HyperTrajectories* in the UK Biobank highlighting their *HyperScore* range, major characteristics, phenotype, and outcomes. Participants were assessed based on their advanced disease state (highest 5% of *HyperScores*) at each trajectory. Predominant characteristics were derived from Extended Data Table 1 as well as survival analysis. Trajectory 6 was not included in the summary due to having low number of participants. SBP = Systolic blood pressure; DBP = Diastolic blood pressure; BP = blood pressure; WMH = White matter hyperintensity; COp = Cardiac output during pulse wave analysis (PWA); Clp = Cardiac index during PWA; LVEF = Left ventricular ejection fraction; BMI = Body mass index; BFM = Body fat mass; LDL = Low-density lipoprotein; HDL = High-density lipoprotein; HbA1c = Glycated haemoglobin; WBC = White blood cells count; RTIC = Reticulocytes count.

changes identified for participants in each *HyperTrajectory* (Figure 3.11). Although all trajectories had a range of events, individuals within Trajectory 1, who had phenotypic changes in cardiac measures, had the highest proportion of cardiac outcomes relative to other trajectories. Participants in trajectory 2, who were identified with changes in cholesterol level, had significantly lower mortality relative to all other trajectories ($p < 0.001$, $c^2 > 10$) but higher rates for most non-mortality events with the highest proportion of metabolic outcomes, mostly lipoprotein disorders. Individuals in trajectory

3, who had changes in inflammatory markers, had relatively more vascular events including peripheral vascular disease and pulmonary thromboembolic disease. Trajectory 4, which had high white matter changes, had the highest proportion of brain disorders. Trajectory 5 individuals, who had cardiac and renal differences, had the highest mortality with high event-free rates for kidney outcomes. Trajectory 6, which had a limited sample but demonstrated differences in body fat and alcohol, had the worst event-free rates for liver diseases. Detailed characteristics are provided in **Table 3.2**.

Table 3.2 | Characteristics of the *HyperTrajectory* in the advanced disease state (top 5% *HyperScores*) in the UK Biobank cohort.

<i>HyperTrajectory</i>	1 (N = 410)	2 (N = 279)	3 (N = 227)	4 (N = 217)	5 (N = 197)	6 (N = 29)
<i>HyperScore</i>	0.75±0.06	0.37±0.07	0.70±0.10	0.52±0.05	0.70±0.06	0.19±0.01
Demographic information						
Female	32 (7.81%)	212 (75.99%)	15 (6.61%)	149 (68.66%)	26 (13.20%)	24 (83.76%)
Age, years	67.42±7.06	63.85±6.68	66.74±7.32	68.53±6.01	69.32±5.80	64.07±8.61
SBP, mmHg	150.78±20.42	149.44±23.97	150.48±20.02	147.63±18.34	151.01±19.08	146.31±23.13
DBP, mmHg	82.25±10.23	81.46±11.49	83.21±10.99	79.64±9.98	81.09±10.59	82.03±10.81
PP, mmHg	68.52±17.07	67.98±20.03	67.27±16.08	67.99±16.13	69.92±15.49	64.28±20.08
MAP, mmHg	105.10±12.00	104.12±13.80	105.63±12.51	102.30±10.98	104.40±11.95	103.46±12.91
Current smoker	16 (3.90%)	7 (2.51%)	3 (1.32%)	7 (3.23%)	6 (3.05%)	0 (0.00%)
Diabetic	61 (14.88%)	11 (3.94%)	25 (11.01%)	8 (3.69%)	17 (8.63%)	0 (0.00%)
Heavy alcohol drinker	321 (78.29%)	194 (69.53%)	178 (78.41%)	142 (65.44%)	154 (78.17%)	23 (79.31%)
Hypertension Medication	166 (40.49%)	62 (22.22%)	100 (44.05%)	74 (34.10%)	78 (39.59%)	24 (24.14%)
Cholesterol Medication	153 (37.32%)	51 (18.28%)	102 (44.93%)	68 (31.34%)	78 (39.59%)	2 (6.90%)
Imaging/non-imaging measurements						
WMH, $\times 10^3 \text{ mm}^3$	9.57±6.73	3.74±3.31	8.29±6.88	17.16±4.85	10.94±7.30	5.12±3.69
GM, $\times 10^4 \text{ mm}^3$	75.92±39.97	80.35±4.37	76.26±4.38	77.63±3.90	75.55±4.10	79.84±4.78
CO, L/min	4.97±1.16	4.50±1.21	5.15±1.09	4.52±1.09	4.73±0.98	4.69±1.34
CI, L/min/m ²	2.46±0.53	2.49±0.58	2.57±0.54	2.49±0.53	2.42±0.50	2.57±0.59
CO _p , L/min	6.94±2.14	8.23±2.80	7.11±2.02	8.06±2.68	6.97±2.16	8.60±2.23
CI _p , L/min/m ²	3.47±1.21	4.64±1.73	3.57±1.12	4.55±1.76	3.61±1.27	4.61±1.03
LVEF, %	54.17±5.98	56.97±6.12	55.23±5.56	56.62±6.09	54.55±5.75	56.96±5.97
VR, bpm	63.00±10.32	61.53±9.53	64.63±10.57	62.33±9.14	60.76±10.53	61.96±9.93

BMI, kg/m²	28.47±3.79	26.14±4.44	28.02±3.95	26.52±4.32	27.28±3.60	28.17±4.24
BFM, kg	25.21±7.54	22.76±7.17	23.86±8.07	24.50±8.57	23.43±6.94	27.82±8.38
Cholesterol, mmol/L	5.49±1.07	5.97±1.08	5.48±1.06	5.83±1.08	5.58±1.02	6.10±0.97
LDL, mmol/L	3.45±0.79	3.71±0.82	3.48±0.82	3.61±0.80	3.52±0.81	3.83±0.70
HDL, mmol/L	1.24±0.29	1.54±0.41	1.27±0.30	1.53±0.39	1.34±0.31	1.56±0.43
Triglycerides, mmol/L	2.04±0.97	1.54±0.77	1.85±0.79	1.54±0.76	1.92±0.96	1.51±0.58
HbA1c, mmol/mol	36.01±4.05	35.18±3.54	35.27±4.16	35.26±3.61	35.66±3.94	34.46±4.29
GLC, mmol/L	5.04±0.62	4.99±0.63	5.02±0.63	5.02±0.64	4.95±0.55	4.85±0.33
ALT, U/L	28.34±11.52	19.83±8.38	29.20±11.76	20.11±8.10	24.93±10.32	22.37±10.12
GGT, U/L	44.04±25.11	27.19±15.81	44.00±25.11	28.61±18.02	40.78±24.62	32.29±25.61
TP, g/L	72.33±3.68	71.87±3.73	73.31±3.78	71.64±3.62	72.05±3.67	73.94±4.11
CRTN, µmol/L	81.46±12.54	67.51±12.63	80.42±11.99	68.25±14.03	79.79±12.58	64.70±10.74
CysC, mg/L	0.93±0.12	0.86±0.12	0.92±0.12	0.87±0.11	0.92±0.12	0.85±0.10
WBC, $\times 10^9$ cells/L	6.85±1.58	6.68±1.55	6.91±1.53	6.46±1.55	6.76±1.55	7.53±1.91
RTIC, $\times 10^{12}$ cells/L	0.068±0.025	0.057±0.025	0.070±0.026	0.056±0.022	0.067±0.026	0.065±0.017
Adverse outcomes						
Death (Circulatory diseases)	4 (3.42%)	0 (0.00%)	4 (3.51%)	2 (3.03%)	2 (2.56%)	0 (0.00%)
Cardiac	16 (13.68%)	9 (14.29%)	20 (17.54%)	12 (18.18%)	15 (19.23%)	0 (0.00%)
Brain	9 (7.69%)	1 (1.59%)	9 (7.90%)	6 (9.09%)	3 (3.84%)	1 (33.33%)
Vasculature	2 (1.71%)	2 (3.18%)	2 (1.75%)	1 (1.52%)	2 (2.56%)	0 (0.00%)
Pulmonary	4 (3.42%)	1 (1.59%)	2 (1.75%)	2 (3.03%)	2 (2.56%)	0 (0.00%)
Liver	6 (5.13%)	8 (12.70%)	9 (7.90%)	7 (10.61%)	8 (10.26%)	1 (33.33%)
Kidney	28 (23.93%)	15 (23.81%)	20 (17.54%)	13 (19.70%)	19 (24.36%)	1 (33.33%)
Metabolic	48 (41.03%)	27 (42.86%)	48 (42.11%)	23 (34.85%)	27 (34.62%)	0 (0.00%)
Overall	117 (28.54%)	63 (22.58%)	114 (50.22%)	66 (30.42%)	78 (39.59%)	3 (10.35%)

Continuous variables are represented as Mean±SD. Categorical variables are represented as n (%). The percentage of organ-specific outcomes is based on the overall number of outcomes. The percentage of the overall outcomes is based on the total number of participants. Bold indicates the trajectory with highest number of outcomes if enough outcomes occur.

3.5 Discussion and conclusion

In this chapter, I have modelled, for the first time, a holistic representation of multi-organ phenotypic changes associated with hypertensive disease development by applying a novel contrastive approach to a large cross-sectional multi-modal imaging dataset. I found six predominant patterns of hypertensive disease progression with distinct demographic, end-organ and outcome profiles. Furthermore, the severity and likely trajectory of end-organ disease can be estimated for any new individual by placing their dataset within this landscape. These findings raise the possibility of personalised treatment plans for primary prevention in those with hypertension, which goes beyond blood pressure level, and takes account of current end-organ disease state.

A key feature of our semi-supervised approach is its ability to discover multiple predominant patterns of phenotypic change in hypertension or *HyperTrajectories*, each reflecting a clinically meaningful pathway of organ damage progression. Unlike previous studies that focused on single-organ phenotyping, *e.g.*, cardiac¹⁸⁷, brain¹⁸⁸, or vasculature¹⁸⁹, our model captures hypertension as a systemic condition. For example, two *HyperTrajectories* with similar cardiac changes may still differ significantly in brain pathology, offering better insight into stratified prediction and treatment. Such a discovery approach, however, is inherently challenging because of a lack of ground truth. To address this, I evaluated model stability within the dataset through multiple internal and external validation steps. Our rationale being that, if the models and trajectories identified are ‘true’ representations of a pathological phenomenon of end-organ change, then the model should remain consistent in any data subset or external dataset examined. As a sense check, I confirmed that known hypertension-related

parameters, such as left ventricular hypertrophy, white matter lesions and grey matter volumes, followed expected patterns across the *HyperScore* range^{190,191}.

The pseudo-temporal interpretation of a cross-sectional dataset with a gradual increase in severity of end-organ damage within the *HyperTrajectory* is fundamentally different from previous clustering approaches aimed at differentiating hypertensive disease states. The latter assumes a similar degree of end-organ change in all individuals¹⁴⁶. The six distinct *HyperTrajectories* emerged from a systematic, data-driven process that took into consideration the functional and structural changes of all major organs simultaneously. The discovery extends the current understanding of how hypertension can affect people in different ways and result in different outcomes. Although all event types occurred to some extent across trajectories, there was proportionally more events associated to the organs that were more affected in each trajectory, *e.g.*, cardiac events were more prevalent in trajectory 1, in which cardiac changes occurred earlier in the trajectory. Furthermore, although cardiac changes were also evident in *HyperTrajectory* 5, both 1 and 5 also displayed renal changes. Whereas *HyperTrajectory* 1 showed higher remodelling in cardiac function, *HyperTrajectory* 5 exhibited severe cardiac and renal pathology, which was linked to higher mortality. Distinct underlying pathologies are likely to be driving these two disease patterns, and this variation may have been difficult to identify if studying cardiac or renal changes in isolation.

There is now an opportunity to explore the distinct drivers and potential treatments for each *HyperTrajectory*, including individual genetic, lifestyle or environmental factors. For example, current treatment guidelines often do not take variation by sex into account¹⁹², whereas I found that women comprised the majority of individuals in *HyperTrajectories* 2, 4, and 6, while *HyperTrajectories* 1, 3, and 5 were predominantly composed of men. Consistent with existing literature, male dominant trajectories had

worse clinical outcomes including cardiac, vasculature, and pulmonary events compared to women¹⁸. However, female dominant trajectories had a sharper increase in white matter lesions with disease progression²⁰. Prior research has shown that incorporating imaging or organ changes into clinical decision-making with patients can significantly improve patient engagement^{172,193}. No such tool exists in hypertension management and further work will be of value to understand whether use of end-organ severity or disease type can impact patient and clinician engagement with management leading to improvements in outcome.

As expected, the multi-organ summary provided by *HyperScore* outperformed carotid intima-media thickness, a single imaging measure¹⁹⁴, in prediction of future events. *HyperScore* also provided prognostic information beyond blood pressure or traditional risk scores, particularly in those with elevated blood pressure. Traditional cardiovascular risk scores¹⁹⁵, which rely on demographic, laboratory, and blood pressure¹⁹⁶, often underperform in individuals with existing risk factors, younger individuals in whom the time horizon for events is less meaningful¹⁹⁷, and for capturing the full spectrum of hypertension-related outcomes.

There are several strengths to the contrastive learning methodology applied within this work. Our approach, based on multi-organ characteristics, allows for a more nuanced understanding of complex interactions across multiple organs, while offering deeper insights into multifaceted progression of hypertensive organ disease. The contrastive semi-supervised methodology enables isolation of variance within phenotypes uniquely related to hypertension, rather than general variance in the measures¹⁹⁸. In addition, to maximise the contrast between the group with normal organ phenotypes and those with hypertensive end-organ disease, I used an unsupervised isolation forest algorithm to remove patients with normal blood pressures but who may have ‘abnormal’ healthy

phenotypes, such as athletic remodelling or inherited phenotypic variation. Previous studies have primarily relied on blood pressure, and time or age, as sole criteria for determining disease state and progression¹⁹⁹. This can lead to limited phenotype characterisation due to imprecise hypertensive disease burden categorisation and tend to be only applicable to single organs^{200,201}. Lastly, the cPCA algorithm was interpretable, enabling data-driven insights into the variables that most strongly influenced the derivation of the *HyperScore* and *HyperTrajectory*.

3.5.1 Study limitations

This study has several limitations. First, the training cohort was drawn from the UK Biobank, which includes relatively “healthy” individuals of mainly white ethnicity²⁰². While our modelled *HyperScore* and *HyperTrajectory* performed well in internal validation, further studies are needed to assess ethnicity-driven differences in hypertension progression. Second, mapping individuals within the end-organ disease landscape requires multi-modality data, which is not widely available in clinical practice. Our cPCA-based model partly addresses this limitation by selecting specific coefficients to handle fewer variables. Third, the model is cross-sectional, lacking a true temporal dimension. While the map provides a pseudo-temporal representation of disease severity based on organ damage, incorporating actual progression time, *e.g.*, age-related trajectories, could improve understanding of aging, duration of hypertension exposure, and organ function relationships. Finally, although the developed cPCA algorithm showed strong discrimination, latest unsupervised models continue to develop both within deep learning²⁰³ and as developments of contrastive PCA²⁰⁴. These methods could be explored for improved dimensionality reduction, capturing non-linear relationships and more granular phenotypes²⁰⁵.

3.5.2 Conclusions

In conclusion, by applying novel computational approaches to large scale imaging datasets, I have developed the first comprehensive and holistic model of end-organ changes related to hypertension. As such, this AI-based cross-organ tool could lay the foundation for future reliable patient-specific hypertension assessment and personalised clinical management, leveraging AI insights into clinical practice.

4. *HYTWIN* EXTERNAL TESTING

4.1 Abstract

Background. Differences across populations arise from variations in ethnicity, sex, geographic location, and other clinical factors. Although machine learning models have exhibited potential across diverse healthcare applications, they are often insufficiently evaluated on out-of-domain populations, limiting insights into their generalisability.

Objectives. To evaluate the validity of the *HyTwin* model, developed in the previous chapter, through external testing in an ethnically diverse cohort.

Methods. I included 5,507 participants from the Atherosclerosis Risk in Communities (ARIC) dataset. The ARIC dataset was matched with the UK Biobank in terms of modalities and organ measures. Identified measures were transformed into coefficients within the baseline contrastive principal component analysis (cPCA) space, based on the available variables for each individual. A *HyperScore* and *HyperTrajectory* were then assigned to each participant using a similarity-based approach with k-nearest neighbours (kNN). I compared the assigned *HyperScores* and *HyperTrajectories* with those defined in the original *HyTwin* model using Jensen-Shannon (JS) distance, regression analysis of organ damage progression, trajectory allocation and characteristics, and survival and event-free probability analysis.

Results. The ARIC cohort had a median age of 75 years (72-80), with 1,067 participants of black-African American ethnicity (19.4%). Jensen-Shannon distances were as low as 0.10 for high *HyperScore* distributions compared with the UK Biobank. No significant deviations were observed in organ damage progression relative to *HyperScore* across multiple organ measures ($p > 0.05$). The predominant end-organ characteristics were captured properly in the ARIC cohort during *HyperTrajectory* allocation. Event-free

probability and the distribution of organ-specific diseases across *HyperTrajectories* were consistent with those in the UK Biobank.

Conclusion. I demonstrated that the *HyTwin* model remains accurate and reproducible when tested on an external cohort. These findings highlight its potential applicability and generalisability in real-world settings, despite minor variations due to modality differences and available markers.

Publication Status. The results presented in this chapter are currently under-review.

1. **Alkhodari M**, Lapidaire W, Kart T, Xiong Z, Krasner S, Fletcher A, Savage N, Baumeister T, Ohuma E, Namburete A, Lamata P, Iturria-Medina Y, Chappel T, Thilaganathan B, Fraser A, Mackillop L, McManus R, Ntusi N, Hadjileontiadis L, Khandoker A, Lewandowski A, Banerjee A, Leeson P. *Contrastive machine learning to quantify hypertensive multi-organ damage and identify new disease phenotypes: a multi-national multi-modal study.*

4.2 Introduction

Human populations are inherently diverse, and this diversity is reflected in a wide range of biological, social, and environmental factors²⁰⁶. Differences across populations arise from variations in characteristics such as ethnicity, gender, geographic location, and other clinical factors²⁰⁷. In healthcare-related research, the majority of studies are conducted solely in the UK, US, Europe, and China²⁰⁸. This geographic concentration raises concerns about the applicability and generalisability of technologies to each other or to other regions of the world.

In recent years, a plethora of AI models have been developed and trained on specific cohorts to perform diverse healthcare tasks. These models have attracted considerable interest for their potential to support personalised medicine, enhance decision-making, and improve risk stratification²⁰⁹. Their reported performance is often high during the development stages, typically demonstrated through cross-validation within the same population, *i.e.*, internal validation, or through limited testing on new individuals with similar characteristics. However, while strong performance in development is encouraging, it is also essential to evaluate these models on populations that differ in ethnicity or geographic location to assess their robustness and clinical applicability.

External testing refers to the validation of model performance on a cohort that is structurally different from the development cohort²¹⁰. This process is described as out-of-domain testing, particularly when the technology is applied to a different country or to different care settings that do not align with the original development domain. Thus, in this chapter, I examine the validity of the *HyTwin* model, developed in the previous chapter, through external testing on an ethnically diverse cohort. Specifically, I assess

whether the *HyperScore* and *HyperTrajectory* metrics defined earlier remain robust and whether the major findings continue to hold in a different population.

4.3 Methods

4.3.1 Dataset for external testing

The external testing dataset was taken from the Atherosclerosis Risk in Communities (ARIC) study²¹¹. ARIC is a population-based surveillance study for myocardial infarction (MI) incidence and coronary heart disease (CHD) mortality in the United States. The main goal of the study is to understand risk factors for subclinical atherosclerosis. The study launched in 1985 in four geographically diverse regions, namely Forsyth County, North Carolina, Jackson, Mississippi, suburban Minneapolis, Minnesota, and Washington County, Maryland. All four centres approved ARIC protocols through their institutional review boards, and participants provided written consent before the enrolment. Recently, ARIC data became part of the data-sharing efforts in the National Heart, Lung, and Blood Institute (NHLBI) TransOmics for Precision Medicine (TOPMed) program²¹².

Nearly 16,000 adults (4,000 from each region) aged between 45 and 64 years old were enrolled prospectively in the study comprising white and black-African Americans early in 1987. They were followed up for more than 35 years through multiple visits, which are identified numerically, in chronological order, from visit 1 to 12. At the early visits, *i.e.*, visits 1 to 3, participants provided demographic and anthropometric data, medical history and lifestyle, clinical laboratory measurements, genetic data, cognitive and mental health details, and hospitalisation and mortality surveillance. From visit 3 onwards, participants start to be enrolled for modality-based examinations such as 12-

lead ECG, brain MR, retinal imaging, ultrasound and carotid IMT measurements, echocardiography, and spirometry. In this work, to validate the *HyTwin* model, I focused on participants in visit 5, *i.e.*, between 2011 and 2013, as this visit had complete data and involved brain MR examinations.

4.3.2 Variables matching between datasets

I initially performed a comparative analysis of end-organ metrics in ARIC compared to UK Biobank to ensure feasibility of model testing. Variables were matched automatically using their extracted name, modality, description, and imaging methods. I also performed manual analysis to ensure proper matching between variables in both datasets. Following this, I identified in ARIC a total of 123 variables that were similar to the ones in the original UK Biobank model. These included 75 brain MR, 7 cardiac imaging variables, 5 body composition, 26 blood biochemistry, 3 ECG, and 5 spirometry, in addition to 2 demographic characteristics, *i.e.*, age and sex.

4.3.3 Modality-wise differences and alignment

The majority of available measures in ARIC had been collected using comparable methods to those used in the UK Biobank. However, there were notable differences. Echocardiography was used within ARIC, as opposed to cardiac MR in UK Biobank. However, as the set of standard cardiac measures available in UK Biobank can also be generated from echocardiography, *e.g.*, LVS and LVED, I substituted the echocardiography variables for cardiac MR measures. In addition, there was a significant difference in brain MR slice thickness between ARIC and UK Biobank. As the identified variation in cardiac and brain indices between studies can lead to recognised differences in normal ranges for these measures, I allowed for this variation

in interpretation. To handle these differences, I ensure proper normalisation of the corresponding variables of these modalities and used normalised ranges when comparing between datasets where appropriate.

4.3.4 Transformation to contrastive principal components

After normalisation, the ARIC dataset was transformed into contrastive principal components (cPCs) using the available variables per participant. While individuals had a maximum of 123 variables, many had fewer. To account for this, the cPC coefficients that correspond to available variables per-individual were selected from the pre-existing cPCA space generated using the full UK Biobank dataset, *i.e.*, the *HyTwin* model, and these were applied for the transformation.

Following the approach described in Chapter 3 during internal validation, I applied a kNN-based method to identify matching coefficients between the unseen cPCs in ARIC and those in the overall UK Biobank model. In the original model, each set of cPCs, *i.e.*, each individual, is associated with a *HyperScore* and a *HyperTrajectory*. Based on the identified matches, I estimated a *HyperScore* for each individual in ARIC using their most similar counterpart in the UK Biobank and assigned a *HyperTrajectory* with corresponding probabilities for trajectory allocation using the $k = 25$ most similar individuals in UK Biobank.

4.3.5 Evaluation of *HyperScore* prediction

I evaluated the prediction of *HyperScore* in ARIC using Jensen-Shannon (JS) distance analysis described in Chapter 3. Here, I compared between the distribution of predicted

scores in the UK Biobank and ARIC datasets at three groups, namely normotensives (background), hypertensives (target), and elevated (intermediate). Furthermore, I evaluated the progression phenotype of organ changes with respect to *HyperScores* between the two datasets through regression analysis. Here, I performed statistical *t*-test analysis on the slopes of the fitted line in data of both datasets for selected variables. A non-significant result would indicate that both datasets represent similar organ changes trend with respect to increment of the estimated *HyperScores*.

4.3.6 Clinical relevance of *HyperScore* and *HyperTrajectory*

I performed the exact same protocol for evaluation of clinical relevance of *HyperScore* and *HyperTrajectory* as used in the internal clinical validation of Chapter 3. This included study of time-to-event according to *HyperScore* ranges. I further looked at *HyperTrajectory* allocation and studied variation in individual phenotypes and time-to-event by *HyperTrajectory*. All findings were compared to those observed in the internal validation to evaluate the consistency of clinical findings in the external dataset. Ideally, both cohorts would provide close findings in terms of organ-specific time-to-event or disease distribution in the defined *HyperTrajectories*.

4.4 Results

4.4.1 External dataset characteristics

The ARIC dataset comprised 5,507 participants who completed their 5th visit. The selected cohort had median age of 75 years old (72—80) with a mean arterial pressure

(MAP) of 87 mmHg (80—95). The dataset included 3,176 female participants (57.67%). I selected a similar subset in terms of age range from the UK Biobank cohort, *i.e.*, 65 to 95 years old, which are the minimum and maximum in the ARIC cohort. Accordingly, among the selected 11,409 individuals, their median age was close to ARIC at 70 years old (68—73), MAP significantly higher at 101 mmHg (93—109), and female proportions slightly lower of 5,590 (49.00%).

As the aim of the analysis was to examine time-to-event and survival, I selected a subset from ARIC that included all participants within an age range comparable to the UK Biobank, *i.e.*, 65 to 70 years. This range represented the lowest ages in ARIC and the third quartile of ages in the overall UK Biobank cohort.

Table 4.1 | Characteristics of the external testing dataset, ARIC, compared with the UK Biobank for selected age ranges.

Age range		65—90		65—70	
Datasets		ARIC (n = 5,507)	UKB (n = 11,409)	ARIC (n = 940)	UKB (n = 6,329)
Female (%)		3,176 (57.67%)	5,590 (49.00%)	568 (60.43%)	3,248 (51.32%)
Age		75 [72-80]	70 [68-73]	69 [68-70]	68 [67-69]
SBP		129 [118-141]	145 [132-159]	125 [115-136]	143 [131-157]
DBP		66 [59-73]	78 [71-85]	68 [62-75]	78 [71-86]
MAP		87 [80-95]	101 [93-109]	87 [80-95]	100 [92-108]
Ethnicity	White	4,440 (80.62%)	11,185 (98.04%)	717 (76.28%)	6209 (98.10%)
	Black	1,067 (19.38%)	32 (0.28%)	223 (23.72%)	18 (0.28%)
	Other	NA	192 (1.68%)	NA	102 (1.62%)

Continuous variables are represented as Median [Q1-Q3]. Categorical variables are represented as n (%). Other ethnicity includes mixed, Caribbean, Chinese, Indian, Bangladeshi, and Pakistani.

The ARIC subset included 940 individuals, of whom 568 (60.4%) were female. In contrast, the same age range in the UK Biobank comprised 6,329 individuals, including 3,248 females (51.32%). Regarding MAP, the UK Biobank consistently showed higher blood pressure than ARIC, with a median of 100 mmHg (92—108) compared to 87 mmHg (80—95), respectively.

The use of ARIC also allowed us to test within a cohort with an increased diversity. In the overall cohort, 19.38% were of Black-African American ethnicity (1,067 individuals), whereas the younger cohort had 23.72% (223 individuals). When compared to their UK Biobank counterparts, 0.28% was included as black ethnicity in the two selected cohorts. Full details of the whole ARIC cohort from visit 5 and the subsets used for external testing are provided in **Table 4.1**.

4.4.1 Allocation of ARIC individuals on *HyperTrajectories*

There was also a consistency of the findings within *HyperTrajectories* in ARIC compared to UK Biobank (**Figure 4.1**). The distribution of males and females across the *HyperTrajectories* was identical to UK Biobank (summarised in Chapter 3) with a predominance of females on trajectories 2, 4, and 6, while there was a predominance of males on trajectories 1, 3, and 5. Moreover, there were no significant differences in median age, SBP, and DBP between ARIC trajectories. When comparing the phenotypical patterns of key brain and heart measurements, *i.e.*, WMH and LVEF, both had similar patterns across trajectories as the UK Biobank.

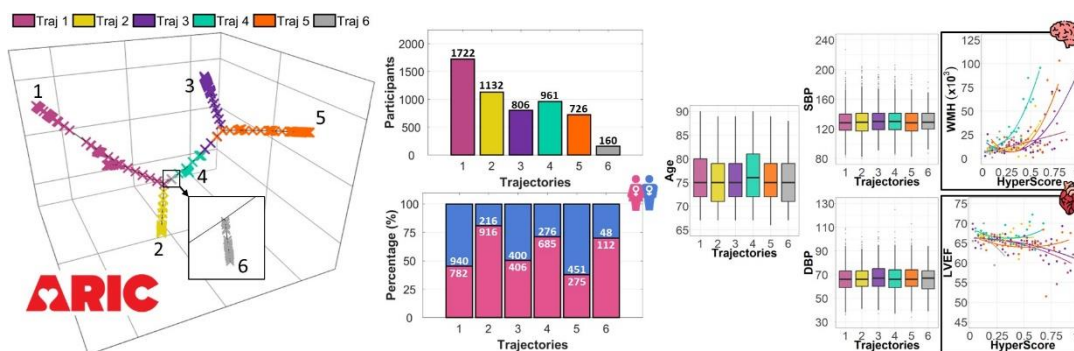


Figure 4.1 | Characterisation of the *HyperTrajectory* in the overall ARIC cohort. Placing each participant from the ARIC dataset within the *HyperTrajectory* and observing their overall representation based on sex, age, blood pressure, progression phenotypes. SBP = systolic blood pressure; DBP = diastolic blood pressure; LVEF = Left ventricular ejection fraction; WMH = White matter hyperintensity.

4.4.2 *HyperScore* prediction similarities

The distribution of *HyperScores* in ARIC across the three groups was consistent with that observed in the UK Biobank counterpart (**Figure 4.2**). Using JS distance to quantify similarity, the overall cohort showed the greatest alignment at the intermediate “E” group (JS distance = 0.14), followed by the target “H” group at 0.28. The background “N” group exhibited the largest divergence at 0.48. In the younger subset, the closest match was observed for the target “H” group at 0.17, followed closely by the intermediate “E” group at 0.18, while the background “N” group showed the greatest distance with 0.48.

I further assessed score distributions by sex and observed similar patterns across male and female groups of individuals, although men consistently exhibited significantly higher *HyperScores* compared to women in the UK Biobank cohort. This sex-related difference was similar but less pronounced within the ARIC cohort. To evaluate clinical correlation, I examined the progression of organ damage (after z-score normalisation) in relation to predicted *HyperScores* for BMI, LVS, and WMH. In the overall cohorts, slope comparisons between datasets were nonsignificant (BMI: $p = 0.135$; LVS: $p =$

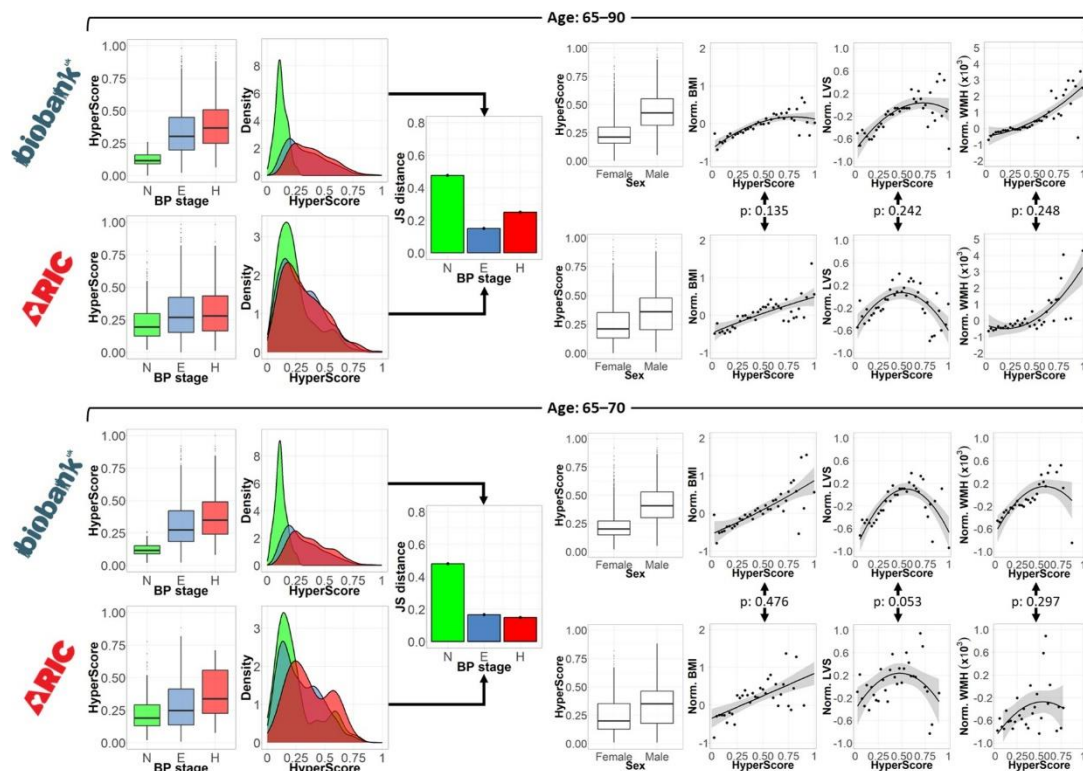


Figure 4.2 | External validation of *HyperScore* distribution and progression patterns in ARIC and UK Biobank. Analysis of ARIC cohort of 5,507 participants for two age groups, namely 65—90 (overall cohort) and 65—70 (younger subset), and participants from the UK Biobank were selected according to the same ranges. *HyperScore* distribution at every stage was compared using Jensen-Shannon (JS) distance calculations. Phenotypic patterns (z-score normalised values) with respect to *HyperScore* for the ARIC subsets were provided alongside their counterparts from the UK Biobank and were statistically compared based on slope. Points within plots represent median value of participants at each 0.025 *HyperScore* step. BP = blood pressure; BMI = Body mass index; LVS = Left ventricular stroke volume; WMH = White matter hyperintensity.

0.242; WMH: $p = 0.248$). Similarly, in the younger subset, slopes did not differ significantly (BMI: $p = 0.476$; LVS: $p = 0.053$; WMH: $p = 0.297$).

4.4.3 Clinical relevance of *HyperScore* and *HyperTrajectory*

I evaluated time-to-event outcomes for mortality and major adverse events, including cardiac, brain, and kidney diseases, stratified by *HyperScore* categories, *i.e.*, low, medium, and high, as defined in Chapter 3. The event-free probability curves demonstrated consistent patterns across cohorts, with significant differences observed

between categories (**Figure 4.3 – upper rows**). As expected, individuals in the high-risk group experienced the lowest probabilities for events and survival, whereas those in the low-risk group had the most favourable outcomes. When comparing cohorts, ARIC participants consistently showed lower event-free probabilities than those in the UK Biobank. For example, among individuals in the high-risk group, event-free probability for cardiac events was approximately 92% in the UK Biobank but declined to 75% in ARIC, with nearly twice the number of observed events.

In terms of event-free probability across the *HyperTrajectories* in ARIC and the UK Biobank (**Figure 4.3 – bottom rows**), events followed similar patterns of incidence across trajectories. For example, ARIC participants in trajectory 1 had the worst probability pattern for cardiac events and trajectory 5 had the highest mortality rates. Moreover, individuals on trajectory 5 had low event-free probabilities for renal diseases

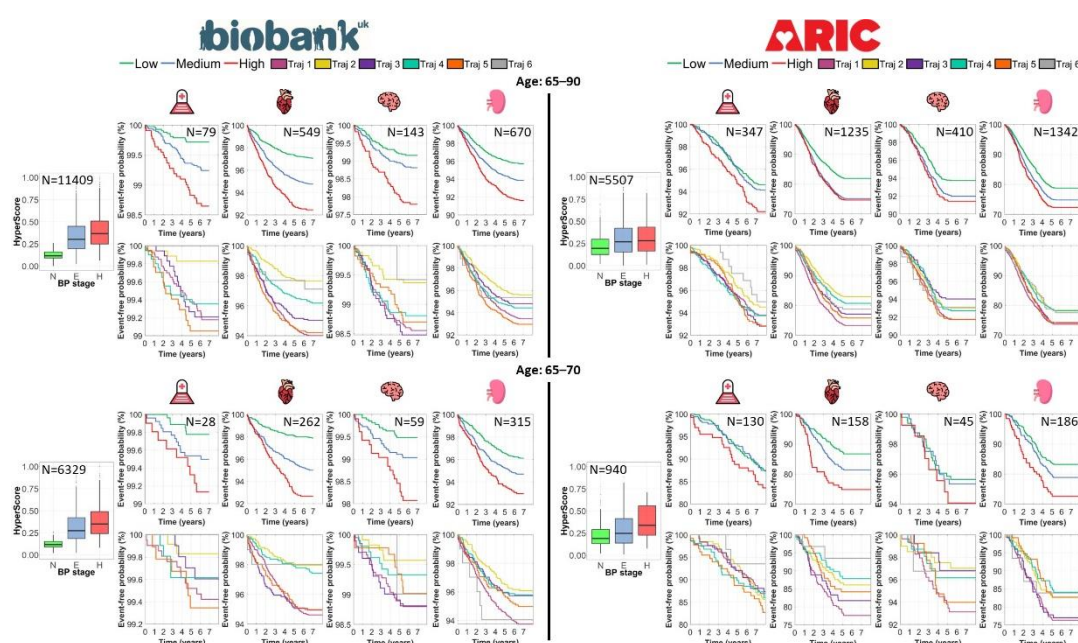


Figure 4.3 | External testing of *HyperScore* and *HyperTrajectory* on the Atherosclerosis Risk in Communities Study (ARIC) dataset (right) and comparison with the UK Biobank dataset (left). The predicted *HyperScores* relative to blood pressure groups along with survival (event-free) probability for death, cardiac, brain, and kidney events. Participants were also located on a *HyperTrajectory* and survival curves were provided for each trajectory. ARIC dataset was divided based on age range as 65-90 (top row - overall cohort) and 65-70 (bottom row - younger subset), and participants from the UK Biobank were selected according to the same ranges.

in the overall cohort. In terms of diseases proportions per trajectory, trajectory 2 in both the UK Biobank and ARIC had the highest proportion of metabolic disease, whereas trajectory 4 had the highest proportion of brain outcomes. More details are provided in **Table 4.2** for the top 5% *HyperScores* of individuals within each trajectory.

Table 4.2 | Characteristics of the *HyperTrajectory* in the advanced disease state (top 5% *HyperScores*) in the ARIC cohort.

<i>HyperTrajectory</i>	1 (N = 87)	2 (N = 57)	3 (N = 41)	4 (N = 49)	5 (N = 37)	6 (N = 8)
<i>HyperScore</i>	0.78±0.06	0.40±0.06	0.76±0.11	0.54±0.06	0.74±0.07	0.28±0.05
Demographic information						
Female	44 (50.57%)	40 (70.18%)	6 (14.63%)	39 (79.59%)	14 (37.84%)	4 (50.00%)
Age, years	77.77±5.76	77.42±4.42	77.71±4.48	77.90±6.42	77.57±4.77	76.25±4.03
SBP, mmHg	131.76±19.19	125.07±14.51	130.12±18.37	132.98±20.64	138.68±20.09	125.00±9.47
DBP, mmHg	66.77±9.85	63.53±7.42	66.85±7.90	65.63±8.76	68.97±10.12	66.75±10.54
PP, mmHg	64.99±16.64	61.54±11.50	63.27±16.27	67.35±17.32	69.70±15.82	58.25±7.72
MAP, mmHg	88.43±11.22	84.04±8.80	87.94±9.76	88.08±11.25	92.21±12.13	86.17±9.52
Current smoker	4 (4.60%)	6 (10.53%)	4 (9.76%)	3 (6.12%)	6 (16.22%)	1 (12.5%)
Diabetic	36 (41.38%)	11 (19.3%)	22 (53.66%)	11 (22.45%)	11 (29.73%)	1 (12.5%)
Heavy alcohol drinker	42 (48.28%)	32 (56.14%)	23 (56.10%)	19 (38.78%)	21 (56.76%)	6 (75.00%)
Hypertension Medication	74 (85.06%)	42 (73.68%)	32 (78.05%)	40 (81.63%)	29 (78.38%)	6 (75.00%)
Cholesterol Medication	69 (79.31%)	31 (54.39%)	31 (75.61%)	24 (48.98%)	23 (62.16%)	6 (75.00%)
Imaging/non-imaging measurements						
WMH, $\times 10^3 \text{ mm}^3$	32.13±24.37	14.86±13.44	45.58±39.92	78.76±18.12	81.13±28.86	16.69±8.28
CO, L/min	3.41±1.03	3.12±0.92	3.80±1.17	3.29±0.91	3.46±0.89	3.87±1.03
CI, L/min/m²	1.81±0.50	1.75±0.43	1.88±0.50	1.76±0.39	1.85±0.39	2.01±0.41
LVEF, %	62.62±8.44	67.22±5.05	62.33±7.00	65.72±5.52	63.13±7.11	61.62±7.14
VR, bpm	68.98±10.72	62.58±7.98	64.80±9.14	65.12±8.25	65.46±7.97	67.00±6.09
BMI, kg/m²	30.12±6.05	26.69±5.61	30.14±4.52	31.57±8.80	28.02±3.37	30.42±6.98
BFM, kg	29.39±12.41	23.74±8.90	30.15±11.64	33.77±16.45	26.16±6.97	30.35±13.47
Cholesterol, mmol/L	4.31±0.95	4.76±0.95	4.08±0.93	4.74±0.98	4.53±1.27	4.56±0.73
LDL, mmol/L	2.28±0.78	2.64±0.79	2.27±0.73	2.70±0.78	2.58±0.98	2.76±0.61
HDL, mmol/L	1.21±0.28	1.56±0.53	1.11±0.26	1.40±0.38	1.29±0.33	1.13±0.20
Triglycerides, mmol/L	1.79±1.00	1.20±0.42	1.54±0.69	1.38±0.56	1.43±0.65	1.45±0.36

HbA1c, mmol/mol	44.66±10.36	38.64±5.62	44.20±11.18	40.09±6.57	42.11±10.65	37.71±3.41
GLC, mmol/L	7.05±2.57	5.89±0.97	6.44±1.18	6.25±1.13	6.57±1.49	5.86±0.59
CRTN, µmol/L	87.20±25.36	87.19±25.97	104.20±50.07	72.33±16.88	97.62±32.14	73.04±14.28
CysC, mg/L	1.24±0.29	1.28±0.40	1.36±0.62	1.21±0.30	1.25±0.38	1.08±0.15
WBC, $\times 10^9$ cells/L	6.54±1.75	5.96±2.04	6.50±1.56	6.62±4.46	5.99±1.50	5.59±1.14
Adverse outcomes						
Death (Circulatory)	9 (10.59%)	2 (4.44%)	4 (10.81%)	5 (12.20%)	5 (13.89%)	0 (0.00%)
Cardiac	27 (31.76%)	7 (15.56%)	6 (16.22%)	9 (21.95%)	9 (25.00%)	2 (40.00%)
Brain	11 (12.94%)	6 (13.33%)	2 (5.41%)	8 (19.51%)	4 (11.11%)	0 (0.00%)
Vasculature	5 (5.88%)	1 (2.22%)	3 (8.11%)	1 (2.44%)	4 (11.11%)	0 (0.00%)
Pulmonary	6 (7.06%)	6 (13.33%)	6 (16.22%)	3 (7.32%)	1 (2.78%)	1 (20.00%)
Liver	2 (2.35%)	1 (2.22%)	1 (2.70%)	3 (7.32%)	1 (2.78%)	0 (0.00%)
Kidney	25 (29.41%)	18 (40.00%)	13 (35.14%)	11 (26.83%)	10 (27.78%)	2 (40.00%)
Metabolic	0 (0.00%)	4 (8.89%)	2 (5.41%)	1 (2.44%)	2 (5.56%)	0 (0.00%)
Overall	85 (97.70%)	45 (78.95%)	37 (90.24%)	41 (83.67%)	36 (97.30%)	5 (62.50%)

Continuous variables are represented as Mean±SD. Categorical variables are represented as n (%). The percentage of organ-specific outcomes is based on the overall number of outcomes. The percentage of the overall outcomes is based on the total number of participants. Bold indicates the trajectory with highest number of outcomes if enough outcomes occur.

4.5 Discussion and conclusion

In this chapter, I tested the performance of the *HyTwin* model in identifying similar phenotypic alterations in another ethnically diverse and geographically distinct cohort, the ARIC dataset. Most importantly, the *HyperScore* estimation and the prediction of *HyperTrajectories* for individuals in ARIC were consistent with those observed in the UK Biobank, highlighting the reproducibility of the *HyTwin* model and its findings in a closely related cohort.

The model performed well despite the limited number of available variables in ARIC. In different populations, it is expected that datasets may not include all imaging modalities utilised by the *HyTwin* model. However, the cTI approach used to build the

model enables the selection of coefficients corresponding to the variables that are available. These coefficients were derived from the full set of variables in the UK Biobank. Thus, in real-world settings, the model is expected to remain functional and deliver stable performance even when only a limited set of modalities is available.

Differences in imaging modalities and settings used to acquire the measures in ARIC did not also appear to significantly impact the stability of the model. For instance, echocardiography was used for cardiac measures in ARIC instead of cardiac magnetic resonance and ARIC used a 5 mm brain MR slice thickness compared to 1 mm in UK Biobank. Despite these technical variations, the proposed model appeared to remain resilient both in terms of range of estimated *HyperScore* across blood pressure stages and allocation of individuals into disease-specific trajectories.

In terms of predicting *HyperScore* in ARIC, which reflects the severity of global organ damage, results were consistent with observations from the UK Biobank. The literature has shown that the ARIC study includes individuals who are generally older and exhibit more severe cardiovascular disease^{211,213}, especially when compared to the UK Biobank cohort at similar visit timepoint. Notably, the JS distance was lowest among individuals with higher blood pressure, indicating strong similarity in predictions. Furthermore, males were at higher risk compared to females which was the same trend observed in the UK Biobank cohort. Most importantly, phenotypic changes in organ function showed similar trends with increasing *HyperScore*, with no significant differences compared to the UK Biobank cohort. This highlights the model's ability to consistently characterise and capture the progression of alterations associated with greater disease severity.

Similarly, across mortality and organ-specific diseases, event-free probabilities were lowest for high-risk individuals in ARIC, consistent with findings from the UK Biobank.

However, in the older cohort (65–90 years), the separation between the medium- and high-risk curves was less pronounced compared to the UK Biobank. For example, in ARIC, the older cohort showed overlapping event-free probability curves for the medium- and high-risk groups, whereas the younger cohort had differences in the estimated probabilities in these groups, similar to the trend in the UK Biobank cohort. This pattern reflects both the greater cardiovascular disease burden in the ARIC population and the effect of age, as ARIC participants were generally older than their UK Biobank counterparts.

Placement of individuals from the ARIC dataset onto a *HyperTrajectory* demonstrates a real-world application of the *HyTwin* model to new patient data. The predominant characteristics of individuals mapped to these trajectories were consistent with those used to characterise the trajectories in the original *HyTwin* model. The correspondence between survival and time-to-event analyses in both cohorts was evident, along with progression patterns in *HyperScore* on each trajectory, *e.g.*, as observed with WMH. While the proportion of diseases differed slightly between ARIC and the UK Biobank, the *HyperTrajectories*, although characterised by a predominant disease, showed overlaps in their disease profiles. Because individuals may present with multiple organ conditions, each person was mapped to the trajectory that best represented their overall profile. This highlights the practical utility of *HyperTrajectories*, where assigning an individual to a trajectory can offer insights not only into their primary condition but also into associated comorbidities and organ alterations.

4.5.1 Study limitations

This study has several limitations. First, while our modelled *HyperScore* and *HyperTrajectory* performed well and showed consistent associations with hypertension

in the ARIC dataset, which included nearly 20% Black-African Americans, further studies are needed to assess ethnicity-driven differences in hypertension progression. Second, although the results in the ARIC dataset were reproducible, our *HyTwin* model operated with fewer variables, incorporating additional variables could potentially improve the accuracy of the findings. Third, while I evaluated performance across different imaging modality settings, further investigation is required to better understand model usability and robustness under varying modality conditions. Fourth and Last, although the reproducibility of *HyperScore* estimation and *HyperTrajectory* was encouraging, the validation focused on identifying similar findings across both datasets, which did not necessarily require a ground truth. However, establishing additional validation strategies would enable the use of standard evaluation metrics to elaborate on the performance.

4.5.2 Conclusions

In conclusion, I evaluated the reproducibility of the *HyTwin* model on an external, out-of-domain dataset and demonstrated that it remained accurate. The ability to reproduce results across different cohorts and settings supports the applicability and generalisability of the model. In real-world clinical practice, *HyTwin* is expected to perform consistently, with only minor variations as outlined in the limitations related to modality differences and available markers, while maintaining overall robustness.

5. *HYTWIN* AND
HYPERTENSIVE
DISORDERS OF
PREGNANCY

5.1 Abstract

Background. Hypertension during pregnancy is associated with severe outcomes later in life due to organ alterations that develop as adaptations to this condition. Understanding pregnancy-related factors and subsequent organ damage may help explain these increased mid-life risks.

Objectives. To identify pregnancy-associated phenotypes related to long-term organ alterations, particularly in the presence of hypertensive disorders of pregnancy (HDP).

Methods. I used data from the Pregnancy Complications Vascular Study (PVS), which included 268 participants (104 normotensive pregnancies, 164 HDP) who underwent multimodal imaging assessments 5-10 years postpartum. Data was projected into the *HyTwin* contrastive principal component analysis (cPCA) space, and each participant was assigned a *HyperScore* and *HyperTrajectory* using a similarity-based k-nearest neighbours (kNN) approach. Accordingly, predominant phenotypes along each trajectory were characterised for these participants based on their pregnancy-related factors. These phenotypes were validated for reproducibility in a pregnancy-specific subset derived from the UK Biobank.

Results. In PVS, the HDP group had a significantly higher proportion of participants with a maternal history of HDP ($n = 22 - 13.41\%$, $p = 0.004$), particularly in early preeclampsia (EPET), and a greater proportion who underwent caesarean section ($n = 32 - 19.51\%$, $p = 0.030$). Individuals with higher *HyperScores* were more likely to have experienced miscarriages, stillbirths, or caesarean section deliveries, or to have been born to a mother previously affected by HDP. The analysis identified predominant phenotypes across *HyperTrajectories*, including normotensive preterm ($p = 0.038$), low

birth centile likely related to twin pregnancy ($p = 0.027$), family history ($p = 0.047$), female babies with evidence of late onset ($p = 0.031$), and foetal loss ($p = 0.022$), all associated with significantly high odds ratios. In the UK Biobank validation set (1,910 participants, 955 controls, 955 HDP), these predominant phenotypes were reproducible, with no significant differences in the ratios of odds ratios between datasets.

Conclusion. I discovered through the *HyTwin* that distinct pregnancy-related characteristics are associated with specific patterns of organ alterations later in life. The results underscore the potential for targeted management of hypertensive pregnancies exhibiting these phenotypes to mitigate adverse long-term outcomes.

Publication Status.

The results presented in this chapter were published in:

1. **Alkhodari M**, Xiong Z, Khandoker A, Hadjileontiadis L, Leeson P, Lapidaire W. *The role of artificial intelligence in hypertensive disorders of pregnancy: towards personalized healthcare*. Expert Review of Cardiovascular Therapy, Taylor & Francis. 2023;21(7):531-543.

More results are currently drafted and prepared for submission.

2. **Alkhodari M**, Krasner S, Cutler HR, Sattwika PD, Kart T, Savage N, Baumeister T, Ohuma E, Namburete A, Lamata P, Iturria-Medina Y, Chappel T, Thilaganathan B, Fraser A, Mackillop L, McManus R, Ntusi N, Hadjileontiadis L, Khandoker A, Banerjee A, Lewandowski A, Lapidaire W, Leeson P. *Discovering pregnancy phenotypes associated with mid-life hypertensive multi-organ damage using contrastive machine learning*.

5.2 Introduction

Hypertension may arise during pregnancy due to the inadequate adjustment of the maternal cardiovascular system to the increased haemodynamic demands of the foetus²¹⁴. As a result, adaptations in placental flow from the mother to the foetus can elevate blood pressure, leading to severe alterations in major organs and causing hypertensive disorders of pregnancy (HDP)²¹⁵. In literature, women with a history of HDP are up to four times more likely to develop vascular diseases later in life, due to persistent elevation in peripheral resistance, remodelling of the left ventricle, alterations in cerebral blood flow, and other renal abnormalities^{216,217}.

Current UK guidelines recommend regular follow-up to monitor cardiovascular health after a hypertensive pregnancy, however, evidence on whether this strategy reduces long-term disease incidence remains limited²¹⁸. Moreover, assessment of blood pressure, body mass index (BMI), and lipid and glucose profiles during and after pregnancy have been used to identify women at risk later in life²¹⁹. Most recently, several trials indicated that postpartum care strategies may alter the risk trajectory, as interventions within the first six weeks after delivery have been shown to lower blood pressure for up to three years^{218,220}, highlighting the importance of early postpartum care in reducing long-term risks.

However, end-organ alterations after pregnancy, along with pregnancy-related and lifestyle factors associated with these changes, remain poorly characterised. Improved understanding of organ damage after a hypertensive pregnancy may enable earlier interventions and proper management of maternal health during pregnancy and before complications develop in the postpartum period²²¹. Furthermore, knowledge about distinct patterns in pregnancy-related and lifestyle factors during pregnancy may provide

valuable insights into disease outcomes later in life and support proactive management to mitigate future risk. Therefore, in this chapter, I apply the *HyTwin* model to identify distinct pregnancy-specific phenotypes based on integrated multi-organ information obtained from women years after delivery, thereby demonstrating *HyTwin*'s potential as an organ damage-specific model to link patterns of organ damage with patient health profiles and risk factors. Within the HDP domain, *HyTwin* achieves this by phenotyping pregnancy-related characteristics during the antepartum period and associating them with their long-term outcomes defined by post-pregnancy organ alterations.

5.3 Methods

5.3.1 Dataset description and participants

The main dataset used in this chapter was obtained from the Pregnancy Complications Vascular Study (PVS), which is part of the Oxford Cardiovascular Clinical Research Facility (CCRF) *HyPrevent* database. This database integrates a series of clinical trials designed to investigate the long-term effects of pregnancy-related hypertension on the heart, brain, and blood vessels across a woman's lifespan, as well as the health of children born following such pregnancies.

The PVS study focuses on investigating the association between pregnancy complications and cardiovascular disease in later life, with the aim of identifying women at increased risk and exploring strategies to reduce that risk²²². The study was conducted at John Radcliffe Hospital (JRH) and sought to recruit over 250 participants. Eligible participants were women who had delivered at Oxford University Hospitals Trust NHS Trust Maternity Unit between 5 and 10 years prior to recruitment (with pregnancies occurring between 1998 and 2010). The study population included women with a history

of hypertensive pregnancy, specifically pregnancy-induced hypertension (PIH) and preeclampsia (PET). Preeclampsia diagnosed before 34 weeks of gestation was classified as early-onset PET (EPET), whereas diagnosis at or after 34 weeks was classified as late-onset PET (LPET)²². In addition, women with normotensive pregnancies (controls) and those with normotensive preterm deliveries were recruited for comparison. The PVS study was approved by the Oxfordshire Research Ethics Committee (REF: 08/H0604/127).

The PVS dataset included a total of 268 participants, of whom 104 had normotensive pregnancies (NP), including 75 controls and 29 preterm, and 164 had HDP, including 59 EPET, 84 LPET, and 21 PIH. All participants had available maternal demographic information, including age at recruitment, blood pressure at recruitment, body mass index (BMI) at recruitment, smoking status at recruitment, maternal birthweight, history of being born to HDP, history of miscarriages/stillbirths, and history of caesarean section. In addition, gestational information was collected, such as gestational age, baby birthweight, birth centile, age at delivery, blood pressure at booking and in the third trimester, BMI at booking, details of twin pregnancies, baby sex, and smoking status during pregnancy. Participants also had undertaken detailed multimodal imaging and functional testing, including cardiac and brain MR, echocardiographic assessment of cardiac function, imaging of small blood vessels and arteries, and routine blood biochemistry testing.

5.3.2 Data preparation

I performed a comparative analysis of end-organ metrics between PVS and the baseline UK Biobank dataset used to develop the *HyTwin* to ensure the inclusion of matching variables. Variables in PVS were aligned with the *HyTwin* by automatically extracting

their names, modalities, description, and imaging methods. In addition, I conducted a manual review to confirm proper matching between the two datasets. Through this process, I identified a total of 188 comparable variables: 148 brain MR, 12 cardiac MR, 2 body composition, 19 blood biochemistry, 1 ECG, and 4 spirometry measures, in addition to 2 demographic characteristics, *i.e.*, age and sex. To handle any differences that may arise between PVS and the UK Biobank, I ensured proper normalisation of the corresponding variables of these modalities using normalised ranges taken from the baseline UK Biobank dataset.

5.3.3 Transformation to contrastive principal components

Using the available variables per participant and the pre-existing baseline contrastive principal components analysis (cPCA) space in *HyTwin*, PVS dataset was transformed into contrastive principal components (cPCs). Participants varied in their available variables depending on the clinical testing they have conducted, thus, many had fewer variables than 188. To account for this, only the cPC coefficients corresponding to the available variables were used for transformation. The transformation was performed by projecting the normalised data onto the cPCA space through matrix multiplication with the selected loading coefficients.

Upon transformation, I used the same kNN-based approach described in **Chapter 3** to identify matching coefficients between the unseen cPCs in PVS and those in the *HyTwin*. Here, our baseline pre-existing cPCA space had a *HyperScore* and a *HyperTrajectory* for each individual in the UK Biobank cohort. By identifying matches, I estimated a *HyperScore* for each individual in PVS using their most similar counterpart in the UK

Biobank and assigned a *HyperTrajectory* with corresponding probabilities for trajectory allocation using the $k = 25$ most similar individuals in the UK Biobank cohort.

5.3.4 Summarisation of information across pregnancy subgroups

I conducted the analysis of PVS by evaluating various pregnancy-related factors through comparisons of medians and proportions across pregnancy subgroups. Specifically, I compared the five subgroups, *i.e.*, control, normotensive preterm, EPET, LPET, and PIH, using the continuous and categorical factors defined in **Section 5.3.1**. The analysis was performed both by summarising characteristics by median and inter-quartile range across the two main groups, *i.e.*, NP and HDP, and within each subgroup. Most importantly, I focused on gestational age, baby birthweight, and baby sex, and calculated the corresponding birth centile using the Intergrowth-21 algorithm²²³. Additionally, for categorical factors, I evaluated the proportions as a percentage of the total within each subgroup, including maternal history of being born to HDP, history of miscarriage/stillbirth, history of caesarean section, and twin pregnancies in the current gestation. Statistical tests were used to identify significantly different factors across different groups (**Section 5.3.9**).

5.3.5 Summarisation of information based on *HyperScore* and *HyperTrajectory*

I categorised PVS participants based on their estimated *HyperScores* according to the three *HyperScore*-based stages defined earlier in **Chapter 3**, namely low-, medium-, and high-risk. I then assessed differences among the pregnancy-related groups both in terms

of medians for continuous factors, *e.g.*, gestational age, and the proportional percentages for categorical factors, *e.g.*, miscarriage/stillbirth rate, across the defined *HyperScore* ranges. In addition, I evaluated these values and proportions at the predicted *HyperTrajectories* for all participants. Both the *HyperScore* and *HyperTrajectory* analyses were conducted for the overall cohort as well as for HDP subgroups only.

5.3.6 Odds ratio analysis of *HyperTrajectories*

To identify distinct pregnancy phenotypes across *HyperTrajectories*, I performed odds ratio (OR) analysis on a range of categorical characteristics. The variables included pregnancy subgroup classifications, namely the five subgroups (control, normotensive preterm, EPET, LPET, and PIH), as well as the combined NP and HDP groups and a combined PET category (merging EPET and LPET). Additionally, I evaluated other clinically relevant categorical outcomes that were available, including maternal history of being born to an HDP, miscarriage/stillbirth and multiple miscarriages/stillbirths, caesarean delivery and multiple caesareans, twin pregnancy, and birth of a female baby. For each characteristic, ORs and 95% confidence intervals (CIs) were calculated by comparing the prevalence of the outcome in a given trajectory with its prevalence in all other trajectories combined.

5.3.7 Characterising the observed phenotypes

I summarised the characteristics of each *HyperTrajectory* by analysing the overall PVS cohort in terms of phenotypic differences in demographics and gestational factors, the proportional distribution of pregnancy-related variables, and odds ratios significance. For each *HyperTrajectory*, I subsequently defined a predominant characteristic based on combined knowledge from the following criteria:

- 1) **Demographic and gestational factors:** I assessed differences in demographic and gestational variables including age, blood pressure, BMI, gestational age, birthweight, and birth centile by comparing their median values across *HyperTrajectories*.
- 2) **Proportional distribution of pregnancy-related factors:** I calculated the rates of categorical pregnancy-related factors, such as maternal history of being born to an HDP, miscarriage/stillbirth, caesarean section, twin pregnancy, female babies, and smoking status, expressed as percentages relative to the total number of participants in each *HyperTrajectory*.
- 3) **Odds ratios significance:** I identified variables with a significantly higher or lower likelihood of occurrence based on odds ratio analysis and considered these significant associations when characterising each *HyperTrajectory*.

5.3.8 UK Biobank validation

To validate the identified predominant characteristics and test reproducibility, I selected an independent pregnancy subset of participants from the UK Biobank. Specifically, I included female participants with documented incidents of HDP, including PET and PIH, identified through ICD-10 codes and hospital episode statistics (HES) records (O13 and O14). In addition, I randomly selected a control group of the same size from the UK Biobank consisting of women with no recorded history of hypertension or HDP, whose blood pressure measurements at both their first assessment and imaging visits were below 120/80 mmHg. I then applied the same procedure described in **Section 5.3.3** to transform their variables into the cPCA space and assign each participant a corresponding *HyperScore* and *HyperTrajectory*. I extracted the available pregnancy-related variables that were comparable to those in the PVS cohort, including history of

miscarriage/stillbirth, multiple miscarriages/stillbirths, caesarean delivery, twin pregnancy, and female baby, alongside reported cases of PET and PIH.

I performed the same odds ratio analysis conducted in **Section 5.3.6** for the available pregnancy factors within the UK Biobank pregnancy subset. Then, I compared the effect sizes between PVS and UK Biobank cohorts using a ratio of odds ratios (ROOR) approach²²⁴. Here, the odds ratios calculated for each outcome per trajectory in the PVS cohort were divided by the corresponding odds ratios derived from the UK Biobank cohort. ROOR values close to one demonstrate similarity in OR between datasets.

5.3.9 Statistical analysis

To compare continuous variables between pregnancy subgroups or between *HyperTrajectories*, I used a linear mixed-effects (LME) model adjusted for age at recruitment. For categorical variables, I applied a generalised linear model (GLM), also adjusted for age at recruitment.

In OR analysis, each selected trajectory was considered the case group while the others were considered as the control group. OR estimation was performed using a standard Wald method on the log-odds scale²²⁵. To address any issues caused by the small sample size, I applied the Agresti-Caffo continuity correction (adding 1 to each cell in the resulting confusion matrix), which improves the stability and interpretability of OR estimates and CI ranges²²⁶. Cases with zero events in the trajectory of interest were reported with an empty OR and CIs.

In ROOR analysis, the logarithm of the ROOR was computed as the difference between the log-transformed odds ratios from each cohort, and its standard error was calculated as the square root of the sum of the variances of the log-transformed odds ratios from

both cohorts. The 95% CI was then derived for the ROOR and a Wald z -test was performed to evaluate whether both datasets had significantly different ORs²²⁷.

For the OR and ROOR significance analysis, I calculated the z -statistic by dividing the log-transformed OR and ROOR by their standard error (SE). The corresponding two-sided p -value was then computed as,

$$p = 2 \times [1 - \phi(|z|)]$$

where ϕ is the cumulative distribution function of the normal distribution, *i.e.*, the z -statistic.

5.4 Results

5.4.1 Characteristics of the PVS cohort

A detailed summary of the PVS cohort is provided in **Table 5.1**. The HDP group had significantly higher ($p < 0.001$) systolic and diastolic blood pressure levels at recruitment, as well as at the initial booking during pregnancy and in the third trimester, compared with the NP group. They also had a significantly higher BMI at recruitment ($p = 0.027$) and at booking ($p = 0.034$), with median values of 25.6 kg/m² [23.0—30.0] and 25.0 kg/m² [22.0—28.0], respectively.

Table 5.1 | Summary of pregnancy characteristics of the PVS cohort.

Groups	NP (N = 104)	HDP (N = 164)	<i>p</i> -value (age adjusted)	Control (N = 75)	Norm. Preterm (N = 29)	EPET (N = 59)	LPET (N = 84)	PIH (N = 21)	<i>p</i> -value (age adjusted)
<i>HyperScore</i>	0.176 [0.114–0.275]	0.192 [0.112–0.338]	0.435	0.158 [0.102–0.267]	0.199 [0.135–0.295]	0.188 [0.119–0.339]	0.187 [0.107–0.338]	0.228 [0.117–0.308]	0.260
Maternal demographic information									
Age at recruitment, <i>years</i>	41.0 [38.0–44.0]	41.0 [37.8–45.0]	NA	40.0 [37.5–43.0]	43.5 [39.0–46.0]	40.0 [37.0–45.0]	41.0 [38.0–46.0]	43.0 [38.0–46.0]	NA
SBP at recruitment, <i>mmHg</i>	111.8 [105.2–120.5]	120.0 [111.6–128.2]	<0.001	111.0 [104.2–117.8]	118.0 [109.0–122.0]	120.0 [112.0–131.8]	120.7 [111.2–128.2]	119.3 [115.3–126.0]	<0.001
DBP at recruitment, <i>mmHg</i>	72.5 [67.9–79.1]	77.7 [73.0–85.4]	<0.001	72.0 [67.5–78.2]	76.0 [71.0–80.0]	78.7 [73.7–86.8]	76.5 [72.5–83.8]	80.7 [70.3–88.0]	<0.001
BMI at recruitment, <i>kg/m²</i>	24.3 [21.9–27.5]	25.6 [23.0–30.0]	0.027	24.4 [22.4–27.3]	24.1 [21.4–28.3]	26.8 [23.0–30.3]	25.5 [22.9–29.3]	25.3 [23.2–30.9]	0.083
Maternal birthweight, <i>g</i>	3174.0 [2920.0–3387.5]	3174.0 [2750.0–3507.5]	0.978	3174.0 [3061.0–3402.0]	2920.0 [2721.0–3274.0]	3203.0 [2778.0–3628.0]	3174.0 [2721.0–3444.0]	3061.0 [2848.5–3416.0]	0.113
Maternal born to HDP, <i>n</i>	5 (4.81%)	22 (13.41%)	0.004	3 (4.00%)	2 (6.90%)	9 (15.25%)	11 (13.10%)	2 (9.52%)	0.147
Miscarriage/stillbirth, <i>n</i>	12 (11.54%)	9 (5.49%)	0.228	3 (4.00%)	9 (31.03%)	3 (5.08%)	4 (4.76%)	2 (9.52%)	0.005
Multiple miscarriages/stillbirths, <i>n</i>	4 (3.85%)	5 (3.05%)	0.916	1 (1.33%)	3 (10.34%)	1 (1.69%)	2 (2.38%)	2 (9.52%)	0.069
Caesarean, <i>n</i>	15 (14.42%)	32 (19.51%)	0.030	0 (0.00%)	15 (51.72%)	15 (25.42%)	13 (15.48%)	4 (19.05%)	<0.001
Multiple caesareans, <i>n</i>	4 (3.85%)	10 (6.10%)	0.169	0 (0.00%)	4 (13.79%)	3 (5.08%)	5 (5.95%)	2 (9.52%)	0.026
Smoker at recruitment, <i>n</i>	11 (10.58%)	11 (6.71%)	0.773	8 (10.67%)	3 (10.34%)	5 (8.47%)	3 (3.57%)	3 (14.29%)	0.346
Gestational information									
Gestational age, <i>weeks</i>	39.0 [34.0–40.0]	36.0 [34.0–38.0]	0.008	40.0 [39.0–41.0]	31.0 [28.0–33.0]	33.0 [31.0–35.0]	37.0 [36.0–39.0]	39.0 [38.0–40.0]	<0.001
Baby birthweight, <i>g</i>	3326.0 [2405.5–3801.0]	2567.0 [1880.0–3237.0]	<0.001	3570.5 [3290.8–3940.5]	1480.0 [1071.0–1992.0]	1855.0 [1188.2–2257.0]	2881.0 [2388.0–3315.0]	3197.5 [2643.0–3537.5]	<0.001
Birth centile, <i>percentile</i>	72.4 [56.5–89.9]	33.8 [10.5–65.2]	<0.001	72.4 [56.5–89.9]	NA	25.4 [10.4–55.4]	43.0 [15.3–72.8]	NA	<0.001
Age at delivery, <i>years</i>	34.0 [29.0–36.0]	33.0 [30.0–37.0]	0.820	31.5 [27.2–35.0]	35.0 [30.5–38.0]	30.0 [28.5–35.0]	36.5 [31.8–39.0]	33.0 [31.0–36.2]	0.137
SBP at booking, <i>mmHg</i>	110.0 [100.0–118.0]	111.5 [105.0–121.5]	<0.001	110.0 [100.0–114.5]	110.0 [105.0–118.0]	110.0 [107.5–123.0]	112.0 [102.0–120.0]	120.0 [108.8–125.2]	0.004
DBP at booking, <i>mmHg</i>	65.0 [60.0–70.0]	70.0 [60.0–75.0]	<0.001	65.0 [60.0–70.0]	65.0 [60.0–68.8]	70.0 [62.0–77.0]	70.0 [60.0–70.5]	71.0 [68.0–78.2]	0.004
SBP at 3 rd trimester, <i>mmHg</i>	110.0 [104.0–120.0]	127.0 [116.5–139.0]	<0.001	110.0 [102.2–118.5]	112.5 [107.5–120.0]	135.0 [124.0–150.0]	120.0 [110.0–130.0]	117.5 [111.5–130.5]	<0.001
DBP at 3 rd trimester, <i>mmHg</i>	66.5 [60.0–70.0]	80.0 [70.0–89.8]	<0.001	65.0 [60.0–70.0]	70.0 [60.8–71.2]	90.0 [80.0–98.0]	78.0 [70.0–80.0]	71.5 [68.0–78.0]	<0.001
BMI at booking, <i>kg/m²</i>	23.0 [22.0–26.0]	25.0 [22.0–28.0]	0.034	24.0 [22.0–26.0]	22.7 [21.1–25.0]	26.0 [22.0–29.5]	24.6 [22.0–27.0]	25.4 [23.6–28.4]	0.122
Twin pregnancy, <i>n</i>	7 (6.73%)	10 (6.10%)	0.722	0 (0.00%)	7 (24.14%)	7 (11.86%)	2 (2.38%)	1 (4.76%)	<0.001
Female babies, <i>n</i>	20 (19.23%)	46 (28.05%)	0.005	20 (26.67%)	0 (0.00%)	16 (27.12%)	30 (35.71%)	0 (0.00%)	0.275
Smoker at pregnancy, <i>n</i>	10 (9.62%)	11 (6.71%)	0.952	6 (8.00%)	4 (13.79%)	5 (8.47%)	4 (4.76%)	2 (9.52%)	0.644

When evaluating pregnancy subgroups, the normotensive preterm group had the lowest maternal birthweight (2.92 kg [2.72—3.27]) and the lowest baby birthweight (1.48 kg [1.07–1.99], $p = <0.001$). Moreover, it showed significantly higher proportions of miscarriages/stillbirths, caesarean deliveries, and twin pregnancies. Notably, the PET group (combined EPET and LPET) had the highest proportion of participants with a maternal history of being born to HDP, while the EPET group had significantly higher blood pressure in the third trimester. A breakdown for major characteristics is provided in **Figure 5.1**.

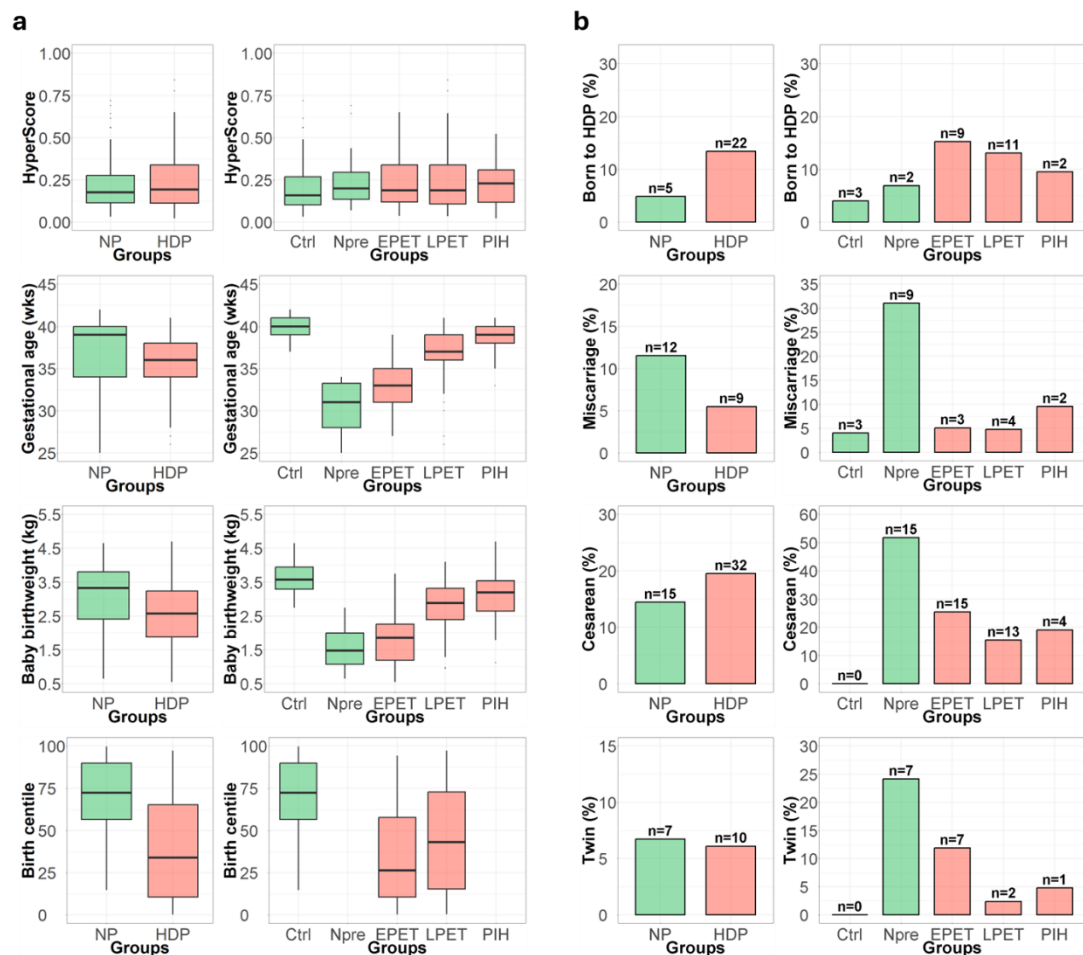


Figure 5.1 | Major characteristics of the PVS cohort for overall normotensive pregnancy (NP) and hypertensive disorder of pregnancy (HDP) groups, alongside their corresponding subgroups. a, Evaluation of continuous variables, including *HyperScore*, gestational age, neonatal birthweight, and birth centile. **b,** Analysis of categorical variables, including maternal history of HDP, miscarriages, caesarean delivery, and twin pregnancies.

5.4.2 PVS and the estimated *HyperScores*

The distribution of pregnancy subgroups within the PVS cohort across the three *HyperScore* stages revealed distinct patterns (**Figure 5.2a**). In total, 175 participants were classified as low risk, 77 as medium risk, and 16 as high risk. Among those in the high-risk category, approximately 31% ($n = 5$) belonged to the EPET subgroup. In contrast, the majority of participants in the low-risk category were from the control and LPET subgroups, each accounting for about 31% ($n = 54$). Notably, most HDP subgroups were concentrated in the medium-risk category, whereas the NP subgroups were predominantly represented in the low-risk category.

In the analysis of major pregnancy factors within the overall PVS cohort (**Figure 5.2b**), gestational age, baby birthweight, and birth centile did not differ significantly across the three *HyperScore* stages. However, categorical factors including maternal history of

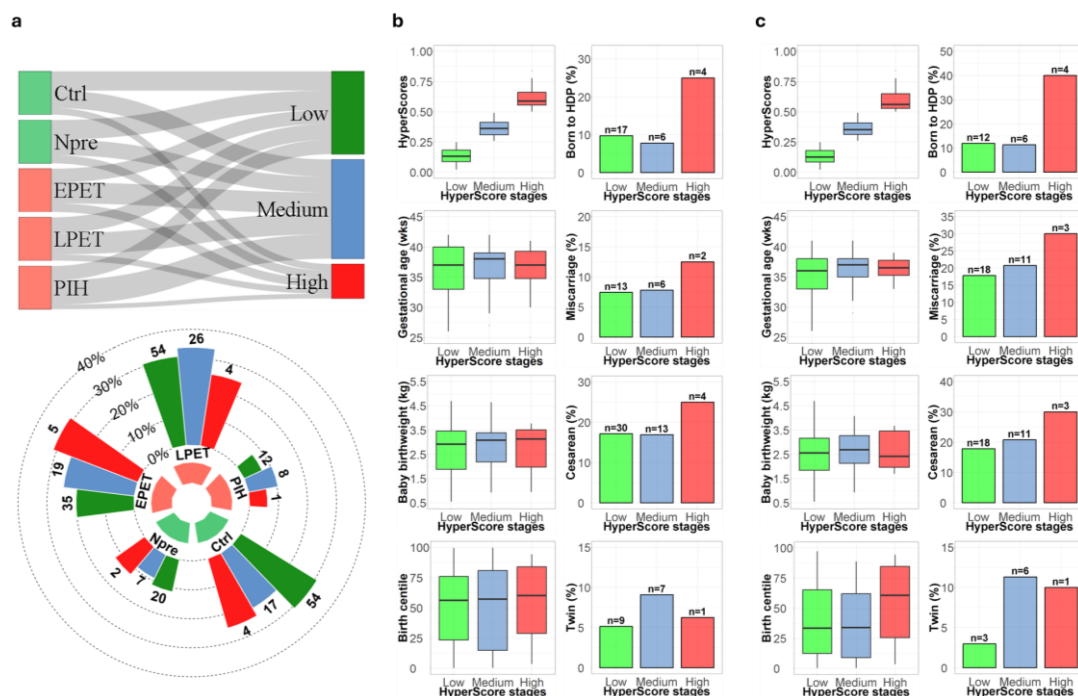


Figure 5.2 | Distribution of pregnancy subgroups and major pregnancy factors based on *HyperScore* stages. **a**, Proportion of pregnancy subgroups within the three *HyperScore* stages relative to total number of participants within subgroups (top Sankey plot) or within *HyperScore* stages (bottom polar plot). **b**, Analysis of major pregnancy factors for the overall PVS cohort. **c**, Analysis of major pregnancy factors for the hypertensive disorders of pregnancy (HDP) subgroups only.

being born to HDP, miscarriage/stillbirth, and caesarean delivery showed the highest proportions in the high-risk stage, at 25%, 12.5%, and 25%, respectively, while twin pregnancies had higher proportions in the medium-risk stage. A similar trend was observed when analysing the HDP subgroups separately (**Figure 5.2c**), with consistent patterns across all factors. Interestingly, birth centile demonstrated a higher median value in the high-risk stage compared with the other stages.

5.4.3 PVS and the *HyperTrajectory* allocation

PVS individuals were distributed across all trajectories on the *HyperTrajectory* map (**Figure 5.3a**). The largest proportion was observed in trajectory 2 ($n = 90$, 33.6%), while the smallest group was found in trajectory 6 ($n = 10$, 3.7%) relative to the total number of individuals in PVS. The distribution of individuals across trajectories also varied according to pregnancy subgroups. Trajectory 1 contained the highest proportion of NP individuals ($n = 25$, 52.1%), including both control and normotensive preterm groups, representing just over half of that trajectory's population. Trajectory 2 was characterised by the highest proportion of PIH cases ($n = 11$, 12.2%). Interestingly, the majority of individuals within trajectory 6 had combined PET ($n = 8$, 80%), while trajectory 3 was high for EPET cases and trajectory 4 for LPET cases. Trajectory 5 did not show a clear predominance of any pregnancy subgroup.

In terms of major pregnancy factors, the overall PVS cohort exhibited notable variations across the *HyperTrajectories* (**Figure 5.3b**). For *HyperScore*, trajectory 5 showed significantly higher values (0.37 [0.18-0.49]), followed by trajectory 3 (0.26 [0.17-0.39]) ($p = 0.030$). Gestational age was highest in trajectory 5 (39 weeks) and lowest in trajectory 3 (36 weeks) ($p < 0.001$). A similar pattern was observed for baby birthweight,

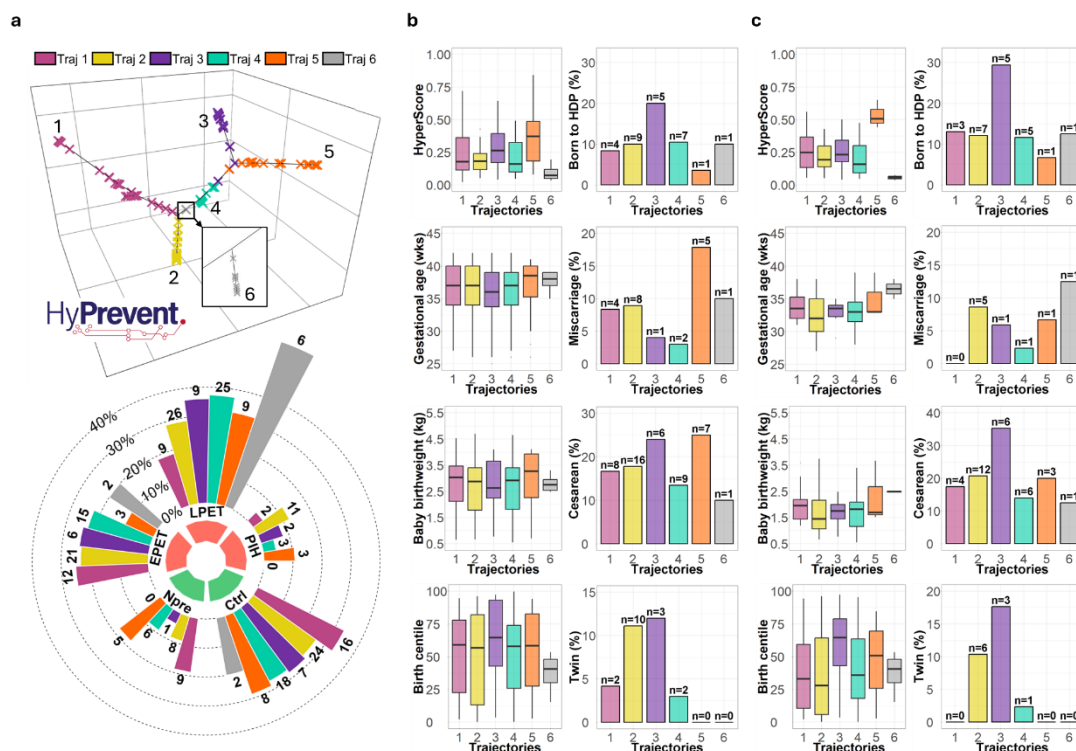


Figure 5.3 | Distribution of pregnancy subgroups and major pregnancy factors based on *HyperTrajectories*. **a.** Proportion of pregnancy subgroups within the three *HyperScore* stages relative to total number of participants within subgroups (top Sankey plot) or within *HyperScore* stages (bottom polar plot). **b.** Analysis of major pregnancy factors for the overall PVS cohort. **c.** Analysis of major pregnancy factors for the hypertensive disorders of pregnancy (HDP) subgroups only.

with the highest values in trajectory 5 (3.3 kg [2.3-3.9]) and the lowest in trajectory 3 (2.6 kg [2.2-3.7]). Regarding categorical factors, trajectory 3 had the highest proportion of participants with a maternal history of being born to HDP (20%), while trajectory 5 showed the highest proportion of miscarriages/stillbirths (17.9%). Both trajectories 3 and 5 recorded the highest caesarean section rates, at nearly 25% each. Notably, twin pregnancies were most common in trajectories 2 and 3, each with a proportion of nearly 12% ($p = 0.047$). The summary of overall cohort characteristics is provided in **Table 5.2**. Analysis based on HDP subgroups separately revealed similar trends with slight differences (**Figure 5.3c**). Most notably, trajectory 5 had a significantly higher *HyperScore* of 0.51 [0.47-0.58] ($p < 0.001$), while trajectory 2 had the lowest birth centile with 27.9 [5.8-64.3] ($p = 0.025$).

Table 5.2 | Characteristics of the *HyperTrajectory* in the overall PVS cohort.

<i>HyperTrajectory</i>	Overall (N = 268)	1 (N = 48)	2 (N = 90)	3 (N = 25)	4 (N = 67)	5 (N = 28)	6 (N = 10)	<i>p</i> -value (age adjusted)
<i>HyperScore</i>	0.186 [0.114– 0.326]	0.178 [0.113– 0.362]	0.181 [0.117– 0.239]	0.263 [0.170– 0.393]	0.160 [0.098– 0.324]	0.371 [0.182– 0.487]	0.072 [0.049– 0.118]	0.030
Maternal demographic information								
Age at recruitment, years	41.0 [38.0–45.0]	40.0 [37.0–45.0]	40.0 [37.0–44.0]	43.5 [39.0–45.0]	42.0 [39.0–46.0]	40.0 [37.0–43.2]	43.0 [38.0–45.0]	NA
SBP at recruitment, mmHg	117.0 [108.0– 125.5]	117.5 [111.0– 126.1]	115.0 [107.2– 124.0]	121.0 [113.5– 131.7]	115.0 [107.0– 126.2]	117.2 [110.5– 122.2]	118.5 [107.5– 124.5]	<0.001
DBP at recruitment, mmHg	76.0 [70.4–83.0]	77.2 [71.0–83.0]	74.8 [68.2–80.5]	79.5 [74.0–85.0]	77.5 [71.6–84.0]	75.8 [68.1–80.0]	79.0 [73.0–80.5]	<0.001
BMI at recruitment, kg/m ²	25.0 [22.5–29.0]	24.9 [22.9–27.2]	24.5 [21.9–27.9]	28.3 [22.5–31.8]	25.2 [22.1–30.2]	27.4 [24.7–31.3]	24.5 [21.3–26.0]	<0.001
Maternal birthweight, g	3174.0 [2813.2– 3430.0]	3033.0 [2721.0– 3195.8]	3174.0 [2863.0– 3486.0]	2990.5 [2608.0– 3274.5]	3174.0 [3068.0– 3402.0]	3260.0 [2920.0– 3656.5]	3118.0 [2927.0– 3288.0]	<0.001
Maternal born to HDP, <i>n</i>	27 (10.07%)	4 (8.33%)	9 (10.00%)	5 (20.00%)	7 (10.45%)	1 (3.57%)	1 (10.00%)	0.458
Miscarriage/stillbirth, <i>n</i>	21 (7.84%)	4 (8.33%)	8 (8.89%)	1 (4.00%)	2 (2.99%)	5 (17.86%)	1 (10.00%)	0.148
Multiple miscarriages/stillbirths, <i>n</i>	9 (3.36%)	1 (2.08%)	3 (3.33%)	1 (4.00%)	1 (1.49%)	2 (7.14%)	1 (10.00%)	0.637
Caesarean, <i>n</i>	47 (17.54%)	8 (16.67%)	16 (17.78%)	6 (24.00%)	9 (13.43%)	7 (25.00%)	1 (10.00%)	0.644
Multiple caesareans, <i>n</i>	14 (5.22%)	4 (8.33%)	4 (4.44%)	3 (12.00%)	2 (2.99%)	1 (3.57%)	0 (0.00%)	0.459
Smoker at recruitment, <i>n</i>	22 (8.21%)	5 (10.42%)	5 (5.56%)	1 (4.00%)	5 (7.46%)	5 (17.86%)	1 (10.00%)	0.462
Gestational information								
Gestational age, weeks	37.0 [34.0–40.0]	37.0 [34.0–40.0]	36.8 [34.0–40.0]	36.0 [33.8–39.0]	37.0 [34.0–39.0]	38.5 [35.2–40.0]	38.0 [37.0–39.0]	<0.001
Baby birthweight, g	2952.5 [1983.8– 3468.2]	3031.5 [2117.8– 3468.8]	2873.0 [1782.8– 3401.2]	2627.5 [2243.8– 3652.8]	2914.0 [1814.0– 3400.0]	3266.0 [2261.2– 3927.0]	2747.0 [2537.0– 2950.0]	<0.001
Birth centile, percentile	56.6 [20.9–79.9]	59.0 [22.5–77.9]	56.6 [13.1–82.0]	60.7 [52.2–92.8]	57.9 [25.8–73.9]	58.4 [27.4–82.7]	40.6 [30.1–47.9]	0.092
Age at delivery, years	33.0 [29.8–37.0]	33.0 [29.0–36.0]	35.0 [29.0–37.0]	35.5 [30.5–37.0]	32.0 [30.2–37.8]	31.0 [26.8–34.2]	35.0 [34.0–35.0]	<0.001
SBP at booking, mmHg	110.0 [101.0– 120.0]	110.0 [110.0– 120.0]	110.0 [100.0– 120.0]	112.0 [110.0– 122.5]	110.0 [100.0– 120.0]	110.0 [100.0– 120.0]	103.0 [100.0– 125.0]	<0.001
DBP at booking, mmHg	68.0 [60.0–71.5]	70.0 [64.0–75.0]	68.0 [60.0–70.5]	70.0 [60.0–74.5]	69.0 [60.0–70.0]	70.0 [60.0–75.8]	60.0 [60.0–68.2]	<0.001
SBP at 3 rd trimester, mmHg	120.0 [110.0– 130.0]	120.0 [110.0– 130.0]	120.5 [110.0– 132.0]	122.0 [110.0– 129.0]	119.0 [110.0– 130.0]	120.0 [130.0– 130.0]	109.5 [110.0– 130.0]	<0.001
DBP at 3 rd trimester, mmHg	70.0 [65.0–80.0]	70.0 [66.0–80.0]	71.5 [60.8–82.5]	75.0 [70.0–83.0]	70.0 [60.0–82.0]	70.0 [68.0–78.0]	73.0 [61.8–88.5]	<0.001
BMI at booking, kg/m ²	24.0 [22.0–27.3]	24.0 [22.0–27.0]	23.7 [21.1–25.9]	27.0 [22.0–30.0]	24.0 [22.0–28.0]	25.0 [22.4–31.6]	24.0 [20.6–25.5]	<0.001
Twin pregnancy, <i>n</i>	17 (6.34%)	2 (4.17%)	10 (11.11%)	3 (12.00%)	2 (2.99%)	0 (0.00%)	0 (0.00%)	0.047
Female babies, <i>n</i>	66 (24.63%)	9 (18.75%)	25 (27.78%)	4 (16.00%)	23 (34.33%)	4 (14.29%)	1 (10.00%)	0.512
Smoker at pregnancy, <i>n</i>	21 (7.84%)	3 (6.25%)	6 (6.67%)	3 (12.00%)	7 (10.45%)	2 (7.14%)	0 (0.00%)	0.693
Pregnancy-related groups								
NP, <i>n</i>	104 (38.81%)	25 (52.08%)	32 (35.56%)	8 (32.00%)	24 (35.82%)	13 (46.43%)	2 (20.00%)	NA
HDP, <i>n</i>	164 (61.94%)	23 (47.92%)	58 (64.44%)	17 (68.00%)	43 (64.18%)	15 (53.57%)	8 (80.00%)	NA
Pregnancy-related subgroups								
Control, <i>n</i>	75 (27.99%)	16 (33.33%)	24 (26.67%)	7 (28.00%)	18 (26.87%)	8 (28.57%)	2 (20.00%)	NA
Normotensive preterm, <i>n</i>	29 (10.82%)	9 (18.75%)	8 (8.89%)	1 (4.00%)	6 (8.96%)	5 (17.86%)	0 (0.00%)	NA
EPET, <i>n</i>	59 (22.01%)	12 (25.00%)	21 (23.33%)	6 (24.00%)	15 (22.39%)	3 (10.71%)	2 (20.00%)	NA
LPET, <i>n</i>	84 (31.34%)	9 (18.75%)	26 (28.89%)	9 (36.00%)	25 (37.31%)	9 (32.14%)	6 (60.00%)	NA
PET, <i>n</i>	143 (53.35%)	21 (43.75%)	47 (52.22%)	15 (60.00%)	40 (59.70%)	12 (42.85%)	8 (80.00%)	NA
PIH, <i>n</i>	21 (7.84%)	2 (4.17%)	11 (12.22%)	2 (8.00%)	3 (4.48%)	3 (10.71%)	0 (0.00%)	NA

5.4.4 Odds ratios of pregnancy factors

Analysis of odds ratios revealed several significant associations across the *HyperTrajectories* (Figure 5.4). Trajectory 1 was significantly associated with increased odds of NP (OR = 1.92 [1.04-3.57], $p = 0.038$) and normotensive preterm outcomes (OR = 2.39 [1.05-5.47], $p = 0.038$), while showing significantly lower odds for HDP. Trajectory 2 was linked to a significantly higher likelihood of twin deliveries (OR = 2.92 [1.13-7.54], $p = 0.027$). Trajectory 3 showed significantly increased odds of a maternal history of being born to HDP (OR = 2.76 [1.01-7.52], $p = 0.047$) and multiple caesarean deliveries (OR = 3.38 [1.01-11.32], $p = 0.048$). Trajectory 4 was significantly associated with delivering female babies (OR = 1.93 [1.06-3.50], $p = 0.031$). Finally, trajectory 5 was associated with significantly increased odds of miscarriage/stillbirth (OR = 3.31 [1.19-9.19], $p = 0.022$).

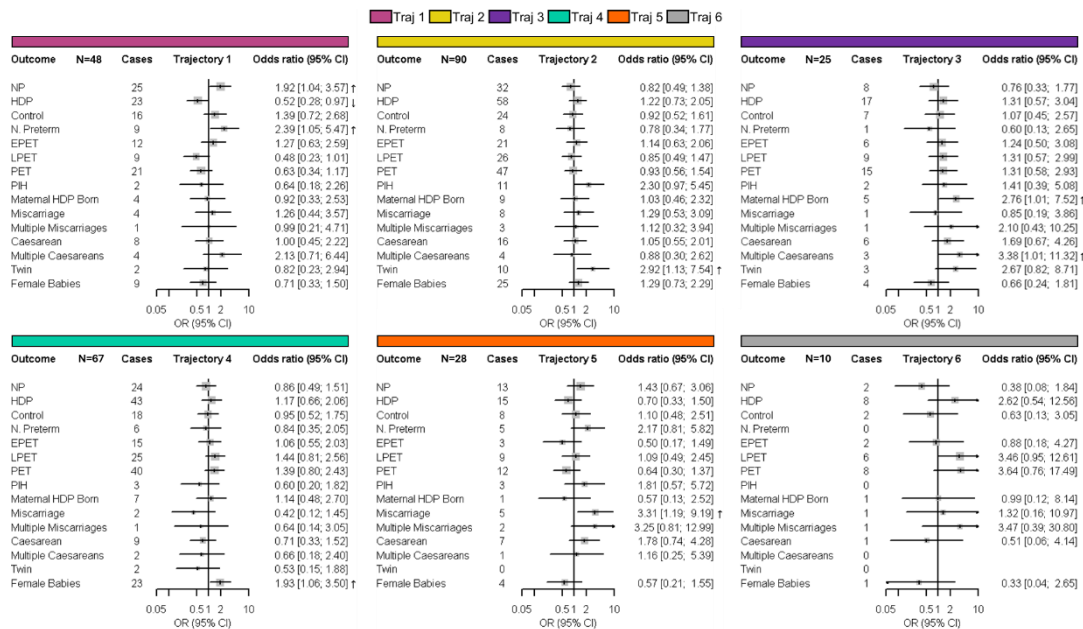


Figure 5.4 | Odds ratios (OR) for pregnancy subgroups distribution and major pregnancy factors across *HyperTrajectories* in PVS. Significantly high and low OR are labelled by upward and downward arrows, respectively.

5.4.5 Predominant characteristics of identified pregnancy phenotypes

Summarising *HyperTrajectories* based on pregnancy factors and odds ratios allowed for identifying unique phenotypical patterns (**Table 5.3**). Specifically, trajectory 1, the normotensive preterm phenotype, was characterised by a predominance of NP, particularly influenced by a higher proportion of normotensive preterm cases with significantly higher odds. Trajectory 2, the low birth centile phenotype, had the lowest birth centile values, which were associated with significantly higher odds of twin pregnancies. Trajectory 3, the family history phenotype, included more cases with significantly higher odds of maternal being born to HDP, linking the lowest maternal birthweight to the lowest baby birthweight. Participants in this trajectory also had higher maternal age, blood pressure, BMI, and caesarean section rates. Trajectory 4, the late onset phenotype, was characterised by the highest proportions of LPET and high proportions for overall PET, alongside significantly higher odds of having female babies. Trajectory 5, the foetal loss phenotype, showed the highest proportions of miscarriages/stillbirths and caesarean sections, with significantly increased odds of miscarriage/stillbirth. Trajectory 6 included a small number of individuals and did not exhibit a distinct phenotype.

Table 5.3 | Overall summary of the HyperTrajectories based on pregnancy factors in PVS.

Trajectory	Total participants	Predominant characteristic	Key phenotype
1	N = 48	Normotensive preterm	- Highest proportion of control - Highest proportion of normotensive preterm - Significantly high odds ratio for normotensive pregnancies - Significantly high odds ratio for normotensive preterm
2	N = 90	Low birth centile	- Second-lowest baby birthweight - Second-lowest gestational age - Lowest birth centile - Second-highest proportion of twin pregnancy - Significantly high odds ratio for twin pregnancy - Highest proportion of PIH
3	N = 25	Family history	- Highest age at recruitment/delivery - Highest SBP/DBP at recruitment/booking/3rd trimester - Highest BMI at recruitment/booking - Lowest maternal birthweight - Lowest baby birthweight - Highest proportion of maternal born to HDP - Significantly high odds ratio for maternal born to HDP - Second-highest proportion of caesareans - Highest proportion of multiple caesareans - Significantly high odds ratio of multiple caesareans
4	N = 67	Female babies	- Highest proportion of female babies - Significantly high odds ratio for having female babies - Highest proportion of LPET
5	N = 28	Foetal loss	- Lowest proportion of maternal born to HDP - Highest proportion of miscarriage/stillbirths - Highest proportion of multiple miscarriages/stillbirths - Significantly high odds ratio for miscarriage/stillbirths - Highest proportion of caesareans - Lowest proportion of twin pregnancy (0%)

5.4.6 Reproducible phenotypes on UKB pregnancy subset

The UK Biobank pregnancy subset included a total of 1,910 participants, including 955 controls and 955 individuals with a history of HDP. Among those with HDP, 437 participants had PET and 518 had PIH. At the time of the imaging visit, participants had a median age of 54 years [47—62]. Their blood pressure measurements were relatively

Table 5.4 | Characteristics of the *HyperTrajectory* in the UK Biobank pregnancy subset.

<i>HyperTrajectory</i>	Overall (N = 1,910)	1 (N = 377)	2 (N = 667)	3 (N = 232)	4 (N = 404)	5 (N = 156)	6 (N = 74)
<i>HyperScore</i>	0.16 [0.11-0.26]	0.20 [0.13-0.40]	0.16 [0.12-0.21]	0.21 [0.14-0.34]	0.13 [0.09-0.22]	0.22 [0.13-0.37]	0.05 [0.03-0.08]
Maternal demographic information							
Age, years	54.00 [47.00-62.00]	54.00 [48.00-61.00]	54.00 [48.00-62.00]	54.00 [47.00-61.00]	54.00 [46.00-61.00]	54.00 [46.75-62.00]	51.00 [45.25-59.00]
SBP, mmHg	117.00 [110.00-132.00]	117.00 [110.75-134.25]	117.00 [110.00-129.75]	116.00 [110.00-131.00]	117.00 [111.00-132.75]	118.00 [110.00-137.00]	118.50 [111.00-134.75]
DBP, mmHg	73.00 [66.00-83.00]	75.00 [67.00-85.00]	73.00 [66.00-81.00]	74.00 [67.00-82.50]	73.50 [66.00-83.00]	74.00 [67.50-86.50]	74.00 [65.00-81.00]
BMI, kg/m ²	23.80 [21.80-26.40]	23.90 [21.60-27.20]	23.80 [21.90-26.08]	24.20 [21.80-27.45]	23.60 [21.60-26.50]	23.50 [22.15-26.00]	23.40 [21.60-25.20]
Gestational information							
Miscarriage/stillbirth, <i>n</i>	606 (31.73%)	119 (31.56%)	198 (29.69%)	70 (30.17%)	129 (31.93%)	69 (44.23%)	21 (28.38%)
Multiple miscarriages/stillbirths, <i>n</i>	201 (10.52%)	44 (11.67%)	69 (10.34%)	25 (10.78%)	34 (8.42%)	19 (12.18%)	10 (13.51%)
Caesarean, <i>n</i>	54 (2.83%)	11 (2.92%)	20 (3.00%)	9 (3.88%)	11 (2.72%)	2 (1.28%)	1 (1.35%)
Twin, <i>n</i>	47 (2.46%)	10 (2.65%)	20 (3.00%)	4 (1.72%)	9 (2.23%)	4 (2.56%)	0 (0.00%)
Female babies, <i>n</i>	465 (24.35%)	103 (27.32%)	147 (22.04%)	43 (18.53%)	115 (28.47%)	37 (23.72%)	20 (27.03%)
Pregnancy-related subgroups							
Control, <i>n</i>	955 (50.00%)	175 (46.42%)	354 (53.07%)	127 (54.74%)	189 (46.78%)	74 (47.44%)	36 (48.65%)
PET, <i>n</i>	437 (22.88%)	92 (24.40%)	146 (21.89%)	43 (18.53%)	101 (25.00%)	39 (25.00%)	16 (21.62%)
PIH, <i>n</i>	518 (27.12%)	110 (29.18%)	167 (25.04%)	62 (26.72%)	114 (28.22%)	43 (27.56%)	22 (29.73%)

low, with a median SBP of 117 mmHg (110—132 mmHg) and a median DBP of 73 mmHg (66—83 mmHg). Detailed characteristics are presented in **Table 5.4**. Across *HyperTrajectories*, age and blood pressure did not differ substantially. However, several notable characteristics were observed. Trajectory 2 was characterised by the highest proportion of twin deliveries ($n = 3$, 20%). Trajectory 3 exhibited the highest BMI levels (24.2 kg/m² [21.8—27.5]) and the greatest proportion of caesarean section deliveries ($n = 9$, 3.9%). Trajectory 4 had the highest proportion of PET ($n = 101$, 25%). Trajectory 5 was associated with the highest proportion of miscarriage/stillbirth ($n = 69$, 44.2%) and relatively the highest *HyperScore* values (0.22 [0.13—0.37]).

In terms of odds ratios (**Figure 5.6**), trajectories 4 and 5 exhibited patterns broadly consistent with those observed in the PVS cohort. Specifically, trajectory 4 showed significantly higher odds of delivering a female baby (OR = 1.32 [1.03—1.69], $p =$

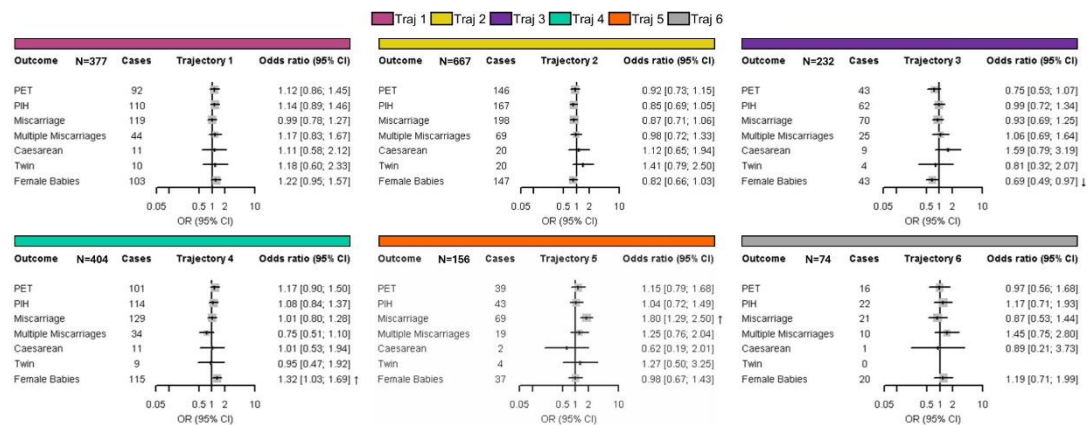


Figure 5.5 | Odds ratios (OR) for available pregnancy subgroups distribution and major pregnancy factors across *HyperTrajectories* in the UK Biobank pregnancy subset. Significantly high and low OR are labelled by upward and downward arrows, respectively.

0.028), while trajectory 5 was associated with significantly increased odds of miscarriage/stillbirth (OR = 1.80 [1.29—2.50], $p < 0.001$). Interestingly, although not statistically significant, trajectory 2 demonstrated a trend of higher odds of twin delivery (OR = 1.41 [0.79—2.50]), and trajectory 3 showed increased odds of caesarean delivery (OR = 1.59 [0.79—3.19]). These findings align with trends previously observed in the PVS cohort. The comparison between the two datasets using ROOR analysis revealed several notable observations (**Figure 5.7**). Here, factors positioned closer to the central line ($\log(\text{ROOR}) = 0$, $\text{ROOR} = 1$) indicate similar associations across datasets, whereas deviations to the right or left reflect stronger odds ratios in favour of the PVS cohort or the UK Biobank pregnancy subset, respectively. Black dots represent factors where the odds ratios in both datasets were in the same direction (either both positive or both negative), while open dots indicate disagreement. Across the *HyperTrajectories*, most pregnancy factors showed no significant differences between the two datasets, particularly for the major characteristics (highlighted in red). The only significantly different factor was PIH in trajectory 2 ($p = 0.028$). Moreover, trajectories 2, 3, and 5 had positive and relatively high median $\log(\text{ROOR})$ values (0.396, 0.355, and 0.581,

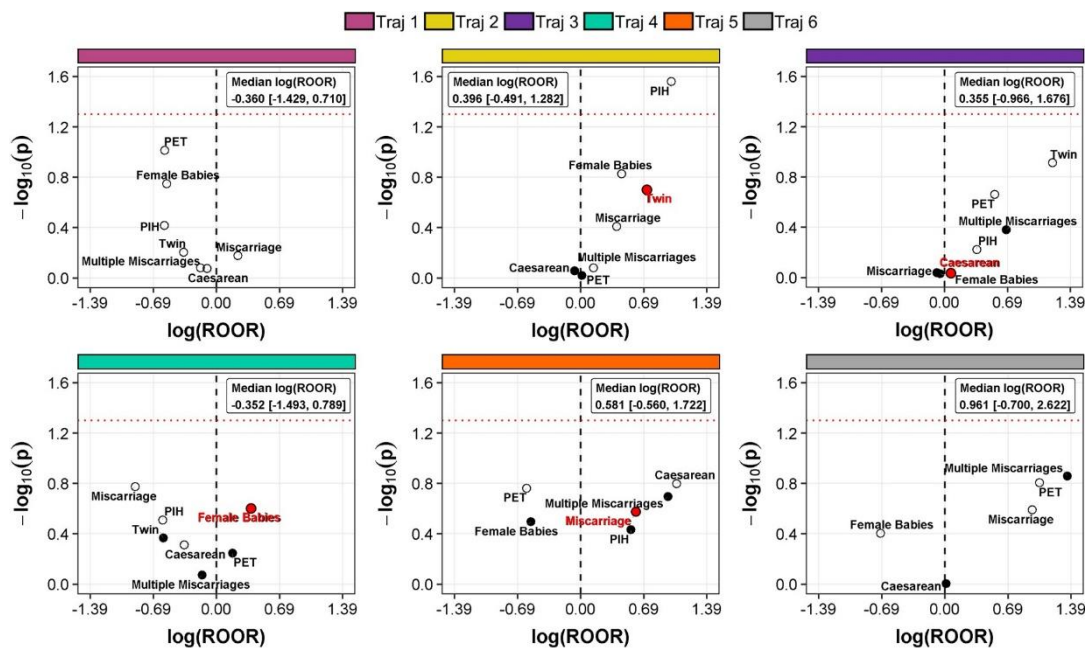


Figure 5.6 | Ratio of odds ratios (ROOR) analysis between PVS and the UK Biobank pregnancy subset across *HyperTrajectories*. Pregnancy factors positioned closer to the central line ($\log(\text{ROOR}) = 0$, $\text{ROOR} = 1$) indicate similar associations across datasets, whereas deviations to the right or left reflect stronger odds ratios in favour of the PVS cohort or the UK Biobank pregnancy subset, respectively. Black dots represent factors where the odds ratios in both datasets were in the same direction (either both positive or both negative), while open dots indicate disagreement. Red dots are key factors for each trajectory.

respectively), suggesting that pregnancy associations tended to be stronger in the PVS cohort. In contrast, trajectories 1 and 4 showed higher associations in the UK Biobank pregnancy subset, as indicated by their negative median $\log(\text{ROOR})$ values (-0.360 and -0.352, respectively). Importantly, key factors across *HyperTrajectories* were represented by black dots, indicating similar odds directionality between the two datasets.

5.5 Discussion and conclusion

In this chapter, I explored the ability of *HyTwin* to identify unique pregnancy phenotypes, particularly in the context of HDP. The findings suggest that organ alterations related to hypertensive disease, observed years after pregnancy, associate with pregnancy-specific characteristics, as illustrated by individuals clustered along *HyperTrajectories*, each

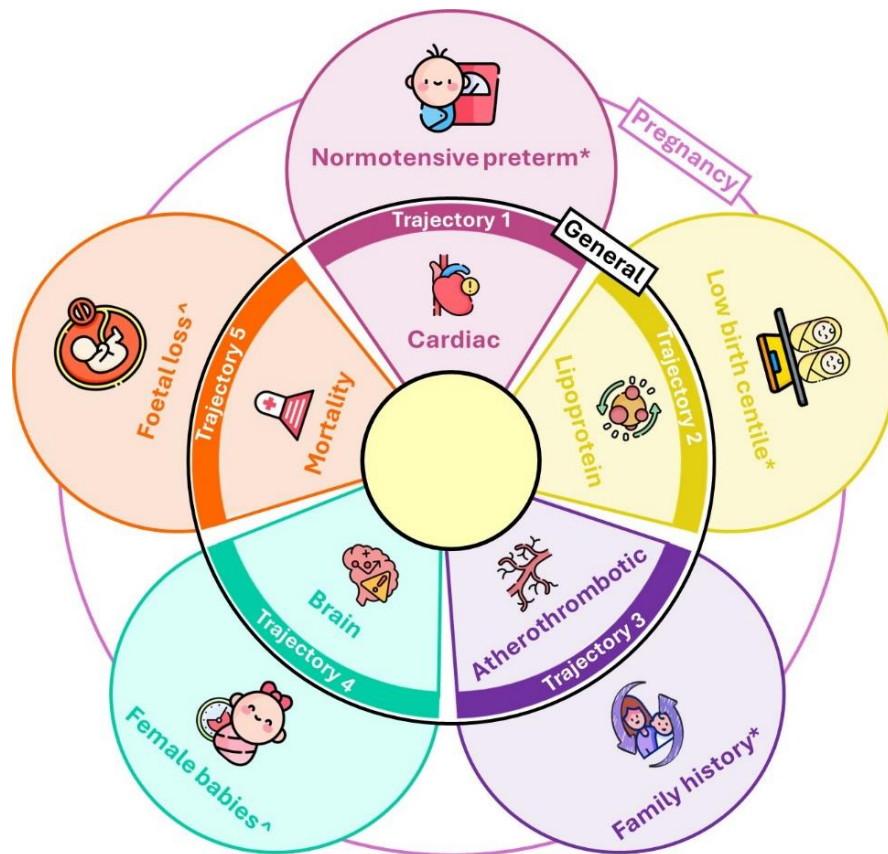


Figure 5.7 | Overall summary of the *HyperTrajectories* linking general hypertensive organ damage characteristics by pregnancy-related phenotypes. The inner characteristics were defined originally according to organ damage and outcomes in the *HyTwin* model. The outer characteristics were defined based on pregnancy-related factors in the overall PVS cohort. * Indicates that these predominant characteristics could not be validated in the UK Biobank pregnancy subset due to data unavailability. ^ Indicates that these predominant characteristics were available and successfully reproduced in the UK Biobank pregnancy subset.

defined by unique predominant features related to their pregnancy. Furthermore, this study demonstrated that females exhibiting greater end-organ damage later in life, as indicated by elevated *HyperScores*, were more likely to have experienced miscarriages/stillbirths, or caesarean section deliveries, or to have been born to a mother previously affected by HDP.

The predominant pregnancy-based phenotypes were associated with the previously defined characteristics of the *HyperTrajectories* (**Figure 5.8**). In particular, women with higher levels of organ damage and a more severe pregnancy phenotype were clustered along trajectory 5. Interestingly, this trajectory exhibited an end-organ profile considered

severe in the original *HyTwin* model and was associated with poorer outcomes, predominantly higher mortality rates. The pregnancy characteristics of these individuals indicated a greater likelihood of miscarriages, stillbirths, and caesarean section deliveries. Existing evidence in the literature supports a strong association between foetal loss and increased maternal mortality later in life, particularly in cases of recurrent pregnancy loss²²⁸. Moreover, as this trajectory was also originally characterised by renal abnormalities, previous studies have shown that stillbirths are linked to a two-fold increase in the risk of adverse renal outcomes²²⁹. Consistent with these observations, individuals within this trajectory exhibited the highest *HyperScores*, reflecting greater organ damage, which was similarly associated with increased rates of miscarriage and caesarean section during the analysis of high-risk *HyperScore* stage.

Among the distinct pregnancy phenotypes identified, trajectory 3 showed evidence of family history associations which were characterised by maternal history of being born to a mother with HDP and low birthweights of themselves and their babies. This observation aligns with emerging evidence suggesting that a family history of HDP constitutes a significant risk factor for developing HDP, particularly PET, in subsequent generations^{230–232}. Such patterns may reflect shared genetic predispositions, vascular maladaptation, or epigenetic influences that persist across maternal generations. However, these findings still require further validation thorough genetic analyses. Furthermore, participants within this trajectory showed increased blood pressure, BMI, and maternal age, which could be associated with the atherothrombotic phenotype described in the original *HyTwin* model.

The identified pregnancy-related phenotypes could trigger the early monitoring of other individuals with similar characteristics, enabling preventive strategies against the development of end-organ damage and related adverse outcomes. For instance,

individuals within trajectory 1 exhibited a predominantly normotensive preterm phenotype, which, in the original *HyTwin* model, was associated with cardiac remodelling later in life. There is growing evidence that preterm delivery independently increases the long-term risk of cardiovascular disease²³³ and is associated with up to a two-fold higher risk of cardiovascular-related mortality²³⁴. Moreover, trajectory 2 was characterised by a higher likelihood of twin pregnancies and lower birth centiles. In the original *HyTwin* model, this trajectory corresponded to the lipoprotein metabolic abnormality phenotype. The literature has shown several studies indicating that serum lipid concentrations in twin pregnancies tend to rise progressively throughout gestation^{235,236}. Thus, understanding mid-life abnormalities and their related pregnancy characteristics highlights the need for tailored management in such pregnancies.

The ability to reproduce findings in another cohort demonstrates that the *HyTwin* model effectively captures organ function dynamics later in life in relation to pregnancy. In this study, the model showed similar significant odds ratios for certain trajectories, namely trajectories 4 and 5, which were associated with predominant female births and miscarriage/stillbirth, respectively. It is worth noting that these characteristics were identified in two different cohorts. Specifically, the PVS cohort consisted of younger individuals with less severe organ damage, whereas the UK Biobank cohort, which was not originally designed for pregnancy studies, included older individuals with potentially higher levels of organ damage. Furthermore, pregnancies in the UK Biobank cohort had occurred much longer before the recruitment compared with those in the PVS cohort (which were 5—10 years post-pregnancy). Despite these differences, the model still captured these phenotypes accurately and reproduced them.

Validation of phenotypes across all *HyperTrajectories* was limited by the absence of relevant pregnancy-specific characteristics in the UK Biobank dataset. For instance,

trajectories 1 and 3 could not be fully validated due to missing information on normotensive preterm deliveries and maternal history of being born to HDP. Furthermore, despite the lack of birth centile calculations in trajectory 2 to directly confirm the findings observed in the PVS cohort, this trajectory showed higher odds of twin pregnancies across both datasets. These results require further evaluation using pregnancy-specific datasets, with the potential for these findings to be reproduced.

5.5.1 Study limitations

This study has several limitations. First, although distinct associations were observed between pregnancy characteristics and mid-life organ damage, the sample size in the PVS study was relatively small. A larger dataset would strengthen the evidence supporting these findings. Second, while key phenotypes were successfully reproduced in the UK Biobank, overall validation across all phenotypes was not feasible due to the limited availability of pregnancy-specific variables in the UK Biobank. Moreover, the UK Biobank is not primarily designed for pregnancy-focused research, necessitating further validation using pregnancy-specific cohorts to ensure accurate reproduction of these phenotypes. Third, both cohorts primarily consisted of individuals of European white ancestry, which may limit the generalisability of the findings to other populations. Lastly, although organ damage observed later in life may reflect phenotypes originating during pregnancy, further investigation into organ alterations that occur during gestation could enhance understanding of how these long-term changes develop.

5.5.2 Conclusions

In conclusion, the *HyTwin* model has been reliably applied to another blood pressure-related domain to characterise organ damage and link it to the predefined disease-

specific phenotypes. Based on patterns of organ damage observed later in life, *HyTwin* identified five distinct pregnancy-related phenotypes that may contribute to these alterations. These findings pave the way for improved management of hypertensive pregnancies exhibiting such characteristics, with the aim of preventing outcomes in later life. Future work should focus on validating these phenotypes in larger pregnancy-specific cohorts, incorporating both organ function assessments and genetic analyses.

6. CONCLUSIONS

In conclusion, this thesis has provided a strong foundation for the *HyTwin* model, which provides a comprehensive and integrative framework for modelling multi-organ alterations induced by hypertension. Through a contrastive machine learning approach, organ damage was characterised by defining two key components: (i) the *HyperScore*, a global risk score quantifying the global end-organ damage caused by hypertension, and (ii) the *HyperTrajectory*, a set of disease progression phenotypes associated with organ-specific adverse outcomes. These two components were internally validated using the model's large-scale baseline data derived from the UK Biobank. Furthermore, external testing was conducted using out-of-domain data from the ARIC study, which included participants from another country and represented a more diverse cohort compared with the baseline population. Lastly, the applicability of the *HyTwin* model within a closely related, blood pressure-based domain, was evaluated, revealing distinct hypertensive disorders of pregnancy (HDP) phenotypes based on mid-life organ alterations.

The first thesis objective (Chapter 2) was to systematically examine the current state-of-the-art in the use of AI algorithms for predicting, phenotyping, and managing hypertension. Here, I conducted a systematic review and meta-analysis by focusing on studies that employed medical imaging to investigate organ-specific functional and structural alterations to characterise the disease. I found that most existing studies focus primarily on predicting hypertension occurrence by comparing organ information between hypertensive and control groups, with the majority lacking external testing. Furthermore, there was a notable absence of models that integrate multi-organ information to provide a holistic understanding of disease manifestation. Most importantly, there remains a need for further discovery research that leverages existing knowledge of hypertension-induced organ alterations to identify organ-based phenotypes and enable personalised, targeted assessment.

The second thesis objective (Chapter 3) was to develop the *HyTwin* model utilising knowledge of current gaps in the literature, which was designed to capture the spectrum of hypertensive multi-organ damage severity and to predict organ-specific adverse outcomes phenotypes. By leveraging contrastive knowledge between healthy normotensive individuals and those with severe hypertension, the model mapped the continuum of organ damage, ranging from normal (no damage) to severe damage caused by hypertension. This contrastive knowledge was established through a contrastive trajectory inference approach, which enabled the derivation of two key components: a global organ damage risk score (*HyperScore*) and an individual's position along one of six major outcome pathways (*HyperTrajectory*). Both components demonstrated high accuracy and robustness across multiple cross-validation folds with low errors. Notably, individuals located further along the *HyperTrajectories* exhibited higher organ damage severity, as reflected by elevated *HyperScores*, and eventually worse outcomes.

The third thesis objective (Chapter 4) was to evaluate the generalisability of the *HyTwin* model by applying it to an external cohort and assessing whether its performance and findings remained consistent. The results showed that both components, the *HyperScore* and *HyperTrajectory*, effectively predicted organ damage severity and phenotypes, respectively, with minimal divergence from the baseline findings. This was demonstrated through analyses of *HyperScore* distribution divergence, outcome proportions of individuals positioned along the *HyperTrajectories*, and organ-specific event-free probability analyses conducted on those trajectories, comparing between UK Biobank and ARIC datasets.

The final objective of this thesis (Chapter 5) was to investigate how an organ damage-based model, the *HyTwin*, could provide insights into distinct phenotypes of another disease that induces organ damage due to elevation in blood pressure. Here, I focused

on HDP using data recorded locally within the lab facilities, and I observed that organ damage later in life could reflect lifestyle, demographic, and pregnancy-related factors from years earlier that could be potentially associated with it. The analysis revealed distinct pregnancy-based phenotypes that may be linked to subsequent cardiovascular outcomes and mid-life organ damage. Notably, a severe pregnancy phenotype was identified, characterised by a higher incidence of foetal loss, which, when traced back using the baseline model, aligned with the high-mortality trajectory and had highest risk scores.

6.1 Future questions and research directions

My thesis work laid the foundation for a hypertensive organ damage-based model, the *HyTwin*, and introduced two key components: the *HyperScore* and *HyperTrajectory*. However, several research questions remain to be investigated.

1. How does hypertension-induced organ damage progress over the years following the initial modelling period?

In this thesis, individuals with higher *HyperScores* exhibited greater organ damage at the time of their imaging visit and recruitment. Furthermore, they were positioned at different points along the *HyperTrajectory*, which was associated with varying disease outcomes. However, there is currently no evidence on how these individuals' organ damage scores or their positions along the same or different trajectories might change over time. Existing literature indicates that delays in initiating antihypertensive therapy are associated with worse outcomes and increased organ damage²³⁷. A longitudinal study examining the progression

of organ damage over time, with and without medication and lifestyle management, is required to address this question. I am interested in extending this work by analysing follow-up visits from the UK Biobank and ARIC cohorts, mapping them back onto the *HyTwin* model to observe longitudinal changes in *HyperScore* and *HyperTrajectory*, and examining how these changes relate to participants' lifestyle factors, medication use, and organ damage characteristics. My advisors and I are submitting *HyTwin* for funding to investigate longitudinal changes of organ damage and the influence of medication and lifestyle factors on its predictions and on the progression or deterioration of organ damage. This will allow us to assess whether factors such as antihypertensive medications, sleep patterns, or physical activity habits modify the long-term severity of organ damage. Currently, *HyTwin* does not account for the temporal dimension, *i.e.*, the duration of hypertension exposure following recruitment or the effects of ageing. To address this, I plan to model the temporal progression of organ damage by integrating information on hypertension exposure duration, follow-up periods, and age between visits. This temporal framework, referred to as *HyperAge*, will potentially provide insights into the physiological age of body organs in relation to their functional and structural alterations.

2. Are there any major ethnicity-based phenotypic variations observed in the *HyTwin* model?

During the external testing of the *HyTwin* model, its generalisability was assessed using the ARIC dataset, where it demonstrated strong performance. In this analysis, I evaluated whether the model's findings remained reproducible in a cohort comprising nearly 20% Black African American participants, in contrast to the predominantly White population of the UK Biobank. Despite this, further

investigation is still needed to determine whether these alterations exhibit specific patterns, severities, or dominant trajectories within the *HyTwin* model across different ethnic groups, such as Chinese or African populations. Recent evidence in the literature indicates that ethnicity plays a significant role in differentiating organ damage patterns in hypertension²³⁸. To achieve this, analysis of ethnically based organ differences would be required, utilising UK Biobank and ARIC.

I have managed to successfully integrate the *HyTwin* model and its associated algorithms into a functional web application capable of estimating *HyperScore* and predicting *HyperTrajectory* using MATLAB Web App Server. I am currently guiding colleagues on its application in related areas of research. At present, a colleague and I are investigating ethnic differences using *HyTwin* across various cohorts, including the UK Biobank, ARIC, and locally recorded datasets. My research interest lies in exploring how these differences influence the allocation of individuals within the identified trajectories.

3. What are the mechanisms of organ alterations that occur in hypertensive pregnancies during and shortly after the gestational period?

The *HyTwin* model revealed distinct pregnancy phenotypes by utilising pregnancy characteristics and mid-life organ alteration information. Despite these novel and reproducible findings, the proposed approach still lacks insight into the mechanisms of organ alterations occurring during the gestational period or immediately after delivery and how these manifest in the presence of hypertensive pregnancies. Evidence suggests that pregnancy-induced organ alterations emerge during gestation and may persist postpartum²³⁹. To address this gap, I am interested in utilising ongoing clinical trials (not yet finalised) in

our lab that provide both imaging and non-imaging measurements of organ alterations during gestation and postpartum. This work could establish a clearer link between the identified pregnancy phenotypes and organ alterations observed during pregnancy, immediately after delivery, and later in life, while associating this knowledge with the original *HyTwin* organ-specific phenotypes.

Moreover, I am collaborating with my advisor to apply *HyTwin* for predicting *HyperScore* and *HyperTrajectory* in a UK Biobank cohort, linking these predictions with the geometrical features of the heart, referred to as *HyperShape*. This emerging work aims to examine variations in cardiac shape and geometry across different trajectories and among individuals with higher risk scores.

Finally, *HyTwin* has been submitted for translational grant funding to support its progression from a laboratory prototype to a clinically validated medical software. Although *HyTwin* currently utilises routinely collected clinical data, it remains a research tool. Comprehensive real-world clinical testing is therefore essential to fully assess its potential impact on patient care. To bridge this gap, I plan to implement targeted enhancements in both technical performance and user functionality, developed through a co-design process involving clinicians, nurses, and patients. These improvements will be embedded within a rigorous software quality management framework. Accordingly, *HyTwin* will undergo prospective evaluation across seven NHS sites in real-world clinical environments. Data generated by *HyTwin* will be validated against detailed measures of end-organ change and analysed to determine its potential influence on patient management. In parallel, I will develop optimised deployment protocols to support software implementation across diverse clinical settings within the UK and internationally.

7. REFERENCES

1. Mills, K. T., Stefanescu, A. & He, J. The global epidemiology of hypertension. *Nature Reviews Nephrology* 2020 16:4 **16**, 223–237 (2020).
2. Harrison, D. G., Coffman, T. M. & Wilcox, C. S. Pathophysiology of Hypertension: The Mosaic Theory and Beyond. *Circ. Res.* **128**, 847–863 (2021).
3. Campbell, E. *et al.* Sociodemographic and health-related differences in undiagnosed hypertension in the health survey for England 2015–2019: a cross-sectional cohort study. *EClinicalMedicine* **65**, 102275 (2023).
4. Forouzanfar, M. H. *et al.* Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: A systematic analysis for the Global Burden of Disease Study 2013. *The Lancet* **386**, 2287–2323 (2015).
5. Health matters: combating high blood pressure - GOV.UK. <https://www.gov.uk/government/publications/health-matters-combating-high-blood-pressure/health-matters-combating-high-blood-pressure>.
6. Ferdinand, K. C. & Nasser, S. A. Management of Essential Hypertension. *Cardiol. Clin.* **35**, 231–246 (2017).
7. Franco, C. *et al.* Essential Hypertension and Oxidative Stress: Novel Future Perspectives. *International Journal of Molecular Sciences* 2022, Vol. 23, Page 14489 **23**, 14489 (2022).
8. Januszewicz, A. *et al.* Cardiac Phenotypes in Secondary Hypertension: JACC State-of-the-Art Review. *J. Am. Coll. Cardiol.* **80**, 1480–1497 (2022).
9. Whelton, P. K. *et al.* 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults a report of the American College of Cardiology/American Heart Association Task Force on Clinical practice guidelines. *Hypertension* **71**, E13–E115 (2018).
10. Johnson, H. M. *et al.* 2025 AHA/ACC/AANP/AAPA/ABC/ACCP/ACPM/AGS/AMA/ASPC/NMA/PCNA/SGIM 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 Joint Committee.... *Circulation* **152**, e114–e218 (2025).
11. Jones, N. R., McManus, R. J., McCormack, T. & Constanti, M. Diagnosis and management of hypertension in adults: NICE guideline update 2019. *British Journal of General Practice* **70**, 90–91 (2020).
12. Williams, B. *et al.* 2018 ESC/ESH Guidelines for the management of arterial hypertension: The Task Force for the management of arterial hypertension of the European Society of Cardiology (ESC) and the European Society of Hypertension (ESH). *Eur. Heart J.* **39**, 3021–3104 (2018).
13. Chobanian, A. V. *et al.* Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. *Hypertension* **42**, 1206–1252 (2003).
14. Kario, K., Thijs, L. & Staessen, J. A. Blood Pressure Measurement and Treatment Decisions: Masked and White-Coat Hypertension. *Circ. Res.* **124**, 990–1008 (2019).

15. Hall, J. E., Do Carmo, J. M., Da Silva, A. A., Wang, Z. & Hall, M. E. Obesity-Induced Hypertension: Interaction of Neurohumoral and Renal Mechanisms. *Circ. Res.* **116**, 991–1006 (2015).
16. Hegde, S. M. & Solomon, S. D. Influence of Physical Activity on Hypertension and Cardiac Structure and Function. *Curr. Hypertens. Rep.* **17**, 77- (2015).
17. Bailey, M. A. & Dhaun, N. Salt Sensitivity: Causes, Consequences, and Recent Advances. *Hypertension* **81**, 476–489 (2024).
18. Everett, B. & Zajacova, A. Gender differences in hypertension and hypertension awareness among young adults. *Biodemography Soc. Biol.* **61**, 1–17 (2015).
19. Chappell, L. C., Cluver, C. A., Kingdom, J. & Tong, S. Pre-eclampsia. *The Lancet* **398**, 341–354 (2021).
20. Alkhodari, M. *et al.* The role of artificial intelligence in hypertensive disorders of pregnancy: towards personalized healthcare. <https://doi.org/10.1080/14779072.2023.2223978> **21**, 531–543 (2023).
21. Collaboration, M.-U. *et al.* *Saving Lives, Improving Mothers' Care - Lessons Learned to Inform Maternity Care from the UK and Ireland Confidential Enquiries into Maternal Deaths and Morbidity 2017-19. National Perinatal Epidemiology Unit: University of Oxford* (National Perinatal Epidemiology Unit: University of Oxford, 2021).
22. Tranquilli, A. L., Brown, M. A., Zeeman, G. G., Dekker, G. & Sibai, B. M. The definition of severe and early-onset preeclampsia. Statements from the International Society for the Study of Hypertension in Pregnancy (ISSHP). *Pregnancy Hypertension: An International Journal of Women's Cardiovascular Health* **3**, 44–47 (2013).
23. Clark, S. L. *et al.* Gestational hypertension and “severe” disease: time for a change. *Am. J. Obstet. Gynecol.* **228**, 547–552 (2023).
24. Erez, O. *et al.* Preeclampsia and eclampsia: the conceptual evolution of a syndrome. *Am. J. Obstet. Gynecol.* **226**, S786–S803 (2022).
25. Hippisley-Cox, J., Coupland, C. & Brindle, P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *BMJ* **357**, (2017).
26. D'Agostino, R. B. *et al.* General cardiovascular risk profile for use in primary care: The Framingham heart study. *Circulation* **117**, 743–753 (2008).
27. Goff, D. C. *et al.* 2013 ACC/AHA guideline on the assessment of cardiovascular risk: A report of the American college of cardiology/American heart association task force on practice guidelines. *Circulation* **129**, 49–73 (2014).
28. Talha, I., Elkhoudri, N. & Hilali, A. Major Limitations of Cardiovascular Risk Scores. *Cardiovasc. Ther.* **2024**, 4133365 (2024).
29. Camm, A. J., Lüscher, T. F., Maurer, G. & Serruys, P. W. Target organ damage, cardiovascular disease risk, and clinical evaluation of the hypertensive patient. *The ESC Textbook of Cardiovascular Medicine* <https://doi.org/10.1093/MED/9780198784906.003.0565> (2018)
doi:10.1093/MED/9780198784906.003.0565.

30. Wang, X. *et al.* Hypertension-mediated organ damage and established cardiovascular disease in patients with hypertension: the China Hypertension Survey, 2012–2015. *Journal of Human Hypertension* 2021 36:12 **36**, 1092–1098 (2021).
31. Zhang, Y. & Moran, A. E. Trends in the prevalence, awareness, treatment, and control of hypertension among young adults in the United States, 1999 to 2014. *Hypertension* **70**, 736–742 (2017).
32. Choudhry, N. K. *et al.* Medication adherence and blood pressure control: A scientific statement from the american heart association. *Hypertension* **79**, E1–E14 (2022).
33. Leow, M. K. S. Environmental origins of hypertension: Phylogeny, ontogeny and epigenetics. *Hypertension Research* **38**, 299–307 (2015).
34. Segal, R. *et al.* Blood pressure variability and organ damage in a general population: Results from the PAMELA study. *Hypertension* **39**, 710–714 (2002).
35. Perrone-Filardi, P. *et al.* Non-invasive cardiovascular imaging for evaluating subclinical target organ damage in hypertensive patients: A consensus paper from the European Association of Cardiovascular Imaging (EACVI), the European Society of Cardiology Council on Hypertension, an.... *Eur. Heart J. Cardiovasc. Imaging* **18**, 945–960 (2017).
36. Williams, B. *et al.* 2018 ESC/ESH Guidelines for the management of arterial hypertensionThe Task Force for the management of arterial hypertension of the European Society of Cardiology (ESC) and the European Society of Hypertension (ESH). *Eur. Heart J.* **39**, 3021–3104 (2018).
37. Nwabuo, C. C. & Vasan, R. S. Pathophysiology of Hypertensive Heart Disease: Beyond Left Ventricular Hypertrophy. *Curr. Hypertens. Rep.* **22**, 11- (2020).
38. Yun, M. *et al.* Blood Pressure and Left Ventricular Geometric Changes: A Directionality Analysis. *Hypertension* **78**, 1259–1266 (2021).
39. Park, C. *et al.* Elevated blood pressure in adolescence is attributable to a combination of elevated cardiac output and total peripheral resistance evidence against a hyperkinetic State. *Hypertension* **72**, 1103–1108 (2018).
40. Hundley, W. G. *et al.* ACCF/ACR/AHA/NASCI/SCMR 2010 expert consensus document on cardiovascular magnetic resonance: A report of the american college of cardiology foundation task force on expert consensus documents. *Circulation* **121**, 2462–2508 (2010).
41. Jiang, M., Lu, M. & Zhao, S. Cardiac Functional Assessment by Magnetic Resonance Imaging. *Cardiology Discovery* **4**, 284–299 (2024).
42. Hoffmann, R. *et al.* Analysis of left ventricular volumes and function: A multicenter comparison of cardiac magnetic resonance imaging, cine ventriculography, and unenhanced and contrast-enhanced two-dimensional and three-dimensional echocardiography. *Journal of the American Society of Echocardiography* **27**, 292–301 (2014).
43. Mavrogeni, S. *et al.* The emerging role of Cardiovascular Magnetic Resonance in the evaluation of hypertensive heart disease. *BMC Cardiovascular Disorders* 2017 17:1 **17**, 132- (2017).

44. González, A. *et al.* Myocardial Interstitial Fibrosis in Hypertensive Heart Disease: From Mechanisms to Clinical Management. *Hypertension* **81**, 218–228 (2024).
45. Gottesman, R. F., Egle, M., Groechel, R. C. & Mughal, A. Blood pressure and the brain: the conundrum of hypertension and dementia. *Cardiovasc. Res.* **120**, 2360–2372 (2025).
46. Amier, R. P. *et al.* Hypertensive Exposure Markers by MRI in Relation to Cerebral Small Vessel Disease and Cognitive Impairment. *JACC Cardiovasc. Imaging* **14**, 176–185 (2021).
47. Pacholko, A. & Iadecola, C. Hypertension, Neurodegeneration, and Cognitive Decline. *Hypertension* **81**, 991–1007 (2024).
48. Iadecola, C. & Gottesman, R. F. Neurovascular and Cognitive Dysfunction in Hypertension: Epidemiology, Pathobiology, and Treatment. *Circ. Res.* **124**, 1025–1044 (2019).
49. Li, Z., Wang, W., Sang, F., Zhang, Z. & Li, X. White matter changes underlie hypertension-related cognitive decline in older adults. *Neuroimage Clin.* **38**, 103389 (2023).
50. Burchell, A. E. *et al.* Comprehensive First-Line Magnetic Resonance Imaging in Hypertension: Experience From a Single-Center Tertiary Referral Clinic. *J. Clin. Hypertens.* **19**, 13–22 (2017).
51. Laporte, J. P. *et al.* Investigating the association between human brainstem microstructural integrity and hypertension using magnetic resonance relaxometry. *Hypertension Research* 2025 48:4 **48**, 1564–1574 (2025).
52. Markus, H. S. & Erik de Leeuw, F. Cerebral small vessel disease: Recent advances and future directions. *International Journal of Stroke* **18**, 4–14 (2023).
53. Gronewold, J. *et al.* Association of Blood Pressure, Its Treatment, and Treatment Efficacy With Volume of White Matter Hyperintensities in the Population-Based 1000BRAINS Study. *Hypertension* **78**, 1490–1501 (2021).
54. Chetouani, A. *et al.* Cross-sectional variations of white and grey matter in older hypertensive patients with subjective memory complaints. *Neuroimage Clin.* **17**, 804–810 (2018).
55. Horowitz, B., Miskulin, D. & Zager, P. Epidemiology of Hypertension in CKD. *Adv. Chronic Kidney Dis.* **22**, 88–95 (2015).
56. Ku, E., Lee, B. J., Wei, J. & Weir, M. R. Hypertension in CKD: Core Curriculum 2019. *American Journal of Kidney Diseases* **74**, 120–131 (2019).
57. Kidney disease: A UK public health emergency - Kidney Research UK. <https://www.kidneyresearchuk.org/about-us/policy/health-economics-report/>.
58. McQuarrie, E. P. *et al.* Association Between Urinary Sodium, Creatinine, Albumin, and Long-Term Survival in Chronic Kidney Disease. *Hypertension* **64**, 111–117 (2014).
59. Peralta, C. A. *et al.* Detection of Chronic Kidney Disease With Creatinine, Cystatin C, and Urine Albumin-to-Creatinine Ratio and Association With Progression to End-Stage Renal Disease and Mortality. *JAMA* **305**, 1545–1552 (2011).

60. Lopez-Giacoman, S. & Madero, M. Biomarkers in chronic kidney disease, from kidney function to kidney damage. *http://www.wjgnet.com/* **4**, 57–73 (2015).
61. Arab, S. F. *et al.* Review of Renal Artery Stenosis and Hypertension: Diagnosis, Management, and Recent Randomized Control Trials. *Saudi Journal of Kidney Diseases and Transplantation* **33**, 147–159 (2022).
62. Cheung, C. Y. L., Ikram, M. K., Sabanayagam, C. & Wong, T. Y. Retinal microvasculature as a model to study the manifestations of hypertension. *Hypertension* **60**, 1094–1103 (2012).
63. Tan, W. *et al.* The New Era of Retinal Imaging in Hypertensive Patients. *Asia-Pacific Journal of Ophthalmology* **11**, 149–159 (2022).
64. Chua, J. *et al.* Impact of hypertension on retinal capillary microvasculature using optical coherence tomographic angiography. *J. Hypertens.* **37**, 572–580 (2019).
65. Sun, C. *et al.* Systemic hypertension associated retinal microvascular changes can be detected with optical coherence tomography angiography. *Scientific Reports* **2020 10:1** **10**, 9580- (2020).
66. Cheung, C. Y. L., Ikram, M. K., Sabanayagam, C. & Wong, T. Y. Retinal microvasculature as a model to study the manifestations of hypertension. *Hypertension* **60**, 1094–1103 (2012).
67. Erden, S. & Bicakci, E. Hypertensive retinopathy: Incidence, risk factors, and comorbidities. *Clin. Exp. Hypertens.* **34**, 397–401 (2012).
68. Dziedziak, J., Zaleska-Zmijewska, A., Szaflik, J. P. & Cudnoch-Jędrzejewska, A. Impact of Arterial Hypertension on the Eye: A Review of the Pathogenesis, Diagnostic Methods, and Treatment of Hypertensive Retinopathy. *Medical Science Monitor* **28**, 0–0 (2022).
69. Mancusi, C. *et al.* Carotid Atherosclerosis Predicts Blood Pressure Control in Patients With Hypertension: The Campania Salute Network Registry. *J. Am. Heart Assoc.* **11**, 22345 (2022).
70. Niu, L. *et al.* Impact of Multiple Cardiovascular Risk Factors on Carotid Intima-Media Thickness and Elasticity. *PLoS One* **8**, e67809 (2013).
71. Magnussen, C. G. Carotid artery intima-media thickness and hypertensive heart disease: a short review. *Clin. Hypertens.* **23**, 7 (2017).
72. Bots, M. L. & Sutton-Tyrrell, K. Lessons from the past and promises for the future for carotid intima-media thickness. *J. Am. Coll. Cardiol.* **60**, 1599–1604 (2012).
73. Ferreira, J. P. *et al.* Intima-media thickness is linearly and continuously associated with systolic blood pressure in a population-based cohort (STANISLAS cohort study). *J. Am. Heart Assoc.* **5**, (2016).
74. Johnson, K. B. *et al.* Precision Medicine, AI, and the Future of Personalized Health Care. *Clin. Transl. Sci.* **14**, 86–93 (2021).
75. Padmanabhan, S., Tran, T. Q. B. & Dominiczak, A. F. Artificial Intelligence in Hypertension: Seeing Through a Glass Darkly. *Circ. Res.* **128**, 1100–1118 (2021).

76. Quer, G., Arnaout, R., Henne, M. & Arnaout, R. Machine Learning and the Future of Cardiovascular Care: JACC State-of-the-Art Review. *J. Am. Coll. Cardiol.* **77**, 300–313 (2021).
77. Krittanawong, C. *et al.* Machine learning prediction in cardiovascular diseases: a meta-analysis. *Sci. Rep.* **10**, 16057- (2020).
78. De Marvao, A., Dawes, T. J. W., Howard, J. P. & O'Regan, D. P. Artificial intelligence and the cardiologist: what you need to know for 2020. *Heart* **106**, 399–400 (2020).
79. Madani, A., Ong, J. R., Tibrewal, A. & Mofrad, M. R. K. Deep echocardiography: data-efficient supervised and semi-supervised deep learning towards automated diagnosis of cardiac disease. *npj Digital Medicine* 2018 1:1 **1**, 59- (2018).
80. Seetharam, K., Raina, S. & Sengupta, P. P. The Role of Artificial Intelligence in Echocardiography. *Curr. Cardiol. Rep.* **22**, 99- (2020).
81. Silva, G. F. S., Fagundes, T. P., Teixeira, B. C. & Chiavegatto Filho, A. D. P. Machine Learning for Hypertension Prediction: a Systematic Review. *Curr. Hypertens. Rep.* **24**, 523–533 (2022).
82. Ahmad, T. *et al.* Machine learning methods improve prognostication, identify clinically distinct phenotypes, and detect heterogeneity in response to therapy in a large cohort of heart failure patients. *J. Am. Heart Assoc.* **7**, (2018).
83. Ghorbani, A. *et al.* Deep learning interpretation of echocardiograms. *npj Digital Medicine* 2020 3:1 **3**, 10- (2020).
84. Chen, C. *et al.* Deep Learning for Cardiac Image Segmentation: A Review. *Front. Cardiovasc. Med.* **7**, 508599 (2020).
85. Somani, S. *et al.* Deep learning and the electrocardiogram: review of the current state-of-the-art. *EP Europace* **23**, 1179–1191 (2021).
86. Zeleznik, R. *et al.* Deep convolutional neural networks to predict cardiovascular risk from computed tomography. *Nature Communications* 2021 12:1 **12**, 715- (2021).
87. Argha, A., Celler, B. G. & Lovell, N. H. Artificial Intelligence Based Blood Pressure Estimation from Auscultatory and Oscillometric Waveforms: A Methodological Review. *IEEE Rev. Biomed. Eng.* **15**, 152–168 (2022).
88. Argha, A., Celler, B. G. & Lovell, N. H. A novel automated blood pressure estimation algorithm using sequences of korotkoff sounds. *IEEE J. Biomed. Health Inform.* **25**, 1257–1264 (2021).
89. Esmaelpoor, J., Moradi, M. H. & Kadkhodamohammadi, A. A multistage deep neural network model for blood pressure estimation using photoplethysmogram signals. *Comput. Biol. Med.* **120**, 103719 (2020).
90. Elgendi, M. *et al.* The use of photoplethysmography for assessing hypertension. *npj Digital Medicine* 2019 2:1 **2**, 60- (2019).
91. Squirrel, D. M. *et al.* Blood Pressure Predicted From Artificial Intelligence Analysis of Retinal Images Correlates With Future Cardiovascular Events. *JACC: Advances* **3**, 101410 (2024).

92. Jana, B., Oswal, K., Mitra, S., Saha, G. & Banerjee, S. Windkessel Model-Based Cuffless Blood Pressure Estimation Using Continuous Wave Doppler Ultrasound System. *IEEE Sens. J.* **20**, 9989–9999 (2020).
93. Layton, A. T. AI, Machine Learning, and ChatGPT in Hypertension. *Hypertension* **81**, 709–716 (2024).
94. Golino, H. F. *et al.* Predicting Increased Blood Pressure Using Machine Learning. *J. Obes.* **2014**, 637635 (2014).
95. Kanegae, H. *et al.* Highly precise risk prediction model for new-onset hypertension using artificial intelligence techniques. *J. Clin. Hypertens.* **22**, 445–450 (2020).
96. Huang, S. *et al.* Evaluating the risk of hypertension using an artificial neural network method in rural residents over the age of 35 years in a Chinese area. *Hypertension Research* 2010 33:7 **33**, 722–726 (2010).
97. Niu, M. *et al.* Identifying the predictive effectiveness of a genetic risk score for incident hypertension using machine learning methods among populations in rural China. *Hypertension Research* 2021 44:11 **44**, 1483–1491 (2021).
98. Pei, Z. *et al.* Risk-Predicting Model for Incident of Essential Hypertension Based on Environmental and Genetic Factors with Support Vector Machine. *Interdisciplinary Sciences: Computational Life Sciences* 2018 10:1 **10**, 126–130 (2018).
99. Wu, X. *et al.* Value of a Machine Learning Approach for Predicting Clinical Outcomes in Young Patients With Hypertension. *Hypertension* **75**, 1271–1278 (2020).
100. Kakadiaris, I. A. *et al.* Machine learning outperforms ACC/AHA CVD risk calculator in MESA. *J. Am. Heart Assoc.* **7**, (2018).
101. Lee, S.-J. ; *et al.* Deep Learning Improves Prediction of Cardiovascular Disease-Related Mortality and Admission in Patients with Hypertension: Analysis of the Korean National Health Information Database. *Journal of Clinical Medicine* 2022, Vol. 11, Page 6677 **11**, 6677 (2022).
102. Alaa, A. M., Bolton, T., Angelantonio, E. Di, Rudd, J. H. F. & van der Schaar, M. Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. *PLoS One* **14**, e0213653 (2019).
103. Cho, J. S. & Park, J. H. Application of artificial intelligence in hypertension. *Clinical Hypertension* 2024 30:1 **30**, 11- (2024).
104. Duan, T., Rajpurkar, P., Laird, D., Ng, A. Y. & Basu, S. Clinical Value of Predicting Individual Treatment Effects for Intensive Blood Pressure Therapy: A Machine Learning Experiment to Estimate Treatment Effects from Randomized Trial Data. *Circ. Cardiovasc. Qual. Outcomes* **12**, (2019).
105. Koren, G., Nordon, G., Radinsky, K. & Shalev, V. Machine learning of big data in gaining insight into successful treatment of hypertension. *Pharmacol. Res. Perspect.* **6**, e00396 (2018).
106. Shimbo, D. *et al.* Transforming Hypertension Diagnosis and Management in The Era of Artificial Intelligence: A 2023 National Heart, Lung, and Blood Institute (NHLBI) Workshop Report. *Hypertension* **82**, 36–45 (2025).

107. Iturria-Medina, Y., Khan, A. F., Adewale, Q. & Shirazi, A. H. Blood and brain gene expression trajectories mirror neuropathology and clinical deterioration in neurodegeneration. *Brain* **143**, 661–673 (2020).
108. McCarthy, J. *et al.* Data-driven staging of genetic frontotemporal dementia using multi-modal MRI. *Hum. Brain Mapp.* **43**, 1821–1835 (2022).
109. Greenacre, M. *et al.* Principal component analysis. *Nature Reviews Methods Primers* **2022 2:1** **2**, 100- (2022).
110. Blomsma, N. *et al.* Minimum spanning tree analysis of brain networks: A systematic review of network size effects, sensitivity for neuropsychiatric pathology, and disorder specificity. *Network Neuroscience* **6**, 301–319 (2022).
111. Wang, X. *et al.* Hypertension-mediated organ damage and established cardiovascular disease in patients with hypertension: the China Hypertension Survey, 2012–2015. *Journal of Human Hypertension* **2021 36:12** **36**, 1092–1098 (2021).
112. Carey, R. M., Wright, J. T., Taler, S. J. & Whelton, P. K. Guideline-Driven Management of Hypertension: An Evidence-Based Update. *Circ. Res.* **128**, 827 (2021).
113. Carey, R. M., Muntner, P., Bosworth, H. B. & Whelton, P. K. Prevention and Control of Hypertension: JACC Health Promotion Series. *J. Am. Coll. Cardiol.* **72**, 1278–1293 (2018).
114. Schutte, A. E., Srinivasapura Venkateshmurthy, N., Mohan, S. & Prabhakaran, D. Hypertension in Low- and Middle-Income Countries. *Circ. Res.* **128**, 808–826 (2021).
115. Whitmore, K. *et al.* Impact of Patient Visualization of Cardiovascular Images on Modification of Cardiovascular Risk Factors: Systematic Review and Meta-Analysis. *Cardiovascular Imaging* **16**, 1069–1081 (2023).
116. Unger, T. *et al.* 2020 International Society of Hypertension Global Hypertension Practice Guidelines. *Hypertension* **75**, 1334–1357 (2020).
117. Dunn, P., Ali, A., Patel, A. P. & Banerjee, S. Brief Review and Primer of Key Terminology for Artificial Intelligence and Machine Learning in Hypertension. *Hypertension* <https://doi.org/10.1161/HYPERTENSIONAHA.123.22347> (2025) doi:10.1161/HYPERTENSIONAHA.123.22347.
118. Chaikijurajai, T., Laffin, L. J. & Wilson Tang, W. H. Artificial Intelligence and Hypertension: Recent Advances and Future Outlook. *Am. J. Hypertens.* **33**, 967–974 (2020).
119. Mello-Thoms, C. & Mello, C. A. B. Clinical applications of artificial intelligence in radiology. *Br. J. Radiol.* **96**, 20221031 (2023).
120. Khera, R. *et al.* Transforming Cardiovascular Care With Artificial Intelligence: From Discovery to Practice: JACC State-of-the-Art Review. *J. Am. Coll. Cardiol.* **84**, 97–114 (2024).
121. Page, M. J. *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **372**, (2021).
122. Clark, J. *et al.* A full systematic review was completed in 2 weeks using automation tools: a case study. *J. Clin. Epidemiol.* **121**, 81–90 (2020).

123. Ouzzani, M., Hammady, H., Fedorowicz, Z. & Elmagarmid, A. Rayyan-a web and mobile app for systematic reviews. *Syst. Rev.* **5**, 1–10 (2016).
124. Drucker, A. M., Fleming, P. & Chan, A. W. Research Techniques Made Simple: Assessing Risk of Bias in Systematic Reviews. *Journal of Investigative Dermatology* **136**, e109–e114 (2016).
125. Sounderajah, V. *et al.* A quality assessment tool for artificial intelligence-centered diagnostic test accuracy studies: QUADAS-AI. *Nature Medicine* **27**, 1663–1665 (2021).
126. Wolff, R. F. *et al.* PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Ann. Intern. Med.* **170**, 51–58 (2019).
127. Higgins, J. P. T., Thompson, S. G., Deeks, J. J. & Altman, D. G. Measuring inconsistency in meta-analyses. *BMJ* **327**, 557–560 (2003).
128. Lian, Q., Hodges, J. S. & Chu, H. A Bayesian Hierarchical Summary Receiver Operating Characteristic Model for Network Meta-Analysis of Diagnostic Tests. *J. Am. Stat. Assoc.* **114**, 949–961 (2019).
129. Alsharqi, M. *et al.* A machine learning-based score for precise echocardiographic assessment of cardiac remodelling in hypertensive young adults. *European Heart Journal - Imaging Methods and Practice* **1**, (2023).
130. Recenti, M. *et al.* Healthy Aging within an Image: Using Muscle Radiodensitometry and Lifestyle Factors to Predict Diabetes and Hypertension. *IEEE J. Biomed. Health Inform.* **25**, 2103–2112 (2021).
131. Pessana, F., Venialgo, E., Rubstein, J. & Furfaro, A. Assessment of human instantaneous arterial diameter using Bmode ultrasound imaging and Artificial Neural Networks: Determination of wall mechanical properties. *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'10* 1409–1412 (2010) doi:10.1109/IEMBS.2010.5626719.
132. Karki, S., Ali, A. E., Alsadoon, O. H. & Rashid, T. A. A novel solution of an enhanced error and loss function using deep learning for hypertension classification in traditional medicine. *CITISIA 2020 - IEEE Conference on Innovative Technologies in Intelligent Systems and Industrial Applications, Proceedings* <https://doi.org/10.1109/CITISIA50690.2020.9371809> (2020) doi:10.1109/CITISIA50690.2020.9371809.
133. Alkhodari, M. *et al.* HyperScore: A unified measure to model hypertension progression using multi-modality measurements and semi-supervised learning. *Proceedings - 2023 2023 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2023* 1886–1889 (2023) doi:10.1109/BIBM58861.2023.10385558.
134. Cetin, I. *et al.* Radiomics Signatures of Cardiovascular Risk Factors in Cardiac MRI: Results From the UK Biobank. *Front. Cardiovasc. Med.* **7**, 591368 (2020).
135. Sabovčik, F. *et al.* Applying machine learning to detect early stages of cardiac remodelling and dysfunction. *Eur. Heart J. Cardiovasc. Imaging* **22**, 1208–1217 (2021).
136. Barbieri, A. *et al.* Three-Dimensional Automated, Machine-Learning-Based Left Heart Chamber Metrics: Associations with Prevalent Vascular Risk Factors and Cardiovascular Diseases. *J. Clin. Med.* **11**, 7363 (2022).

137. Dai, G. *et al.* Exploring the effect of hypertension on retinal microvasculature using deep learning on East Asian population. *PLoS One* **15**, e0230111 (2020).
138. Katz, D. H. *et al.* Phenomapping for the Identification of Hypertensive Patients with the Myocardial Substrate for Heart Failure with Preserved Ejection Fraction. *J. Cardiovasc. Transl. Res.* **10**, 275–284 (2017).
139. Raghavendra, U. *et al.* Automated Diagnosis and Assessment of Cardiac Structural Alteration in Hypertension Ultrasound Images. *Contrast Media Mol. Imaging* **2022**, (2022).
140. Diao, K. *et al.* Multi-channel deep learning model-based myocardial spatial–temporal morphology feature on cardiac MRI cine images diagnoses the cause of LVH. *Insights Imaging* **14**, 1–11 (2023).
141. Wang, Y. R. *et al.* Screening and diagnosis of cardiovascular disease using artificial intelligence-enabled cardiac magnetic resonance imaging. *Nature Medicine* **2024** 30:5 30, 1471–1480 (2024).
142. Wang, Y. *et al.* Artificial intelligence measuring the aortic diameter assist in identifying adverse blood pressure status including masked hypertension. *Postgrad. Med.* **134**, 111–121 (2022).
143. Xu, Z., Yu, F., Zhang, B. & Zhang, Q. Intelligent diagnosis of left ventricular hypertrophy using transthoracic echocardiography videos. *Comput. Methods Programs Biomed.* **226**, 107182 (2022).
144. Yang, Z. *et al.* Gene-SGAN: discovering disease subtypes with imaging and genetic signatures via multi-view weakly-supervised deep clustering. *Nature Communications* **2024** 15:1 15, 1–16 (2024).
145. Morales, S., Naranjo, V., Navea, A. & Alcañiz, M. Computer-aided diagnosis software for hypertensive risk determination through fundus image processing. *IEEE J. Biomed. Health Inform.* **18**, 1757–1763 (2014).
146. Loncaric, F. *et al.* Automated Pattern Recognition in Whole-Cardiac Cycle Echocardiographic Data: Capturing Functional Phenotypes with Machine Learning. *Journal of the American Society of Echocardiography* **34**, 1170–1183 (2021).
147. Lo, Y. *et al.* AML-Net: A Preliminary Screening Model for Mild Hypertension. *Proceedings - 2021 14th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, CISP-BMEI 2021* <https://doi.org/10.1109/CISP-BMEI53629.2021.9624382> (2021) doi:10.1109/CISP-BMEI53629.2021.9624382.
148. Yu, X. *et al.* Analyzing brain structural differences associated with categories of blood pressure in adults using empirical kernel mapping-based kernel ELM. *Biomed. Eng. Online* **18**, (2019).
149. Kandil, H. *et al.* A CAD System for the Early Prediction of Hypertension based on Changes in Cerebral Vasculature. *IST 2019 - IEEE International Conference on Imaging Systems and Techniques, Proceedings* <https://doi.org/10.1109/IST48021.2019.9010179> (2019) doi:10.1109/IST48021.2019.9010179.
150. Kandil, H. *et al.* A novel computer-aided diagnosis system for the early detection of hypertension based on cerebrovascular alterations. *Neuroimage Clin.* **25**, 102107 (2020).

151. Kandil, H. *et al.* Analysis of the Importance of Systolic Blood Pressure Versus Diastolic Blood Pressure in Diagnosing Hypertension: MRA Study. *Proceedings - International Conference on Image Processing, ICIP 2020-October*, 443–447 (2020).
152. Peng, B., Yu, X., Ma, X., Zhu, J. & Dai, Y. Self-paced learning based multi-kernel KRR for brain structure analysis in patients with different blood pressure levels. *Proceedings - 2020 13th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, CISP-BMEI 2020* 169–174 (2020) doi:10.1109/CISP-BMEI51763.2020.9263541.
153. Cetin, I. *et al.* A radiomics approach to analyze cardiac alterations in hypertension. *Proceedings - International Symposium on Biomedical Imaging 2019-April*, 640–643 (2019).
154. Rose, B., Kavya, S., Rachana, S. & Manisha, E. Cardiovascular Disease Prediction from Retinal Images using Machine Learning. *2nd International Conference on Sustainable Computing and Data Communication Systems, ICSCDS 2023 - Proceedings* <https://doi.org/10.1109/ICSCDS56580.2023.10104816> (2023) doi:10.1109/ICSCDS56580.2023.10104816.
155. Koduri, H. & Ma, M. Diabetic and Hypertensive Retinopathy Classification from Retinal Images Using Dual Vision Transformer. *Proceedings - 2024 4th International Conference on Digital Data Processing, DDP 2024* 87–91 (2024) doi:10.1109/DDP64453.2024.00034.
156. Azhari, H., Gath, I., Beyar, R., Marcus, M. L. & Sideman, S. Discrimination Between Healthy and Diseased Hearts by Spectral Decomposition of Their Left Ventricular Three-Dimensional Geometry. *IEEE Trans. Med. Imaging* **10**, 207–215 (1991).
157. Neisius, U. *et al.* Radiomic Analysis of Myocardial Native T1 Imaging Discriminates Between Hypertensive Heart Disease and Hypertrophic Cardiomyopathy. *JACC Cardiovasc. Imaging* **12**, 1946–1954 (2019).
158. Vidal-Sospedra, I. *et al.* Determination of Image-based Biomarkers for the Diagnosis of Hypertrophic Cardiomyopathy, Hypertensive Cardiomyopathy and Amyloidosis from Texture Analysis in Cardiac MRI. *Proceedings - IEEE 20th International Conference on Bioinformatics and Bioengineering, BIBE 2020* 230–235 (2020) doi:10.1109/BIBE50027.2020.00045.
159. Yu, F. *et al.* Artificial intelligence-based myocardial texture analysis in etiological differentiation of left ventricular hypertrophy. *Ann. Transl. Med.* **9**, 108–108 (2021).
160. Shi, R. Y. *et al.* Texture analysis applied in T1 maps and extracellular volume obtained using cardiac MRI in the diagnosis of hypertrophic cardiomyopathy and hypertensive heart disease compared with normal controls. *Clin. Radiol.* **76**, 236.e9–236.e19 (2021).
161. Hwang, I. C. *et al.* Differential diagnosis of common etiologies of left ventricular hypertrophy using a hybrid CNN-LSTM model. *Scientific Reports 2022 12:1* **12**, 1–12 (2022).
162. Zhang, X. *et al.* Deep learn-based computer-assisted transthoracic echocardiography: approach to the diagnosis of cardiac amyloidosis. *International Journal of Cardiovascular Imaging* **39**, 955–965 (2023).

163. Wang, Z. C. *et al.* Deep Learning for Discrimination of Hypertrophic Cardiomyopathy and Hypertensive Heart Disease on MRI Native T1 Maps. *Journal of Magnetic Resonance Imaging* **59**, 837–848 (2024).
164. Kaupp, A. *et al.* Measuring morphologic properties of the human retinal vessel system using a two-stage image processing approach. *Proceedings - International Conference on Image Processing, ICIP* **1**, 431–435 (1994).
165. Khitran, S., Akram, M. U., Usman, A. & Yasin, U. Automated system for the detection of hypertensive retinopathy. *2014 4th International Conference on Image Processing Theory, Tools and Applications, IPTA 2014* <https://doi.org/10.1109/IPTA.2014.7001984> (2015) doi:10.1109/IPTA.2014.7001984.
166. Ahmad, F., Khan Sial, M. R., Yousaf, A. & Khan, F. Textural and Intensity Feature Based Retinal Vessels Classification for the Identification of Hypertensive Retinopathy. *Proceedings of the 21st International Multi Topic Conference, INMIC 2018* <https://doi.org/10.1109/INMIC.2018.8595667> (2018) doi:10.1109/INMIC.2018.8595667.
167. Kiruthika, M., Swapna, T. R., Santhosh, K. C. & Peeyush, K. P. Artery and vein classification for hypertensive retinopathy. *Proceedings of the International Conference on Trends in Electronics and Informatics, ICOEI 2019* 244–248 (2019) doi:10.1109/ICOEI.2019.8862719.
168. Bhimavarapu, U., Chintalapudi, N. & Battineni, G. Automatic Detection and Classification of Hypertensive Retinopathy with Improved Convolution Neural Network and Improved SVM. *Bioengineering 2024, Vol. 11, Page 56* **11**, 56 (2024).
169. Triwijoyo, B. K., Adil, A. & Zulfikri, M. Detection and classification of hypertensive retinopathy based on retinal image analysis using a deep learning approach. *Computer Methods and Programs in Biomedicine Update* **7**, 100191 (2025).
170. Leoncini, G. *et al.* Microalbuminuria identifies overall cardiovascular risk in essential hypertension: an artificial neural network-based approach. *J. Hypertens.* **20**, 1315–1321 (2002).
171. Viazzi, F. *et al.* Predicting cardiovascular risk using creatinine clearance and an artificial neural network in primary hypertension. *J. Hypertens.* **24**, 1281–1286 (2006).
172. Suvila, K. *et al.* Early Onset Hypertension Is Associated with Hypertensive End-Organ Damage Already by MidLife. *Hypertension* **74**, 305–312 (2019).
173. Berek, P. AL, Irawati, D. & Hamid, A. Y. S. Hypertension: A global health crisis. *Annals of Clinical Hypertension* **5**, 008–011 (2021).
174. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018 562:7726 **562**, 203–209 (2018).
175. Littlejohns, T. J. *et al.* The UK Biobank imaging enhancement of 100,000 participants: rationale, data collection, management and future directions. *Nature Communications* 2020 11:1 **11**, 1–12 (2020).
176. Orhac, F. *et al.* A Guide to ComBat Harmonization of Imaging Biomarkers in Multicenter Studies. *Journal of Nuclear Medicine* **63**, 172 (2022).

177. Liu, F. T., Ting, K. M. & Zhou, Z. H. Isolation forest. *Proceedings - IEEE International Conference on Data Mining, ICDM* 413–422 (2008) doi:10.1109/ICDM.2008.17.
178. Abid, A., Zhang, M. J., Bagaria, V. K. & Zou, J. Exploring patterns enriched in a dataset with contrastive principal component analysis. *Nature Communications* 2018 9:1 **9**, 1–7 (2018).
179. Belkin, M. & Niyogi, P. Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering.
180. Martínez-Camblor, P. & Pardo-Fernández, J. C. The Youden Index in the Generalized Receiver Operating Characteristic Curve Context. *International Journal of Biostatistics* **15**, (2019).
181. Cunningham, P. & Delany, S. J. K-Nearest Neighbour Classifiers-A Tutorial. *ACM Comput. Surv.* **54**, (2022).
182. Angelopoulos, A. N. & Bates, S. A Gentle Introduction to Conformal Prediction and Distribution-Free Uncertainty Quantification. <https://arxiv.org/abs/2107.07511v6> (2021).
183. Menéndez, M. L., Pardo, J. A., Pardo, L. & Pardo, M. C. The Jensen-Shannon divergence. *J. Franklin Inst.* **334**, 307–318 (1997).
184. Clim, A., Zota, R. D. & Tinica, G. The Kullback-Leibler Divergence Used in Machine Learning Algorithms for Health Care Applications and Hypertension Prediction: A Literature Review. *Procedia Comput. Sci.* **141**, 448–453 (2018).
185. D'Arrigo, G. *et al.* Methods to Analyse Time-to-Event Data: The Kaplan-Meier Survival Curve. *Oxid. Med. Cell. Longev.* **2021**, 2290120 (2021).
186. McClelland, R. L. *et al.* 10-Year Coronary Heart Disease Risk Prediction Using Coronary Artery Calcium and Traditional Risk Factors Derivation in the MESA (Multi-Ethnic Study of Atherosclerosis) with Validation in the HNR (Heinz Nixdorf Recall) Study and the DHS (Dallas Heart Stu.... *J. Am. Coll. Cardiol.* **66**, 1643–1653 (2015).
187. Elghazaly, H. *et al.* Characterizing the hypertensive cardiovascular phenotype in the UK Biobank. *Eur. Heart J. Cardiovasc. Imaging* **00**, 1–9 (2023).
188. Siedlinski, M. *et al.* Genetic analyses identify brain structures related to cognitive impairment associated with elevated blood pressure. *Eur. Heart J.* **44**, 2114–2125 (2023).
189. Murray, E. C. *et al.* Vascular phenotypes in early hypertension. *Journal of Human Hypertension* 2022 37:10 **37**, 898–906 (2022).
190. Kart, T., Alkhodari, M., Lapidaire, W. & Leeson, P. Modelling relations between blood pressure, cardiovascular phenotype, and clinical factors using large scale imaging data. *Eur. Heart J. Cardiovasc. Imaging* <https://doi.org/10.1093/EHJCI/JEAD161> (2023) doi:10.1093/EHJCI/JEAD161.
191. Carnevale, L. *et al.* Brain MRI fiber-tracking reveals white matter alterations in hypertensive patients without damage at conventional neuroimaging. *Cardiovasc. Res.* **114**, 1536–1546 (2018).
192. Gillis, E. E. & Sullivan, J. C. Sex Differences in Hypertension: Recent Advances. *Hypertension* **68**, 1322–1327 (2016).

193. Ma, H. *et al.* Clinical value of multiorgan damage in hypertensive crises: A prospective follow-up study. *The Journal of Clinical Hypertension* **22**, 914 (2020).
194. Ren, L. *et al.* Correlation between hypertension and common carotid artery intima-media thickness in rural China: a population-based study. *Journal of Human Hypertension* **2018** 32:8 **32**, 548–554 (2018).
195. Chapman, N. *et al.* Absolute cardiovascular risk assessment using ‘real world’ clinic blood pressures compared to standardized unobserved and ambulatory methods: an observational study. *Hypertension Research* **2024** 47:10 **47**, 2855–2863 (2024).
196. Dehghan, A., Ahmadnia Motlagh, S., Khezri, R., Rezaei, F. & Aune, D. A comparison of laboratory-based and office-based Framingham risk scores to predict 10-year risk of cardiovascular diseases: a population-based study. *J. Transl. Med.* **21**, 687 (2023).
197. van Daalen, K. R. *et al.* Risk estimation for the primary prevention of cardiovascular disease: considerations for appropriate risk prediction model selection. *Lancet Glob. Health* **12**, e1343–e1358 (2024).
198. Yang, S., Varghese, P., Stephenson, E., Tu, K. & Gronsbell, J. Machine learning approaches for electronic health records phenotyping: a methodical review. *Journal of the American Medical Informatics Association* **30**, 367–381 (2023).
199. Bisong, E. *et al.* Predicting high blood pressure using machine learning models in low- and middle-income countries. *BMC Med. Inform. Decis. Mak.* **24**, 234 (2024).
200. Cho, S. M. J. *et al.* Measured Blood Pressure, Genetically Predicted Blood Pressure, and Cardiovascular Disease Risk in the UK Biobank. *JAMA Cardiol.* **7**, 1129–1137 (2022).
201. Wan, E. Y. F. *et al.* Blood Pressure and Risk of Cardiovascular Disease in UK Biobank: A Mendelian Randomization Study. *Hypertension* **77**, 367–375 (2021).
202. Fry, A. *et al.* Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *Am. J. Epidemiol.* **186**, 1026–1034 (2017).
203. Wang, X. *et al.* Transformer-based unsupervised contrastive learning for histopathological image classification. *Med. Image Anal.* **81**, 102559 (2022).
204. Oliveira, de E. *et al.* Identifying patterns differing between high-dimensional datasets with generalized contrastive PCA. *PLoS Comput. Biol.* **21**, e1012747 (2025).
205. Kart, T. *et al.* Deep Learning-based Modelling of Complex Hypertensive Multi-Organ Damage with Uncertainty Quantification from Simple Clinical Measures. *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* 1542–1547 (2024) doi:10.1109/BIBM62325.2024.10822397.
206. Zou, J. & Schiebinger, L. Ensuring that biomedical AI benefits diverse populations. *EBioMedicine* **67**, 103358 (2021).
207. Noseworthy, P. A. *et al.* Assessing and Mitigating Bias in Medical Artificial Intelligence: The Effects of Race and Ethnicity on a Deep Learning Model for ECG Analysis. *Circ. Arrhythm. Electrophysiol.* **13**, E007988 (2020).
208. Tran, B. X. *et al.* Global Evolution of Research in Artificial Intelligence in Health and Medicine: A Bibliometric Study. *Journal of Clinical Medicine* **2019**, Vol. 8, Page 360 **8**, 360 (2019).

209. Ramspek, C. L., Jager, K. J., Dekker, F. W., Zoccali, C. & Van Diepen, M. External validation of prognostic models: what, why, how, when and where? *Clin. Kidney J.* **14**, 49–58 (2021).
210. Cabitza, F. *et al.* The importance of being external. methodological insights for the external validation of machine learning models in medicine. *Comput. Methods Programs Biomed.* **208**, 106288 (2021).
211. Wright, J. D. *et al.* The ARIC (Atherosclerosis Risk In Communities) Study: JACC Focus Seminar 3/8. *J. Am. Coll. Cardiol.* **77**, 2939–2959 (2021).
212. Taliun, D. *et al.* Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *Nature* **2021 590:7845 590**, 290–299 (2021).
213. Kalbaugh, C. A. *et al.* Frequency of care and mortality following an incident diagnosis of peripheral artery disease in the inpatient or outpatient setting: The aric (atherosclerosis risk in communities) study. *J. Am. Heart Assoc.* **7**, (2018).
214. Soma-Pillay, P., Nelson-Piercy, C., Tolppanen, H. & Mebazaa, A. Physiological changes in pregnancy. *Cardiovasc. J. Afr.* **27**, 89–94 (2016).
215. Khedagi, A. M. & Bello, N. A. Hypertensive Disorders of Pregnancy. *Cardiol. Clin.* **39**, 77–90 (2021).
216. Ashworth, D. C. *et al.* Postnatal health and care following hypertensive disorders in pregnancy: a prospective cohort study (BPiPP study). *BMC Pregnancy and Childbirth* **2022 22:1 22**, 1–12 (2022).
217. Poon, L. C. *et al.* Hypertensive disorders of pregnancy and long-term cardiovascular health: FIGO Best Practice Advice. *International Journal of Gynecology and Obstetrics* **160**, 22–34 (2023).
218. Wu, P. *et al.* Preeclampsia and future cardiovascular health. *Circ. Cardiovasc. Qual. Outcomes* **10**, (2017).
219. Haug, E. B. *et al.* Association of Conventional Cardiovascular Risk Factors With Cardiovascular Disease After Hypertensive Disorders of Pregnancy: Analysis of the Nord-Trøndelag Health Study. *JAMA Cardiol.* **4**, 628–635 (2019).
220. Cutler, H. R. *et al.* Temporal patterns of pre- and post-natal target organ damage associated with hypertensive pregnancy: a systematic review. *Eur. J. Prev. Cardiol.* **31**, 77–99 (2024).
221. Khosla, K. *et al.* Long-Term Cardiovascular Disease Risk in Women After Hypertensive Disorders of Pregnancy: Recent Advances in Hypertension. *Hypertension* **78**, 927–935 (2021).
222. Boardman, H. *et al.* Variations in Cardiovascular Structure, Function, and Geometry in Midlife Associated with a History of Hypertensive Pregnancy. *Hypertension* **75**, 1542–1550 (2020).
223. Papageorgiou, A. T. *et al.* The INTERGROWTH-21 st fetal growth standards: toward the global integration of pregnancy and pediatric care. *Am. J. Obstet. Gynecol.* **218**, S630–S640 (2018).

224. Da Costa, B. R. *et al.* Methods to convert continuous outcomes into odds ratios of treatment response and numbers needed to treat: meta-epidemiological study. *Int. J. Epidemiol.* **41**, 1445–1459 (2012).
225. Gnona, K. M. & Stewart, W. C. L. Revisiting the Wald Test in Small Case-Control Studies With a Skewed Covariate. *Am. J. Epidemiol.* **191**, 1508–1518 (2022).
226. Agresti, A. & Caffo, B. Simple and effective confidence intervals for proportions and differences of proportions result from adding two successes and two failures. *American Statistician* **54**, 280–288 (2000).
227. Altman, D. G. & Bland, J. M. Interaction revisited: the difference between two estimates. *BMJ* **326**, 219 (2003).
228. Grandi, S. M. *et al.* Long-Term Mortality in Women With Pregnancy Loss and Modification by Race/Ethnicity. *Am. J. Epidemiol.* **191**, 787–799 (2022).
229. Vlachou, F. *et al.* Fetal loss and long-term maternal morbidity and mortality: A systematic review and meta-analysis. *PLoS Med.* **21**, e1004342 (2024).
230. Nie, X., Xu, Z. & Ren, H. Analysis of risk factors of preeclampsia in pregnant women with chronic hypertension and its impact on pregnancy outcomes. *BMC Pregnancy and Childbirth* 2024 24:1 **24**, 1–8 (2024).
231. Boyd, H. A., Tahir, H., Wohlfahrt, J. & Melbye, M. Associations of Personal and Family Preeclampsia History With the Risk of Early-, Intermediate- and Late-Onset Preeclampsia. *Am. J. Epidemiol.* **178**, 1611–1619 (2013).
232. Skjærven, R. *et al.* Recurrence of pre-eclampsia across generations: exploring fetal and maternal genetic components in a population based cohort. *BMJ* **331**, 877 (2005).
233. Tanz, L. J. *et al.* Preterm Delivery and Maternal Cardiovascular Disease in Young and Middle-Aged Adult Women. *Circulation* **135**, 578–589 (2017).
234. Wu, P. *et al.* Preterm delivery and future risk of maternal cardiovascular disease: A systematic review and meta-analysis. *J. Am. Heart Assoc.* **7**, (2018).
235. Lou, Y., He, P., Jiang, H., Xiang, L. & Gao, X. Analysis of the characteristics of blood lipid metabolism in twin pregnancy. *Journal of Investigative Medicine* **71**, 53–57 (2023).
236. Wei, J. *et al.* Exploration of Serum lipid levels during twin pregnancy. *Journal of Maternal-Fetal and Neonatal Medicine* **36**, (2023).
237. Lembo, M. *et al.* Prolonged Time-to-antihypertensive Therapy Worsens Organ Damage and Blood Pressure Control in Arterial Hypertension. *High Blood Pressure & Cardiovascular Prevention* 2024 31:6 **31**, 639–648 (2024).
238. Hernandez-Rubio, A. *et al.* Role of Ethnicity and Sex in Hypertension-Mediated Organ Damage in a Dual-Ethnic Cohort of Individuals With Hypertension. *Hypertension* <https://doi.org/10.1161/HYPERTENSIONAHA.125.25032> (2025)
doi:10.1161/HYPERTENSIONAHA.125.25032.
239. Cutler, H. R. *et al.* Temporal patterns of pre- and post-natal target organ damage associated with hypertensive pregnancy: a systematic review. *Eur. J. Prev. Cardiol.* **31**, 77–99 (2024).