

Trans-ethnic Associations of Plasma Lipid Traits and Blood Pressure and Risk of Myocardial Infarction

Hanyu Wang, MPhil

Kellogg College, University of Oxford

A thesis submitted for the degree of Doctor of Philosophy

Thesis word count: 44,558 words

Michaelmas Term 2025

Abstract

Background: Myocardial infarction (MI), the most serious form of ischaemic heart disease (IHD), is a leading cause of premature death and disability worldwide, and while stroke mortality has declined in recent decades, MI mortality continues to increase in China. Elevated levels of low-density lipoprotein cholesterol (LDL-C) and systolic blood pressure (SBP) are among the most important modifiable risk factors for MI. However, uncertainties remain about the strength and shape of their associations with risk of MI across diverse ancestry populations with different baseline plasma lipid levels, and about the combined effects on risk of MI of lowering both risk factors simultaneously.

Methods: This thesis used observational and genetic approaches to investigate the individual and combined associations of plasma lipid traits and SBP with risk of MI in Chinese and European (White British) populations. First, an umbrella review summarised evidence from prospective cohort studies, Mendelian randomisation (MR) analyses, and randomised controlled trials (RCTs). Second, observational analyses were conducted in a nested case-control study with ~1,000 MI cases within the China Kadoorie Biobank (CKB, ~0.5 million adults recruited from 10 regions in China) and in the UK Biobank (UKB, ~0.5 million adults from the UK with >5,000 MI cases). Logistic and Cox regression models were applied to estimate rate ratios (RRs) in both ancestry populations. Third, MR analyses were undertaken in CKB and UKB to assess the causal relevance of plasma lipid traits for risk of MI, with additional factorial MR to test the independent and joint effects of LDL-C and SBP on risk of MI.

Results: In observational analyses, mean levels of LDL-C were ~30% lower in CKB than in UKB (2.3 [SD 0.6] vs 3.7 [SD 0.8] mmol/L), whereas mean levels of HDL-C and triglycerides were comparable. The RRs of MI per 1-SD higher usual levels of LDL-C were about half as strong in Chinese as in Europeans (1.27; 95% CI 1.13–1.44 vs 1.55; 1.49–1.61), but the corresponding RRs for ApoB and non-HDL-C were similar across populations. In MR analyses using restricted genetic risk scores (GRS), which excluded variants highly correlated with lipid traits other than LDL-C, higher levels of genetically-predicted LDL-C were equally strongly associated with risk of MI in both populations (RR per 1-SD higher levels of genetically-predicted LDL-C: 1.23, 95% CI 1.01–1.50 in Chinese; 1.37, 1.04–1.79 in Europeans). Higher levels of genetically-predicted triglycerides and HDL-C were not significantly associated with risk of MI when using restricted GRS. Levels of SBP were more strongly associated with risk of MI in Chinese than in Europeans. Each 1-SD higher usual level of SBP was associated with a 70% higher risk of MI in Chinese (RR 1.70, 1.64–1.76) and a 45% higher risk in Europeans (RR 1.45, 1.38–1.51). The positive associations of higher levels of LDL-C with risk of MI were consistent across SBP strata, with no evidence of interaction (P for heterogeneity = 0.573 in Chinese and 0.698 in Europeans). Factorial MR confirmed that the associations of genetically-predicted lower LDL-C and lower SBP with risk of MI were independent (P for interaction = 0.14 in Chinese and 0.62 in Europeans).

Conclusions: This thesis provides observational and genetic evidence from both Chinese and European populations that higher levels of LDL-C and SBP were each strongly and causally associated with risk of MI, and that their combined effects were independent of each other. The results demonstrated that, despite lower levels of LDL-related traits in the Chinese population compared with Europeans, the

proportional effects of lowering LDL-C on risk of MI were comparable across populations. The findings add to the accumulating evidence that primary prevention of atherosclerotic cardiovascular disease (ASCVD) should prioritise initiation of lipid-lowering drug treatments in individuals at high overall cardiovascular risk rather than basing it on elevated LDL-C levels. Moreover, this thesis provides strong evidence for combined lowering of LDL-C and SBP for primary prevention of MI, complementing recent RCT evidence on combination drug treatments. Together, these results could inform global strategies for lipid and blood pressure control to reduce the burden of MI across diverse populations.

Statement of contribution to work

This thesis is my own work unless stated otherwise. The idea to investigate the individual and combined associations of blood pressure and plasma lipid traits with risk of myocardial infarction was conceived by my supervisors. I established the research objectives for this thesis with their input. I developed the data analysis plans, conducted all statistical analyses, produced the tables and figures, and interpreted the findings with feedback from my supervisors, Professor Derrick Bennett, Professor Robert Clarke, and Dr Neil Wright. I wrote all sections of this thesis.

I was not involved in the design or data collection of the UK Biobank or China Kadoorie Biobank.

For the umbrella review presented in Chapter 2, I developed the search strategy with feedback from Nia Roberts, librarian at Bodleian Health Care Libraries, who provided comments on the search terms.

Chapter 4 (observational analysis of plasma lipid traits and risk of myocardial infarction) has been published in European Heart Journal Open. As the first author, I wrote the initial draft of the manuscript and revised it based on comments from all co-authors.

I received funding from the Nuffield Department of Population Health, University of Oxford, to undertake this DPhil project.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisors, Professors Derrick Bennett and Robert Clarke, and Dr Neil Wright. The most valuable lessons I have learned from them extend beyond technical knowledge. They have taught me the importance of dedication to clinically relevant research and meticulous attention to detail, qualities that I will carry with me throughout my career. I would especially like to thank Professor Clarke for his continuous support after his retirement. His dedication to research and tenacious attitude in addressing scientific questions have been truly inspirational.

It has been a great privilege to work with data from the China Kadoorie Biobank and UK Biobank, two of the most valuable resources for epidemiological research. I am deeply grateful to all the participants of these biobanks, and to the many individuals, principal investigators, database managers, and fieldworkers whose efforts made these studies possible.

I would also like to thank the Nuffield Department of Population Health, University of Oxford, for generously funding my DPhil studies. It is, without doubt, the best department to work in at Oxford, not least because of the free lunch scheme that is the envy of all other departments.

This work would not have been possible without the support of many friends I met during my studies. One thing I looked forward to throughout my DPhil was lunch with the BDI lunch group. I am grateful to Jia, Qingze, Kuangyi, Zhaojing, Steven, Michael, Annie, Bowen, Yiming, and many others for the countless hours we spent talking about everything except science during our long (mean >2 hrs?) office lunches. It is true that it is procrastination from work, so if anything in this thesis falls

below the standard, I have them to blame. I would like to give special thanks to my CKB buddies, Charles, Sam, and Valirie, with whom I can always reach out to about anything (mostly venting) in the research group.

I would also like to thank Kellogg College for making my time in Oxford so enjoyable, and for the many friends I met there. Special thanks to Alena, JC, and Abraham, whose home-cooked meals nourished me in this culinary desert; and to Linyi and James, who made my first year in Oxford especially memorable. Most importantly, I will remember my drinking buddies, especially Henderson, Charles, Keyi, Keyu, and many others, and perhaps I should also thank the many bottles of wines, beers, whiskies, and cocktails we shared, and the late nights that brought much-needed joy during the long Oxford winters.

It has been a magical experience living in Oxford, a city I have truly fallen in love with. I will always remember the punting trips along the River Cherwell to the Victoria Arms, the walks around the Radcliffe Camera, which never failed to evoke a sense of mystery, and the long runs from Port Meadow, with its cows and horses (they don’t seem to appear together somehow), to Uni Park and South Park, where I could see astonishing sunsets.

Last but not least, I would like to thank my family for their support and faith in me. I grew up in a small city in northeast China, and my parents and family have always been proud of how far I have come. They have been nothing but supportive throughout all these years of study, and I owe them more than words can express.

Publications

Wang H, Clarke R, Kartsonaki C, Millwood I, Walters R, Hill M, Avery D, Yu C, Sun D-JY, Lv J, Li S, Li L, Chen Z, Wright N, Bennett DA, on behalf of the China Kadoorie Biobank Collaborative Group. Contrasting associations of blood lipids with risk of myocardial infarction in Chinese and European adults. European Heart Journal Open, 2025, oeaf119, https://doi.org/10.1093/ehjopen/oeaf119

Table of Contents

Abstract 2
Statement of contribution to work 5
Acknowledgements 6
Publications 8
Table of Contents 9
List of Abbreviations 13
List of Tables 15
List of Figures 17
Chapter 1: Introduction 19
1.1. Myocardial infarction (MI) 21
1.1.1. Definition 21
1.1.2. Diagnosis and clinical presentation 22
1.1.3. Pathophysiology 23
1.2. Burden and Risk factors for MI 24
1.2.1. Burden of MI globally and in China 24
1.2.2. Major modifiable risk factors for MI 25
1.3. Plasma lipid traits and risk of MI 26
1.3.1 Types of plasma lipids and their metabolism 26
1.3.2. Lipid metabolism and cardiovascular disease 29
1.3.3. Epidemiological evidence of plasma lipid traits and risk of MI 30
1.4. Blood pressure (BP) and risk of MI 32
1.4.1. Physiological link of BP and cardiovascular system 32
1.4.2. Epidemiological evidence of SBP and risk of MI 32
1.5. Combined effects of LDL-C and SBP on risk of MI 33
1.6. The need for further research 35
1.7. Aims and objectives 37
1.8. Structure of the thesis 38
Chapter 2: Umbrella reviews of LDL-C and SBP and risk of IHD 40
2.1. Background 40
2.2. Methods 42
2.2.1. Eligibility criteria 42
2.2.2. Study deduplication and ranking 44
2.2.3. Information sources 46
2.2.4. Data Extraction 48
2.3. Results 50

2.3.1. Overview of the included studies 50
2.3.2. Associations between SBP and IHD 56
2.3.3. Associations between LDL-C and IHD 63
2.3.4. Combined effects of BP and LDL-C 71
2.4. Discussion 74
Chapter 3 Overview of Data Sources and Study Procedures 78
3.1. Background 78
3.2. China Kadoorie Biobank (CKB) design and procedures 79
3.2.1. CKB study design 79
3.2.2. CKB survey data collection 83
3.2.3. Genetic study of CKB 88
3.3. UK Biobank (UKB) design and procedures 89
3.3.1. UKB study design 89
3.3.2. UKB survey data collection 91
3.3.3. Genetic study of UKB 95
3.4. Baseline characteristics of data sources 96
3.4.1. Baseline characteristics of CKB participants by study subsets 96
3.4.2. Baseline characteristics of UKB participants by study subsets 106
3.5. Summary 110
Chapter 4 Observational associations of plasma lipid traits and MI 111
4.1. Background 111
4.2. Methods 113
4.2.1. Study populations and settings 113
4.2.2. Statistical methods 115
4.3. Results 119
4.3.1. Baseline characteristics 119
4.3.2. LDL-C-related lipid traits 126
4.3.3. Triglyceride-related lipid traits 132
4.3.4. HDL-C-related lipid traits 134
4.3.5. Lipid ratios 136
4.4. Discussion 140
Chapter 5 Causal relevance of plasma lipid traits with risk of MI 144
5.1. Background 144
5.2. Methods 150
5.2.1. Study populations and genotyping 150
5.2.2. Instrumental variable selection and restriction 153
5.2.3. Statistical methods 156

5.2.4. Sensitivity analyses 159
5.3. Results 163
5.3.1. Genetic risk score and lipid traits 163
5.3.2. Genetically-predicted LDL-C and MI risk 169
5.3.3. Genetically-predicted triglycerides and MI risk 169
5.3.4. Genetically-predicted HDL-C and MI risk 169
5.3.5. Sensitivity Analyses 172
5.4. Discussion 176
Chapter 6 Observational associations of SBP and LDL-C with MI risk 181
6.1. Background 181
6.2. Methods 183
6.2.1. Study populations and settings 183
6.2.2. Statistical methods 185
6.3. Results 187
6.3.1. Baseline characteristics 187
6.3.2. Association of SBP with risk of MI 189
6.3.3. Association of LDL-C with risk of MI stratified by SBP levels. 191
6.4. Discussion 195
Chapter 7 Causal relevance of combined effects of LDL-C and SBP with MI risk 199
7.1. Background 199
7.2. Methods 201
7.2.1. Study sample 201
7.2.2. Instrumental variable selection 203
7.2.3. Statistical methods 207
7.3. Results 210
7.3.1. Genetically-predicted SBP and MI risk 210
7.3.2. Combined exposure to genetically-predicted lower LDL-C and SBP 212
7.3.3. Dose-response relationships and meta-regression results 219
7.4. Discussion 223
Chapter 8 Discussion 226
8.1. Summary of main findings 226
8.2. Population health and clinical implications 231
8.2.1. Optimal plasma lipid traits as risk markers of MI. 231
8.2.2. Generalisability of lipid-modifying and BP-lowering therapies. 233
8.2.3. Multi-factorial intervention for the primary prevention of MI. 236
8.3. Strengths and limitations 243
8.3.1. Strengths 243

8.3.2. Limitations 244
8.4. Future research 247
8.4.1. Replication of results in larger and more diverse databases 247
8.4.2. Drug target MR 247
8.4.3. Mediation and analysis of underlying mechanisms 248
8.4.4. Novel lipids and other markers to address residual risk of MI 249
8.4.5. Modifiable factors for primary prevention of MI 250
8.5. Conclusions 251
References 253
Supplementary materials 269

List of Abbreviations

ASCVD: Atherosclerotic Cardiovascular Disease

ApoA1: Apolipoprotein A1

APCSC: Asia Pacific Cohort Studies Collaboration

ApoB: Apolipoprotein B

BMI: Body Mass Index

BP: Blood Pressure

BPLTTC: Blood Pressure Lowering Treatment Trialists' Collaboration

CAD: Coronary Artery Disease

CARDIoGRAM: Coronary ARtery DIsease Genome-Wide Replication And Meta-Analysis

CCHS: Copenhagen City Heart Study

CHO-POD: Cholesterol Oxidase-Peroxidase

CGPS: Copenhagen General Population Study

CC4D: Coronary ARtery DIsease Genome wide Replication and Meta-analysis plus The Coronary Artery Disease Genetics

CHD: Coronary Heart Disease

CI: Confidence Interval

CKB: China Kadoorie Biobank

CTTC: The Cholesterol Treatment Trialists' Collaboration

CVD: Cardiovascular Disease

DBP: Diastolic Blood Pressure

EPOCH-Japan: Evidence for Cardiovascular Prevention from Observational Cohorts in Japan

ERFC: The Emerging Risk Factors Collaboration

GLGC: Global Lipids Genetics Consortium

GRS: Genetic Risk Score

HbA1c: Haemoglobin A1c

HDL-C: High-Density Lipoprotein Cholesterol

ICBP: International Consortium for Blood Pressure

IHD: Ischaemic Heart Disease

IPD: Individual Participant Data

IS: ischaemic stroke

KCPS-II: Korean Cancer Prevention Study-II

KoGES: Korean Genome and Epidemiology Study

LDL-C: Low-Density Lipoprotein Cholesterol

M-A: Meta-Analysis

MACE: Major Adverse Cardiovascular Events

MI: Myocardial Infarction

MR: Mendelian Randomisation

MVE: Major Vascular Event

MVP: Million Veteran Program

OR: Odds Ratio

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

PRS: Polygenic Risk Score

PSC: Prospective Studies Collaboration

RCTs: Randomized Controlled Trials

Remnant-C: Remnant Cholesterol

RR: Rate Ratio

SBP: Systolic Blood Pressure

SiGN: Stroke Genetics Network

SNPs: Single Nucleotide Polymorphisms

TChol: Total Cholesterol

UKB: UK Biobank

UK BiLEVE: UK Biobank Lung Exome Variant Evaluation

VLDL: Very Low-Density Lipoprotein

List of Tables

Table 2.1. Inclusion and exclusion criteria for studies included in the umbrella review 43
Table 2.2. Search strategy for the umbrella review 47
Table 2.3. Characteristics of the included observational studies 53
Table 2.4. Characteristics of the included randomised studies* 54
Table 2.5. Characteristics of the included MR studies 55
Table 2.6. Observational studies of the associations of SBP with risk of IHD, by ancestry 60
Table 2.7. Genetic studies on the association of SBP with risk of IHD, by ancestry 61
Table 2.8. RCTs on the association of SBP with risk of IHD 62
Table 2.9. Observational evidence on the association between LDL-C and the risk of IHD 68
Table 2.10. Genetic evidence on the association between LDL-C and the risk of IHD 69
Table 2.11. RCT evidence on the association between LDL-C and the risk of IHD . 70
Table 2.12. Combined effects of cholesterol and SBP on risk of IHD 73
Table 3.1. Baseline variable definitions in CKB 85
Table 3.2. Definition of prevalent and incident CVD outcomes in CKB 87
Table 3.3. Baseline variable definitions in UKB 93
Table 3.4. Definition of prevalent and incident CVD outcomes in UKB 94
Table 3.5. Baseline characteristics of participants in CKB cohort 97
Table 3.6. Baseline characteristics of participants in CKB genetic study 99
Table 3.7. Baseline characteristics of participants in the CKB biochemistry subset101
Table 3.8. Mean biochemical plasma lipid traits (SD) by region in CKB biochemistry subset 103
Table 3.9. Mean SBP and biochemical plasma lipid traits (SD) by age and sex in CKB biochemistry subset 105
Table 3.10. Baseline characteristics of participants in the UKB cohort study 107
Table 3.11. Mean SBP and biochemical plasma lipid traits (SD) by age and sex in UKB 109
Table 4.1. Baseline characteristics of participants in CKB and UKB studies 120

Table 4.2. Mean plasma levels of lipids at baseline and resurvey 122
Table 4.3 Prevalence of diabetes, by lipid traits 124
Table 4.4. Multivariable risk ratios (95%CI) of MI after adjustment for additional lipids or apolipoproteins 139
Table 5.1. Power calculation for unrestricted and restricted MR analyses 168
Table 5.2. Multivariable MR results for the association of lipid traits with MI in European and Chinese populations 174
Table 5.3. MR Sensitivity analyses for lipid traits with MI in Chinese and European populations 175
Table 6.1. Baseline characteristics of participants 188
Table 7.1. Comparison of different variant selection for SBP 206
Table 7.2. Associations of genetically-predicted SBP and risk of MI in Chinese and European populations 211
Table 7.3 Factorial MR with combined exposure to lower LDL-C and SBP (adjusted for BMI SBP GRS in CKB) 213
Table 7.4. Genetically-predicted per SD difference (0.65 mmol/L in Chinese and 0.80 mmol/L in Europeans) in LDL-C and risk of MI, stratified by SBP GRS (adjusted for BMI) 215
Table 7.5. Genetically-predicted per SD difference in SBP (21.54 mmHg in Chinese and 18.49 mmHg in Europeans) and risk of MI, stratified by LDL-C GRS (adjusted for BMI for SBP SNPs in CKB) 216
Table 7.6. Genetically-predicted per SD difference (0.65 mmol/L in Chinese and 0.80 mmol/L in Europeans) in LDL-C and risk of MI, stratified by SBP GRS (unadjusted for BMI) 217
Table 7.7. Genetically-predicted per SD difference in SBP (21.54 mmHg in Chinese and 18.49 mmHg in Europeans) and risk of MI, stratified by LDL-C GRS (unadjusted for BMI in CKB for SBP SNPs) 218
Table 7.8. Dose-response association of combined exposure to lower LDL-C and SBP in the Chinese population (GRS adjusted for BMI) 220
Table 7.9. Dose-response association of combined exposure to lower LDL-C and SBP in the Chinese population (GRS unadjusted for BMI) 221
Table 7.10. Dose-response association of combined exposure to lower LDL-C and SBP in the European populations 222
Table 8.1. Characteristics of three major polypill randomised clinical trials 240

List of Figures

Figure 1.1. Structure and relationships of major lipoprotein particles 28
Figure 2.1. MR study deduplication process 45
Figure 2.2. PRISMA Flow diagram for study selection 52
Figure 2.3. Risk of IHD per 10 mmHg lower SBP in East Asian and European populations 59
Figure 2.4. Risk of IHD per 1 mmol/L lower LDL-C in Asian and European populations 67
Figure 3.1. Map of China showing locations of the 10 CKB study sites and number of participants recruited by each site 81
Figure 3.2. Locations of UK Biobank assessment centres throughout the United Kingdom 90
Figure 4.1. Flowchart for the inclusion of study participants in observational analysis 114
Figure 4.2. Correlations between different lipid measures in Chinese and Europeans 125
Figure 4.3. Adjusted RRs (95% CI) of MI for LDL-related measures in Chinese and Europeans 128
Figure 4.4. RR (95% CI) of MI by fifths of LDL-C in CKB, excluding individuals with diabetes 129
Figure 4.5. RR (95% CI) of MI by fifths of LDL-C in UKB, after dividing the first quintile 130
Figure 4.6. Risk ratios (95%CI) of MI before and after additional adjustment for lipids in other classes 131
Figure 4.7. Adjusted RRs (95% CI) of MI for triglyceride-related measures in Chinese and Europeans 133
Figure 4.8. Adjusted RRs (95% CI) of MI for HDL-related measures in Chinese and Europeans 135
Figure 4.9. Adjusted RRs (95% CI) of MI in Chinese and Europeans 137
Figure 4.10. Risk ratios (95%CI) of MI for ratio of triglycerides to HDL-C in Chinese and Europeans 138
Figure 5.1. Mendelian randomisation and instrument variables 147

Figure 5.2. Inclusion criteria for CKB and UKB 152
Figure 5.3. Instrument selection in Chinese and European populations..... 155
Figure 5.4 Association of unrestricted and restricted GRS with lipid traits in European populations..... 165
Figure 5.5. Association of unrestricted and restricted GRS with lipid traits in the Chinese population..... 166
Figure 5.6. Genetically-predicted lipid traits and MI in Chinese and European populations..... 171
Figure 6.1. Flowchart for inclusion of study participants in Chinese and Europeans in observational analysis..... 184
Figure 6.2. Adjusted RRs for risk of MI by quintiles of SBP in Chinese and European populations..... 190
Figure 6.3. Association of LDL-C with risk of MI stratified by SBP quartiles..... 193
Figure 6.4. Association of ApoB with risk of MI stratified by SBP quartiles 194
Figure 7.1. Flowchart for inclusion in CKB and UKB 202
Figure 7.2. SNP selection for LDL-C and SBP in East Asian and European populations..... 205
Figure 7.3. Factorial design (2 × 2) with participants in four groups based on GRS 208
Figure 8.1. Summary of main findings..... 227

Chapter 1: Introduction

Myocardial infarction (MI), the most serious form of ischaemic heart disease (IHD), is a major component of cardiovascular disease (CVD), and a leading cause of death and disability worldwide.[1] In China, despite the reduction in stroke mortality over the past two decades, mortality due to MI is increasing.[2]

Elevated levels of plasma low-density lipoprotein cholesterol (LDL-C) and systolic blood pressure (SBP) are each associated with a substantial proportion of the population-attributable fraction for MI worldwide.[3] Lipid-lowering medications, such as statins, are widely used in Western populations to reduce plasma levels of LDL-C for primary prevention of MI.[4,5] However, mean LDL-C levels are about 30% lower in Chinese than in European populations, prompting an ongoing debate in China about the need for lipid-lowering medications for primary prevention of MI.[6]

There is persistent uncertainty about whether the findings from randomised controlled trials (RCTs) of statin therapy conducted in European ancestry populations can be extrapolated to the Chinese population.[7,8] Mean plasma levels of LDL-related traits, including apolipoprotein B (ApoB), non-high-density lipoprotein cholesterol (non-HDL-C), and LDL-C, are all about 30% lower in Chinese compared to European populations. While some studies have suggested that all LDL-related measures are equally reliable as risk markers for MI, there has been particular uncertainty about the shape and strength of the associations of LDL-related markers with risk of MI in the Chinese population.[9-20]

The polypill strategy was first proposed by Wald and Law as a population-level approach to cardiovascular prevention.[21] It involves the use of a single daily pill that combines several agents to simultaneously lower LDL-C, SBP, serum homocysteine,

and platelet function, thereby targeting multiple major causal risk factors at once. Their modelling suggested that such a formulation, including a statin, three classes of blood pressure lowering drugs at half dose, folic acid, and aspirin, could reduce the incidence of ischaemic heart disease by nearly 90% and stroke by around 80%.[21] This has prompted considerable interest in multi-factorial interventions, including combination therapies targeting different components of the original polypill proposal, e.g., lowering SBP and LDL-C simultaneously.[22]

Several RCTs have assessed the combined effects of lowering LDL-C and SBP for primary prevention of MI,[23] although few have included the Chinese populations. Emerging evidence also suggests that individuals from different ancestry groups may respond differently to the joint reduction of SBP and LDL-C levels.[24,25] Mendelian randomisation (MR) studies, which use genetic data as instrumental variables, provide an alternative approach to study these effects.[26]

This chapter provides an introduction to the clinical definition and characteristics of MI, including clinical subtypes, diagnosis, pathophysiology, and an overview of major risk factors for MI. In addition, selected details of lipid metabolism are outlined, along with a brief review of the observational evidence on the associations of plasma lipid traits and SBP with risk of MI. The chapter concludes with an overview of the aims and objectives of the thesis and a brief outline of the content of each chapter in the thesis.

1.1. Myocardial infarction (MI)

1.1.1. Definition

MI, also known as “heart attack”, is the most serious subtype of IHD, and typically presents with acute chest pain, breathlessness with or without systemic upset or collapse.[1] MI is characterised by the presence of acute damage to the cardiac muscle.[1] The diagnostic criteria of MI include elevated plasma levels of cardiac biomarkers, including cardiac troponin I and T levels, and electrocardiographic changes in the setting of acute myocardial ischaemia.

In clinical practice, an MI is typically classified into ST elevation MI (STEMI) or non-ST elevation MI (NSTEMI) based on ST elevation on electrocardiography (ECG) records.[27] MI can also be classified into five subtypes based on aetiology. Type 1 MI is a spontaneous event resulting from a primary coronary artery occlusion, including atherosclerotic plaque rupture.[1] Type 2 MI occurs when there is insufficient oxygen supply to the cardiac muscle, often without direct blockage of a coronary artery. Type 3 MI refers to sudden cardiac death with symptoms suggestive of ischaemia. Type 4 MI is associated with percutaneous coronary interventions. Type 5 MI is associated with coronary artery bypass grafting surgery.

1.1.2. Diagnosis and clinical presentation

Diagnosis of MI is based on elevated plasma levels of cardiac biomarkers (greater than the upper limit of normal), usually Troponin I or T, and additional evidence supporting the presence of myocardial ischaemia.[1] Myocardial ischaemia is diagnosed by the presence of clinical symptoms (chest pain or shortness of breath), ischaemic changes on the ECG, imaging tests (echocardiography and cardiac magnetic resonance imaging) that show evidence of myocardial damage consistent with occlusion of a coronary artery.

Patients with MI typically present with chest pain in the centre or left side of the chest, which radiates to the shoulders, neck, jaw, arms, or back.[1,28] A systematic review and meta-analysis involving over 1.2 million participants across 27 observational studies found that chest pain was reported in more than 70% of MI cases.[29] While previous research suggested that women may present with atypical symptoms more frequently than men, recent evidence indicates that sex differences in symptom presentation are minimal and may reflect differences in description of symptoms and interpretation rather than actual physiological differences.[29] Common non-chest pain symptoms also include pain between the shoulder blades, neck pain, palpitations, jaw pain, nausea or vomiting, fatigue, shortness of breath, indigestion, dizziness, syncope, epigastric pain, left arm and shoulder pain, and diaphoresis.[28,29]

1.1.3. Pathophysiology

A majority of MI cases are caused by the rupture of lipid-laden atherosclerotic plaques,[28,30] where inflammation triggers the formation of a platelet-rich thrombus, potentially leading to coronary artery occlusion. Complete coronary artery blockage typically results in STEMI, while partial occlusion of a coronary artery is more often associated with NSTEMI. Alternative mechanisms contributing to MI include plaque erosion, which involves thrombus formation on a fibro-intimal plaque rather than the necrotic core of a plaque.[28]

Calcific nodules, characterised by protruding calcifications that lead to thrombus formation, also cause a minority of MI cases. In addition, less common causes such as coronary artery spasm, spontaneous coronary dissection, and myocardial infarction with nonobstructive coronary arteries account for a much smaller proportion of MI cases.[28]

1.2. Burden and Risk factors for MI

1.2.1. Burden of MI globally and in China

MI is a major type of CVD. CVD is the leading cause of death and a major determinant of disability globally and accounts for one-third of all deaths in the world.[31] The Global Burden of Disease (GBD) study reported that the global burden of CVD has been steadily increasing in the 21st century, with 12.1 million CVD deaths in 1990 and 19.4 million CVD deaths in 2021, of which 8.99 million deaths were caused by IHD and 7.25 million deaths were caused by stroke.[31,32] Global age-standardised rates for deaths and prevalence of CVD have declined over the last two decades, suggesting that the overall rise in CVD worldwide has been driven by population growth and ageing. Consistent with the overall trends of reduction in CVD death rates, the age-standardised death rates of IHD have also declined globally. However, age-standardised death rates for IHD are still increasing in East Asia, including China.[31]

It has been estimated that CVD accounts for more than 40% of deaths in China.[2] Overall, the age-standardised total CVD mortality rates have declined over the past three decades.[31,32] The reduction in overall CVD mortality rates is driven chiefly by a decline in stroke mortality rates. However, the age-standardised mortality rates for IHD have been increasing in China.[2] According to the China PEACE study, there has been a rapid increase in STEMI hospital admissions from 2001 to 2011.[33] The increase in IHD mortality is further complicated by heterogeneity of disease rates between different provinces in China. The northeast region in China has the highest rates of IHD, which are almost 4-fold greater than those in provinces in the South of China.[2]

1.2.2. Major modifiable risk factors for MI

A recent report estimated that more than half of incident MI cases could be attributable to five major modifiable risk factors, including body mass index (BMI), SBP, non-HDL-C (non-high-density lipoprotein cholesterol, calculated by subtracting HDL-C from total cholesterol), smoking, and diabetes.[3] There has been a marked increase in metabolic risk factors including high fasting plasma glucose, high BMI, and high blood pressure over the past two decades, although the prevalence of high LDL-C has remained largely constant in the past decade.[34]

Of these risk factors, elevated SBP and atherogenic dyslipidaemia are the two most important contributors to the global burden of CVD.[3] According to the GBD Study,[34] high SBP is the single leading risk factor for CVD deaths worldwide, while the Global Cardiovascular Risk Consortium identified it as the largest contributor to the population-attributable fraction of CVD.[3]

Similarly, elevated levels of atherogenic lipoproteins, measured by non-HDL-C or LDL-C, are established causal risk factors and a major target for preventative strategies for CVD.[34,35]

More importantly, levels of SBP and plasma lipids are highly modifiable through widely available and effective low-cost medications.[21,36] Therefore, a better understanding of their individual and combined effects on CVD risk, particularly across diverse ancestry populations with differing baseline risk profiles, is crucial for informing global CVD prevention strategies.

1.3. Plasma lipid traits and risk of MI

1.3.1 Types of plasma lipids and their metabolism

Lipids refer to a diverse group of hydrophobic compounds circulating in plasma. The most clinically relevant lipids include cholesterol and triglycerides. Cholesterol has an important biological role in human physiology. They are precursors for synthesis of steroids, vitamin D, oxysterol, and bile acids and are integral components of cell membranes and are required for the activation of signalling molecules in neurons. Triglycerides, which are made up of free fatty acids and glycerol, are important sources of energy in the body.[37]

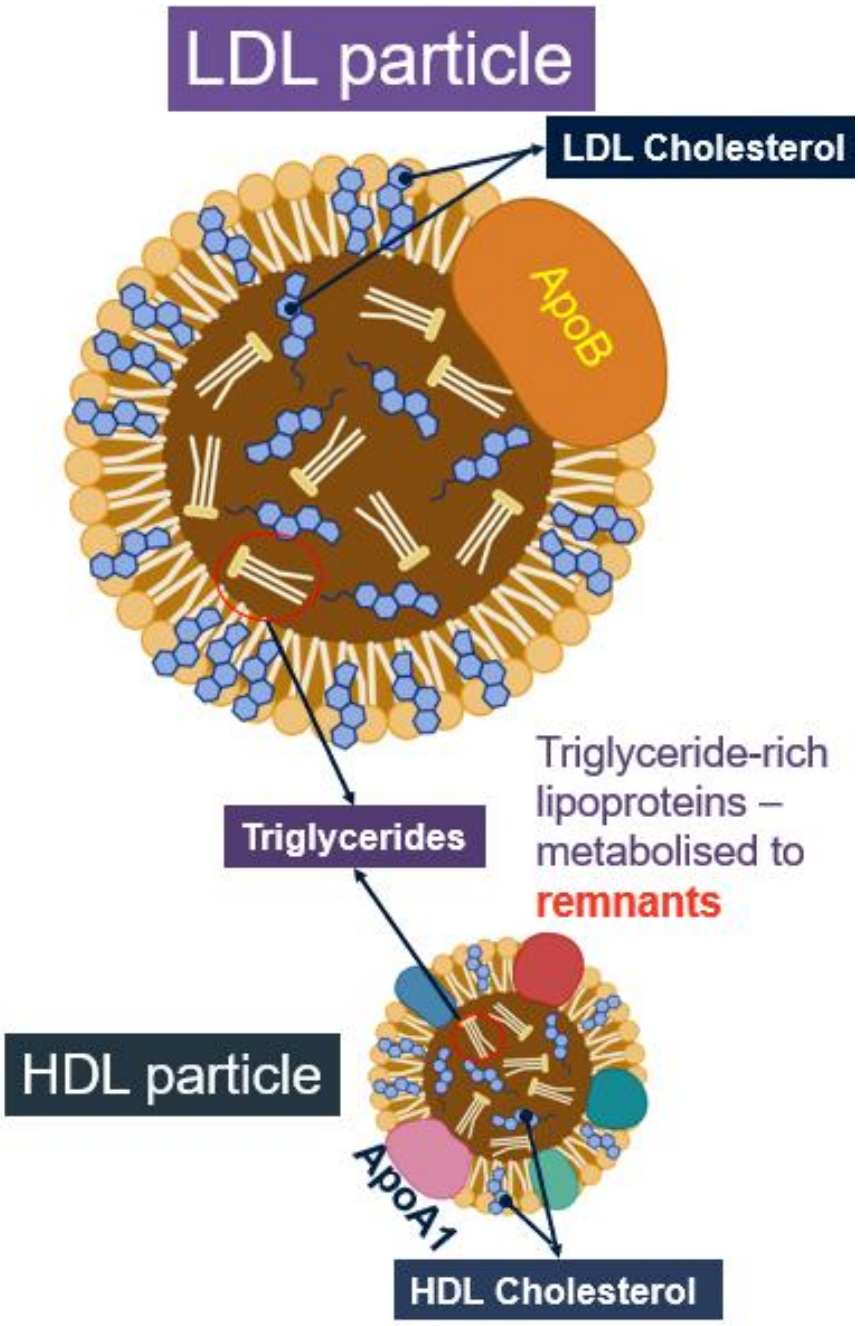
Since lipids are hydrophobic, they are transported by macromolecules called lipoproteins.[37] Lipoproteins are distinguished by their size and density. Lipoproteins have hydrophobic cores with cholesteryl esters and hydrophilic coats with free cholesterol. Triglycerides are mainly carried by chylomicrons (CM) and very low-density lipoprotein (VLDL). Cholesterol is mainly carried by low-density lipoprotein (LDL) and high-density lipoproteins (HDL).

Lipoproteins are complexes that transport cholesterol and triglycerides through the bloodstream, as these lipids are not water soluble.[37] Apolipoproteins refer to the protein components of lipoproteins. Apolipoprotein B (ApoB) is the primary structural protein of atherogenic lipoproteins, including VLDL, IDL, LDL, and lipoprotein(a).[37] In the plasma, LDL has an exact 1 to 1 ratio with Apolipoprotein B (ApoB) since one LDL particle has one ApoB particle. Hence, plasma LDL-C and ApoB levels are directly correlated with each other. Likewise, one HDL particle typically has two to four molecules of apolipoprotein A1 (ApoA1), depending on the size and maturity of the particles. Plasma HDL-C levels are also highly correlated with ApoA1 levels.[37]

Figure 1.1 shows the structural components and relationships between LDL and

HDL particles, with their key apolipoproteins (ApoB and ApoA1), cholesterol, and triglyceride content.

Figure 1.1. Structure and relationships of major lipoprotein particles



Abbreviations: **ApoA1**: apolipoprotein A1; **ApoB**: apolipoprotein B; **HDL**: high-density lipoprotein; **LDL**: low-density lipoprotein.
Created with BioRender.com.

1.3.2. Lipid metabolism and cardiovascular disease

The pathological link between lipid metabolism and CVD is well established, as dyslipidaemia, defined as abnormal plasma levels of lipids in the bloodstream, contributes to the development of atherosclerosis.[38] Lipoproteins play a pivotal role in the maintenance of cellular function and energy homeostasis. However, when there is an imbalance in lipid metabolism, cholesterol accumulates in arterial walls, leading to plaque formation and vascular inflammation. These atherosclerotic plaques result in narrowing of the coronary arteries and reduction in blood supply to the heart muscle.[38]

In clinical practice, plasma levels of total cholesterol (TChol), LDL-C, HDL-C, and non-HDL-C are widely used biochemical measures to assess CVD risk.[5] TChol represents the overall amount of cholesterol present in the blood, and includes LDL-C, HDL-C, intermediate-density lipoproteins (IDL), and very low-density lipoproteins (VLDL). LDL-C is chiefly composed of cholesterol and ApoB, and these allow it to bind to cell receptors and deliver cholesterol to cells. Non-HDL-C levels are calculated by subtracting HDL-C levels from TChol levels and represent the TChol carried by all atherogenic lipoproteins, including LDL-C, IDL, and VLDL. Remnant cholesterol (remnant-C) refers to cholesterol contained in triglyceride-rich lipoproteins and is estimated by subtracting HDL-C and LDL-C from TChol.[39]

1.3.3. Epidemiological evidence of plasma lipid traits and risk of MI

LDL-related lipid traits, including LDL-C, non-HDL-C, and ApoB, are well-established risk factors for MI. The Emerging Risk Factors Collaboration (ERFC)[9] meta-analysis combined data from 58 prospective cohort studies, conducted mainly in Europe and North America. They reported that 1 mmol/L lower non-HDL-C was associated with about 35% lower risk of IHD.[9] ERFC also reported that the strength of associations with IHD for ApoB and non-HDL-C were similar. MR studies confirmed the causal relevance of LDL-C for risk of MI.[40,41] Evidence from RCTs further demonstrated that lowering LDL-C levels reduced the risk of MI. The Cholesterol Treatment Trialists' Collaboration (CTTC),[42] which aggregated individual-level data from 27 RCTs, reported that a 1 mmol/L reduction in LDL-C achieved by statin therapy was associated with a 25% reduction in risk of major vascular events. The larger effect sizes reported by the ERFC compared with the CTT trials likely reflect methodological and population differences. The stronger associations in ERFC could reflect the longer duration of follow-up and correction for regression dilution, compared with the shorter (~5 years) durations in higher-risk populations in RCTs.[9,42]

The results of observational studies of the associations of plasma levels of triglycerides with risk of MI have been conflicting.[9] Observational studies have indicated that higher plasma levels of triglycerides are associated with higher risk of MI. However, these associations were greatly attenuated after adjusting for HDL-C and attenuated to the null after further adjustment for non-HDL-C.[9] Commonly available lipid-modifying medications, including Proprotein convertase subtilisin/kexin type 9 (PCSK9) inhibitors and fibrates, lower triglyceride levels.[43] Fibrates can significantly reduce triglyceride levels and have been shown to reduce CVD risk as

monotherapy. However, when used in combination with statins, no further risk reduction was observed.[43] Genetic analyses have provided additional support for triglyceride-rich lipoproteins or remnant cholesterol as causal risk factors for CVD, rather than triglycerides themselves.[44,45]

HDL-C has been proposed as an inversely related lipid marker of CVD risk since the 1960s (the HDL hypothesis).[46] The Framingham Heart Study reported an inverse association between plasma levels of HDL-C and CHD,[47] and these observations have been replicated by multiple prospective studies since then. Prospective studies conducted in diverse populations have also confirmed that low plasma levels of HDL-C were strongly and independently associated with risk of MI.[46] However, genetic studies of HDL-C, after adjusting for triglyceride levels, have reported no associations between HDL-C levels and MI, which challenges the causal relevance of HDL-C for MI.[46] Several RCTs testing HDL-raising medications have failed to demonstrate a reduction in CVD risk. Reports related to cholesteryl ester transfer protein (CETP) inhibitors, including torcetrapib[48] and dalcetrapib,[49] showed no benefits on risk of CVD despite significant increases in HDL-C levels. The discrepant findings between observational associations and null results observed in RCTs have prompted analyses of associations of related biomarkers, including ApoA1, and HDL function with CVD risk.[49]

1.4. Blood pressure (BP) and risk of MI

1.4.1. Physiological link of BP and cardiovascular system

BP is essential for maintaining adequate perfusion of all tissues and organs in the body, including the heart. However, elevated levels of BP, or hypertension, cause excessive mechanical stress on the arterial walls, contributing to the pathogenesis of CVD.[50] Elevated levels of BP induce vascular endothelial dysfunction, characterised by impaired vasodilation, increased oxidative stress, and inflammation. These physiological processes accelerate the development of atherosclerosis. In the coronary arteries, hypertension promotes atheroma formation and can destabilise plaques, thereby increasing the risk of plaque rupture and thrombus formation. Acute occlusion of coronary arteries ultimately leads to MI and other ischaemic events.[50]

1.4.2. Epidemiological evidence of SBP and risk of MI

Elevated SBP is one of the leading contributors to the global burden of CVD.[3] High SBP, or hypertension, is considered a highly prevalent risk factor that is causally and strongly associated with risk of MI.[51]

The Prospective Studies Collaboration (PSC) estimated that higher usual levels of SBP were positively and linearly associated with higher risk of mortality from IHD.[52] At ages 40–69 years, each 20 mmHg higher usual SBP was associated with approximately a twofold difference in IHD death rates, with no evidence of a threshold down to at least 115 mmHg SBP. Other observational studies conducted in diverse ancestry populations have provided additional support for the strong and independent role of SBP for MI risk.[52-55] MR analyses conducted in diverse ancestry populations have provided additional support for the causal relevance of SBP for risk of IHD.[56-58] The Blood Pressure Lowering Treatment Trialists Collaboration

(BPLTTC) estimated that lowering SBP by 10 mmHg resulted in a 10% risk reduction in IHD over 5 years.[59] The smaller risk reduction observed in RCTs compared with observational studies is likely due to the shorter duration of follow-up in trials, during which the full benefits of long-term BP-lowering may not yet have been realised.

1.5. Combined effects of LDL-C and SBP on risk of MI

Studies on population-level CVD prevention have explored multi-factorial interventions that simultaneously lower several major cardiovascular risk factors, including plasma levels of LDL-C and levels of SBP.[21] A "polypill" strategy was proposed to prevent over 80% of CVD cases.[21] The original polypill strategy suggested a combination of a statin, several BP-lowering agents at low doses, folic acid to reduce homocysteine, and aspirin to modify platelet function.[21]

Since the original proposal of the polypill strategy, major meta-analyses of observational studies have examined the combined associations of levels of LDL-C and SBP with risk of CVD.[9,52,54,60] These meta-analyses reported differences in these associations across CVD subtypes and in diverse ancestry populations.[9,54] However, several meta-analyses suggested possible interactions or effect modifications between SBP and LDL-C on risk of CVD.[9,54,60] The latter findings question whether a polypill strategy could have uniform proportional effects on different CVD subtypes or in diverse ancestry populations. Several RCTs have tested the efficacy of some polypill formulations and reported promising results.[23] In some of the early trials of fixed-dose combinations, including the UMPIRE trial, the primary endpoints included adherence, reflecting the rationale that fixed-dose combinations could improve long-term compliance with preventive therapy.[61,62] This is particularly relevant as, despite

the availability of effective and relatively low-cost treatments such as BP-lowering drugs and statins, uptake remains low in many low- and middle-income countries (LMICs) such as China.[63,64] Evidence from a systematic review also confirms that polypill strategies are consistently associated with improved adherence compared to multi-pill regimens.[62] However, these RCTs have been constrained by short duration of treatment and limited ethnic diversity, leaving uncertainty about the long-term efficacy of such strategies in diverse populations.

1.6. The need for further research

Mean levels of LDL-C are lower in Chinese than in Europeans, but whether such differences influence the strength of associations of plasma lipid traits with MI between Chinese and European populations is uncertain.[6] The ERFC meta-analysis of prospective studies advocated limiting lipid measures for prediction of risk of vascular disease to total cholesterol and HDL-C or apolipoproteins without any requirement for fasting or including triglyceride measurements.[9] In particular, the ERFC reported that additional lipid measures including ApoB and non-HDL-C yielded only modest improvements in risk prediction of CVD beyond conventional lipids (LDL-C, HDL-C and triglycerides).[9] Subsequently, in contrast to the conclusions from ERFC, several studies reported that ApoB is a more accurate marker of cardiovascular risk than LDL-C,[10-20] but no previous studies compared the associations of different lipid fractions and apolipoproteins with risk of MI in the Chinese population. In addition, recent reports have advocated targeting of remnant-C as an adjunct to LDL-C-lowering.[44] However, little is known about the comparative importance of remnant-C versus triglycerides and ApoA1 versus HDL-C for risk of MI between Chinese and Europeans.

The causal relevance of LDL-C is well-established in Western populations and has been supported by extensive randomised trial evidence.[42] Statins are the standard drug treatment for primary prevention of MI in all guidelines for CVD prevention in all ancestry populations.[4,5] However, there is an ongoing uncertainty about the efficacy and indications for lipid-lowering medications for prevention of MI in Chinese adults.[65,66] The mean plasma levels of LDL-C are lower in Chinese populations compared to European populations, raising questions about the efficacy of LDL-lowering at lower baseline levels of LDL-C.[6] No large-scale RCTs have evaluated

the effects of LDL-lowering on risk of first incident MI in Chinese populations, leaving clinical guidelines in China reliant on RCTs from Western populations.[8,67] Nonetheless, uncertainty persists about whether the findings from RCTs conducted primarily in European populations can be extrapolated to the Chinese population.

Several RCTs have tested the efficacy of some polypill formulations, yet the three major trials testing such treatments each had important limitations.[23] Collectively, these RCTs included 18,000 participants and had a median duration of follow-up of 5 years. The treatment duration of LDL-lowering in the RCTs is not sufficient to assess their lifelong effects. Furthermore, participants selected for inclusion in trials typically were people with intermediate CVD risk, restricting the generalisability of findings to population-wide interventions. Ancestry differences among participants in these trials were limited. Most participants were recruited from Europe, North America, or Australia, with fewer than 5% from China. These discrepancies highlight uncertainty about whether findings from RCTs in Western populations can be generalised to the Chinese population.

1.7. Aims and objectives

The overall aim of this thesis is to provide a comprehensive trans-ethnic investigation of the separate and combined effects of major plasma lipid traits and SBP on the risk of MI in Chinese and European populations.

The objectives of this thesis are:

- To systematically review the epidemiological evidence on the associations of LDL-C and SBP with risk of IHD in European and East Asian populations.
- To compare the shape and strength of observational associations of LDL-related, triglyceride-related, and HDL-related traits with risk of MI in Chinese and European populations.
- To investigate the causal relevance of genetically-predicted differences in plasma lipid traits on risk of MI in both Chinese and European populations.
- To assess the associations of levels of SBP with risk of MI in Chinese and European populations and examine possible interactions for the associations of levels of SBP and plasma levels of LDL-C with risk of MI.
- To estimate the combined associations of genetically-predicted levels of SBP and LDL-C on risk of MI in Chinese and European populations.

1.8. Structure of the thesis

Chapter 2 provides an umbrella review of the available evidence on the associations of LDL-C and SBP with MI risk, including evidence from observational, genetic (MR), and randomised sources from East Asian and European ancestry populations. The review also includes studies on the combined associations of levels of SBP and LDL-C with risk of MI. The aim of Chapter 2 is to review the available worldwide evidence on these associations and identify gaps in the available evidence to prioritise for analyses in subsequent chapters in this thesis.

Chapter 3 describes the study populations and methods. It provides details of the design, data collection, and available genetic data for the China Kadoorie Biobank (CKB) and UK Biobank (UKB), and describes the statistical methods used in the analytical chapters.

Chapter 4 compares the shape and strength of associations of plasma lipids, lipid fractions and apolipoproteins with risk of MI, including: (i) non-HDL-C or ApoB vs. LDL-C, (ii) remnant-C vs. triglycerides, and (iii) ApoA1 vs. HDL-C in CKB and UKB. This chapter provides comparisons of the associations of major plasma lipid traits with MI in the Chinese population where the distribution of lipids differs from European populations and provides updated observational evidence for the associations of different lipid traits with risk of MI in both ancestry populations.

Chapter 5 uses MR approaches to evaluate the causal relevance of major plasma lipid traits (LDL-C, triglycerides, and HDL-C) for risk of MI in Chinese and European populations. This chapter triangulates the observational evidence from Chapter 4 and addresses research gaps in the literature on the causal relevance of conventional and novel lipid traits for MI in the Chinese population.

Chapter 6 assesses the associations of SBP levels with risk of MI in Chinese and European populations, together with the combined associations of levels of SBP and LDL-related lipid traits with risk of MI. This chapter addresses a gap in the available evidence on the combined effects of SBP and LDL-C on MI risk in the Chinese population.

Chapter 7 examines the combined effects of genetically-predicted lowering of LDL-C and SBP using a factorial MR approach, which should predict the effects of combined lowering of LDL-C and SBP on risk of MI. This chapter starts by examining the associations of genetically-predicted SBP in Chinese and European populations, replicating some of the previous analyses. Participants were then divided into factorial groups for the assessment of the joint effect of lowering SBP and LDL-C. The analyses in this chapter provide direct comparisons of the observational evidence from Chapter 6 and indications of the likely effects on MI of a combined intervention aimed to lower levels of both LDL-C and SBP.

Chapter 8 provides a summary and interpretation of the main findings of the thesis and an evaluation of their consistency with results of previous studies in the worldwide literature. It highlights the value of research on the separate and combined effects of lowering both lipids and BP on risk of MI in both Chinese and European populations. The final chapter also discusses the broader public health and clinical implications of the findings, outlining potential applications for improving strategies for prevention of CVD. It also identifies directions for future research and reviews the strengths and limitations of this thesis.

Chapter 2: Umbrella reviews of LDL-C and SBP and risk of IHD

2.1. Background

Umbrella reviews aim to systematically collect and assess multiple systematic reviews and meta-analyses conducted on a broad, but well-defined, research topic.[68,69] Umbrella reviews are designed to summarise the most reliable evidence and combine the results of previously published reviews and provide an overview of a particular topic. An umbrella review can also compare multiple interventions and evidence sources, therefore provides an overview of the strengths and weaknesses of the evidence from multiple sources.[68,69]

This chapter includes an umbrella review of the best available worldwide evidence on the individual and combined associations of LDL-C and SBP with risk of IHD. Because MI subtyping in epidemiological studies has only become common in recent years, this chapter uses IHD as the main outcome to include a broader and more relevant range of studies.[70,71]

SBP and LDL-C are both well-established causal risk factors for IHD.[36,42] However, the estimates of associations of SBP and LDL-C with risk of IHD differ between different sources, including observational studies, RCTs, and MR studies.[36,40,42,52,57,60] Recent studies have also suggested potential ethnic differences in the associations of SBP and LDL-C with risk of IHD.[53,57,72,73] In addition, the combined effects of lowering levels of LDL-C and SBP on risk of IHD require further investigation, given the results from RCTs on polypills appear to show weaker effects on IHD risk reduction compared to those that were predicted based on observational studies.[21,23] There have been multiple meta-analyses of observational studies, RCTs

and MR studies of the associations of LDL-C and SBP with risk of IHD.[36,40,42,52,57,60] However, there have been no overviews of the individual and combined effects of these markers for risk of IHD. This review intends to identify evidence gaps and provide a foundation for comparing the results from the different types of evidence in each ancestry population.

The aim of this chapter is to summarise the most reliable available evidence on the individual and combined associations of plasma levels of LDL-C and levels of SBP with risk of IHD from different sources of evidence, including observational, randomised, and genetic studies. This umbrella review was designed to address the following questions:

1. To compare the observational, genetic, and randomised evidence for the associations of SBP with risk of first incident cases of IHD in Chinese and European populations across different evidence types.
2. To contrast the observational, genetic, and randomised evidence for the associations of plasma levels of LDL-C with risk of first incident cases of IHD in Chinese and European populations across different evidence types.
3. To investigate the observational, genetic, and randomised evidence for the joint associations of levels of SBP and plasma levels of LDL-C with risk of first incident cases of IHD in Chinese and European populations across different evidence types.

2.2. Methods

2.2.1. Eligibility criteria

Table 2.1 outlines the selection criteria for included studies in this umbrella review. Large prospective biobank studies, defined as having more than 100,000 participants, meta-analyses of cohort studies and RCTs, and MR studies were included. Exposures included SBP, LDL-C, non-HDL-C or TChol in meta-analyses of observational studies, BP- or LDL-C-lowering medication in meta-analyses of RCTs, and genetically-predicted SBP levels or genetically-predicted LDL-C levels in MR studies. Outcomes included IHD (including CHD, MI and CAD). Studies on composite CVD outcomes, for example, major vascular events (MVE, defined as a composite of coronary death, nonfatal MI, or coronary revascularisation),[74] were also included and assessed on an individual study basis. Only English language publications were considered due to feasibility constraints in translation and because most large-scale cardiovascular studies are published in English. No temporal limits were applied.

Narrative and literature reviews were excluded. Cohorts restricted to specific subgroups based on disease status, for example, in diabetic populations, or secondary-prevention populations were excluded. Studies on lifestyle or behavioural interventions were also excluded. Studies focusing exclusively on other CVD endpoints (e.g., stroke or heart failure) were excluded.

Table 2.1. Inclusion and exclusion criteria for studies included in the umbrella review

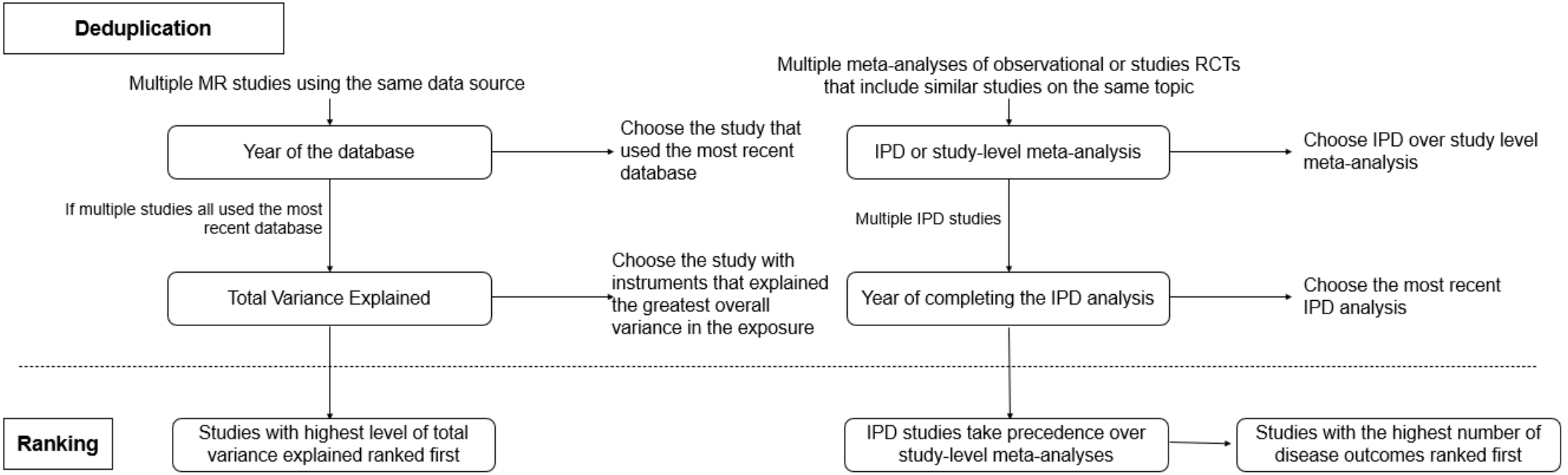
Criteria	Inclusion	Exclusion
Study Design	• Large prospective biobank studies (with > 100,000 participants) • Meta-analyses of prospective cohort studies • Meta-analyses of RCTs • MR studies	• All other types of studies, for example, narrative synthesis and literature review.
Participants	• Population-based studies of healthy adults (age>18 yrs old) • Studies comparing different populations and one of the groups including the general adult population	• Studies that were restricted to populations with specific disease statuses. • Studies conducted exclusively in populations with a CVD history (secondary prevention).
Exposures	• For meta-analyses of prospective cohort studies, studies with exposures as SBP, LDL-C, non-HDL-C, or TChol • For meta-analyses of RCTs, studies with interventions as BP- or LDL-C-lowering medication • For MR studies, studies using genetically-predicted differences in SBP or LDL-C levels as exposures	• Studies that did not report differences or reductions in SBP or LDL-C levels but only reported drug effects • Studies focusing solely on lifestyle or behavioural interventions
Comparators	• Studies that compare different levels of SBP and LDL-C, including different baseline levels in observational analyses, reduction by medications in RCTs, and genetically-predicted differences in MR studies	• Studies that did not quantitatively compare different levels of SBP and LDL-C
Outcomes	• Studies with IHD (including CHD, MI, and CAD) as outcomes • Studies reported composite CVD outcomes, such as MVE, defined as a composite of coronary death, nonfatal MI, or coronary revascularisation.	• Studies focusing exclusively on other types of CVD, such as stroke, peripheral arterial disease, heart failure, etc.
Year and Language	• Studies published in English • No year restriction	• Studies published in languages other than English

Abbreviations: **BP**: Blood Pressure; **CVD**: Cardiovascular Disease; **CHD**: Coronary Heart Disease; **IHD**: Ischaemic Heart Disease; **LDL-C**: Low-Density Lipoprotein Cholesterol; **MR**: Mendelian Randomisation; **RCTs**: Randomised Controlled Trials; **SBP**: Systolic Blood Pressure; **TChol**: Total Cholesterol; **MI**: Myocardial Infarction; **CAD**: Coronary Artery Disease.

2.2.2. Study deduplication and ranking

Details of the criteria used for selection and deduplication of studies are presented in ***Figure 2.1.*** Duplicates were removed because some meta-analyses were updated, and some MR studies used the same databases published in different years. A hierarchical set of rules was applied. First, among multiple MR studies using the same data sources, the study based on the most recent database was chosen, and if more than one study used that database, the study whose instruments explained the greatest overall variance in the exposure was selected. Second, among multiple meta-analyses of observational studies or RCTs addressing the same topic, IPD meta-analyses took priority over study-level meta-analyses. If more than one IPD analysis existed, the most recently published report was chosen. Finally, for ranking, all IPD meta-analyses were prioritised above study-level meta-analyses, and within each category, studies were ranked by the highest number of disease outcomes.

Figure 2.1. MR study deduplication process



Abbreviations: **MR**: Mendelian randomisation; **IPD**: individual participant data; **RCTs**: randomised controlled trials.

2.2.3. Information sources

Literature search strategies were developed using medical subject headings and texts related to the exposures and outcomes in Medline (Ovid interface) and Embase (Ovid interface) from inception to 2023. A medical librarian was consulted to refine the search terms. The literature search was limited to English language and human subjects. The search strategy consisted of four parts: 1) BP; 2) cholesterol; 3) CVD; 4) meta-analysis, biobank, or MR studies. The detailed search strategy is presented in ***Table 2.2***.

Table 2.2. Search strategy for the umbrella review

Item	Search term	Number of studies
1	blood pressure.ti.	155,232
2	cholesterol.ti.	126,277
3	lipid$.ti.	338,092
4	low-density lipoprotein.ti.	26,629
5	LDL.ti.	23,887
6	1 or 2 or 3 or 4 or 5	643,498
7	cardiovascular disease$.ab,ti.	508,537
8	coronary artery disease.ab,ti.	245,021
9	coronary heart disease.ab,ti.	125,844
10	myocardial infarction.ab,ti.	500,490
11	Ischaemic heart disease.ab,ti.	69,694
12	ischaemic heart disease.ab,ti.	22,305
13	acute coronary syndrome.ab,ti.	75,781
14	stroke.ab,ti.	753,774
15	cerebrovascular accident$.ab,ti.	19,815
16	cerebrovascular disease$.ab,ti.	58,442
17	cardiovascular event$.ab,ti.	124,289
18	cardiovascular death$.ab,ti.	27,311
19	cardiovascular mortality.ab,ti.	40,767
20	7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19	2,055,879
21	meta-analysis.ti,pt.	409,749
22	individual participant$.ab,ti.	8,148
23	consortium.ab,ti.	88,533
24	biobank.af.	45,701
25	mendelian randomi?ation.ab,kf,ti.	12,208
26	21 or 22 or 23 or 24 or 25	553,330
27	6 and 20 and 26	2,378
28	limit 27 to english language	2,341
29	limit 28 to humans	2,175
30	remove duplicates from 29	1,407

*Searches were conducted in: Embase 1974 to January 2023; Medline (Ovid MEDLINE® Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE® Daily and Ovid MEDLINE®) 1946 to January 2023.

2.2.4. Data Extraction

The results of the literature search were uploaded to Covidence.[75] Two rounds of screening were conducted. The first round of screening focused on titles and abstracts to screen for initial eligibility. The second round focused on full-text screening. All the studies were further evaluated after full-text screening to check for eligibility prior to their inclusion.

Firstly, we extracted descriptive information about the included studies, including author-year, collaborations, region, and the study title to present an overview of the selected studies. Secondly, detailed data on the effect estimates were extracted. Information was extracted on the number of studies, participants, events, and study characteristics including length of follow-up, mean age of the participants, and percentage of females. All the reported effect sizes were transformed to estimates for effects on IHD per 10 mmHg lower SBP or 1 mmol/L lower LDL-C, assuming a log-linear association with disease outcomes. The effect sizes were represented as Rate Ratios (RRs).

Furthermore, for observational analyses, we recorded if the analyses had corrected for regression dilution bias (denoted as usual BP or cholesterol level and marked with ✓✓ in the tables). Regression dilution bias refers to the systematic underestimation of risk associations caused by within-person variability and measurement error in baseline risk factor levels, which leads to weaker observed associations than those with the “usual” long-term levels.[76] The degree of adjustment was recorded and denoted as + for age- and sex-adjusted, ++ for additionally adjusting for lifestyle factors (smoking and/or alcohol use) and/or BMI, +++ for also adjusting for other vascular risk factors (e.g., BP and history of diabetes). For MR studies, the percentage of the variance of the exposure explained by the genetic

instrument (R^2) was also recorded. R^2 was recorded instead of the F statistic, because F is strongly influenced by sample size and can appear large in very large studies even when the instrument explains only a negligible proportion of the exposure variance, whereas R^2 directly reflects the proportion of variance explained and is therefore more informative for comparing instrument strength across studies.[77]

2.3. Results

2.3.1. Overview of the included studies

The PRISMA diagram provided the study selection process as presented in ***Figure 2.2***. There were 1,002 and 1,173 reports identified from Medline and Embase, respectively. After removing duplicates, 1,324 studies were included in the first round of screening. After removing 1,150 studies deemed irrelevant from the title and abstract screening, 174 full-texts were screened for eligibility. The most common reason for full-text exclusion was that the studies did not report differences in SBP or LDL-C level (n = 60) and other reasons included unsuitable study design (e.g., conference abstracts or scoping reviews, n = 27), duplicated data sources (n = 29), irrelevant endpoints (e.g., outcome is other CVD, n = 19), no reporting in the healthy population (studies that focused on a specific morbidity group, n = 2), and focusing only on secondary prevention (n = 4). After full-text selection, 33 studies fulfilled the inclusion criteria, including 9 meta-analyses of observational studies or large biobanks published from 2002 to 2018 (3 on SBP,[52,53,55] 4 on LDL-C,[9,60,72,78] and 1 on SBP and LDL-C[54] and 1 on their combined effects[79]), 17 MR studies published from 2012 to 2022 (7 on SBP,[56,57,80-84] 8 on LDL-C,[40,41,85-90] 1 included the individual effects of both LDL-C and SBP,[58] and 1 focused on the combined effects of LDL-C and SBP,[74] and 7 meta-analyses of RCTs published from 2009 to 2021 (2 on SBP[36,59] and 5 on LDL-C[42,91-94]). ***Table 2.3*** summarises the characteristics of the included meta-analyses of the observational studies, including details on region, study design, number of participants, exposures assessed, and outcomes. ***Table 2.4*** presents the characteristics of the included meta-analyses of RCTs, outlining the interventions (e.g., statin or BP-lowering therapy) and their reported outcomes.

Table 2.5 provides an overview of the MR studies with the exposures, outcome databases, and the variance in exposure explained by the genetic instruments.

For the associations between SBP and LDL-C with IHD, the study with the largest number of participants was PSC (61 studies with a total of around one million participants).[60] When studying the associations between LDL-C and IHD, the meta-analysis with the largest number of included studies was ERFC (68 studies with 302,430 participants).[9] For the effects of BP-lowering medications on the risk of CVD, BPLTTC[59] is the most recent IPD analysis, with 38 primary prevention trials (186,128 participants). CTTC[42] represents the most recent IPD analysis for the effects of LDL-C-lowering on CVD risks, with 27 primary prevention trials (174,149 participants). We further deduplicated two meta-analyses of RCTs,[92,93] and six MR studies.[80,81,83,84,86,87,95]

Figure 2.2. PRISMA Flow diagram for study selection

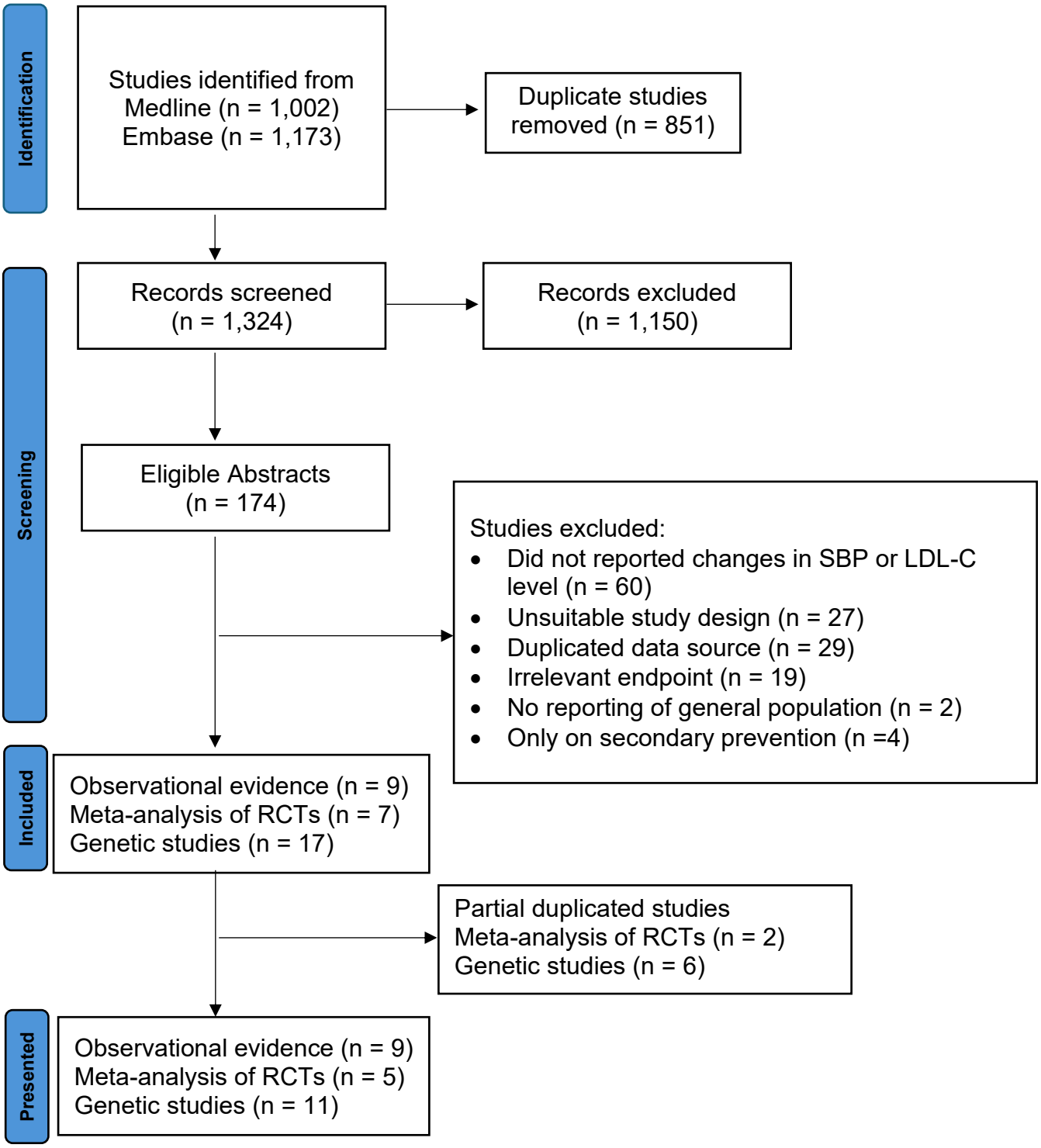


Table 2.3. Characteristics of the included observational studies

Author, Year	Collaboration	Region	Design	No. of Studies	No. of Participants	Exposure	Outcome
Studies in predominantly European populations							
Di Angelantonio et al., 2009[9]	ERFC	Mainly Europe	IPD	68	302,430	Usual non-HDL-C	IHD
Lewington et al., 2007[60]	PSC	Mainly Europe	IPD	61	892,337	Usual TChol	Fatal IHD
Lewington et al., 2002[52]	PSC	Mainly Europe	IPD	61	958,074	Usual SBP	Fatal IHD
Studies in predominantly Asian populations							
Holmes et al., 2018[78]	CKB	China	Biobank	1	4,662	Measured LDL-C	IHD
Satoh et al., 2015	EPOCH–Japan	Japan	IPD	11	73,916	Measured SBP/TChol	Fatal IHD
Kengne et al., 2007[72]	APCSC	Asia-Pacific	IPD	30	333,533	Usual TChol	IHD
Rodgers et al., 2005[79]	APCSC	Asia-Pacific	IPD	36	380,216	Usual TChol/SBP	IHD
Lawes et al., 2003[53]	APCSC	Asia-Pacific	IPD	37	425,325	Usual SBP	IHD
Lacey et al., 2018[55]	CKB	China	Biobank	1	489,125	Usual SBP	IHD

Abbreviations: **ERFC**: Emerging Risk Factors Collaboration; **PSC**: Prospective Studies Collaboration; **CKB**: China Kadoorie Biobank; **EPOCH-JAPAN**: Evidence for Cardiovascular Prevention from Observational Cohorts in Japan; **APCSC**: Asia Pacific Cohort Studies Collaboration; **IPD**: individual participant data; **SBP**: systolic blood pressure; **LDL-C**: low-density lipoprotein cholesterol; **non-HDL-C**: non-high-density lipoprotein cholesterol; **TChol**: total cholesterol; **IHD**: ischaemic heart disease.

Table 2.4. Characteristics of the included randomised studies*

Author, Year	Collaboration	Design	No. of Studies	No. of Participants	Intervention	Outcome
Wang et al., 2020[94]	N/A	M-A	11	81,786	Statin	MVE†
Genser et al., 2019[92]	*N/A*	*M-A*	*17*	*97,861*	*Statin*	*IHD*
Law et al., 2009[36]	N/A	M-A	26	108,297	BP-lowering**	IHD
Silverman et al., 2016[93]	*N/A*	*M-A*	25	177,088	*LDL-C-lowering§*	*MVE†*
Briel et al., 2009[91]	N/A	M-A	95	299,310	LDL-C-lowering§	IHD
Mihaylova et al., 2012[42]	CTTC	IPD	27	174,149	Statin	IHD
Rahimi et al., 2021[59]	BPLTTC	IPD	38	186,128	BP-lowering**	IHD

* No identified studies reported results separated by ethnicities.
**Included all BP-lowering medication classes
†MVE: composite of cardiovascular mortality, MI, coronary revascularization, or stroke
§Included trials on statins and non-statin medications to lower LDL-C.
Abbreviations: **N/A**: not applicable; **M-A**: meta-analysis; **IPD**: individual participant data; **CTTC**: Cholesterol Treatment Trialists' Collaboration; **BPLTTC**: Blood Pressure Lowering Treatment Trialists' Collaboration; **BP**: blood pressure; **LDL-C**: low-density lipoprotein cholesterol; **IHD**: ischaemic heart disease; **MVE**: major vascular events; **MI**: myocardial infarction.

Table 2.5. Characteristics of the included MR studies

Author, Year	Population	Exposures	Exposure Database	Outcome	Outcome Database	No. of SNPs	Variance Explained, %
Studies in predominantly European populations							
Cho et al., 2022[56]	European	SBP	MVP	CVD*	UKB	272	NR
Gill et al., 2020[84]	European	BP	CplusC4D	IHD	UKB	34	NR
Ference et al., 2014[80]	European	SBP	ICBP	IHD	CARDIoGRAM	12	NR
Higgins et al., 2021[81]	European	SBP	ICBP	IHD	UKB	256	2.1
Malik et al., 2021[82]	European	SBP	ICBP	IHD	UKB	253	4.8
Wan et al., 2021	European	SBP	ICBP	IHD	UKB	327	5.7
Ference et al., 2012	European	LDL-C	GLGC	IHD	CARDIoGRAM	9	NR
Burgess et al., 2014[86]	European	LDL-C	GLGC	IHD	CARDIoGRAM	162	NR
Richardson et al., 2020	European	LDL-C	UKB	IHD	CC4D	209	NR
Valdes-Marquez et al., 2019[89]	European	LDL-C	GLGC	IHD	CARDIoGRAM MEGASTROKE	62	4.0
Holmes et al., 2015	European	LDL-C	-	IHD	-	42	4.3
White et al., 2016	European	LDL-C	GLGC	IHD	CC4D	130	7.9
Allara et al., 2019[85]	European	LDL-C	GLGC	IHD	UKB	184	14.6
Ference et al., 2019[74]	European	SBP+LDL-C	Liu et al.	IHD	UKB	61/100	NR
Studies in predominantly East Asian populations							
Clarke et al., 2014[57]	Chinese	SBP	-	IHD	CKB	460	NR
Wang et al., 2022[58]	Japanese	SBP+LDL-C	Biobank Japan	IHD	Biobank Japan	61/29	NR
Lee et al., 2020[88]	Korean	LDL-C	KoGES	IHD	KCPS-II	20	NR

*CVD was defined as a composite of first incident myocardial infarction, stroke, heart failure, or cardiovascular-related death.

Abbreviations: **MR**: Mendelian randomisation; **SBP**: systolic blood pressure; **BP**: blood pressure; **LDL-C**: low-density lipoprotein cholesterol; **IHD**: ischaemic heart disease; **CVD**: cardiovascular disease; **SNP**: single nucleotide polymorphism; **NR**: not reported; **MVP**: Million Veteran Program; **ICBP**: International Consortium for Blood Pressure; **GLGC**: Global Lipids Genetics Consortium; **UKB**: UK Biobank; **CARDIoGRAM**: Coronary ARtery DIsease Genome wide Replication and Meta-analysis; **CC4D**: CARDIoGRAM plus Coronary Artery Disease Genetics consortium; **MEGASTROKE**: International Stroke Genetics Consortium meta-analysis; **CKB**: China Kadoorie Biobank; **KoGES**: Korean Genome and Epidemiology Study; **KCPS-II**: Korean Cancer Prevention Study-II.

2.3.2. Associations between SBP and IHD

This section summarises the available evidence on the associations of SBP with risk of IHD. The findings are presented sequentially by study design, beginning with observational studies, followed by genetic studies and randomised trials.

Observational studies: The meta-analyses of the observational associations between SBP and IHD are presented in **Table 2.6**. Four large prospective studies were included, including three IPD meta-analyses (PSC,[52] APCSC,[53] and EPOCH-Japan[54]) and a biobank study (CKB[55]). PSC reported studies conducted in Europe and Asia separately and, as such, there was one RR for the European population and four RRs for the Asian population in total. All studies reported significant log-linear associations between higher levels of SBP and higher risks of IHD.[52-55] The estimate with the highest number of events (number of fatal IHD = 28,797) was PSC Europe,[52] which reported that a 10 mmHg lower usual SBP was associated with a 25% lower risk of fatal IHD (HR: 0.75, 95% CI: 0.75-0.76). The estimates were comparable between studies conducted in Asia, including PSC Asia,[52] APCSC,[53] and CKB,[55] with an estimated 24-26% lower risk of incident IHD or IHD mortality per 10 mmHg lower SBP. EPOCH-Japan[54] reported a more modest effect size (HR: 0.86, 95% CI: 0.83-0.89) compared to the estimates from other studies, but the study did not correct for regression dilution. The regression dilution ratio in CKB was 0.6 and if EPOCH-Japan had corrected for regression dilution bias with a comparable regression dilution ratio as in the CKB analysis, the RR per 10 mmHg lower SBP would have been 0.77 (95% CI: 0.73-0.82), which is concordant with the estimates from other meta-analyses.

Genetic studies: Four MR studies exploring the association between SBP and the risk of IHD are shown in ***Table 2.7***. The study by Cho et al.[56] constructed polygenic

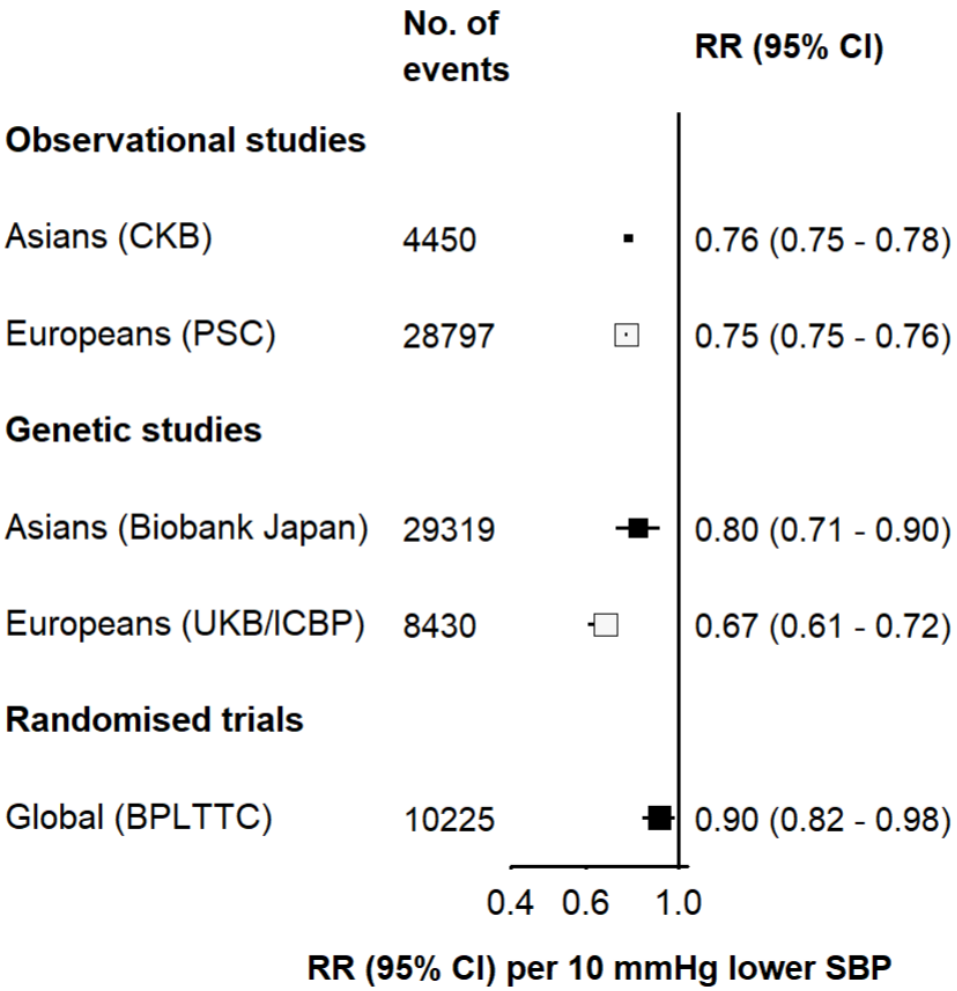
risk scores (PRS) using MVP data and estimated the gene-outcome association in the UKB. The study excluded participants with a history of CVD but included participants with antihypertensive medication use. The study also did not limit participants to White British ancestry, but 94.2% of the included participants were White British. Only composite CVD outcomes were reported, and it was estimated that a 10 mmHg lower PRS-predicted SBP was associated with a HR of 0.77 (95% CI: 0.72-0.86) for CVD. The analyses by Malik et al.[82] among White British participants without CVD history and antihypertensive drug use demonstrated that 10 mmHg lower genetically-predicted SBP was associated with a 33% lower risk of IHD (HR: 0.67, 95% CI: 0.61-0.72). For the MR studies conducted in populations of Asian ancestry, analyses in CKB demonstrated that 10 mmHg lower genetically-predicted SBP was associated with a 22% lower risk of IHD (HR: 0.78, 95% CI: 0.70-0.85).[57] Estimates using Biobank Japan showed that a 10 mmHg lower SBP was associated with a 20% lower odds of IHD (OR: 0.80, 95% CI: 0.71-0.90).[58]

Randomised trials: Two meta-analyses of RCTs of BP-lowering medications were summarised in ***Table 2.8***. IPD analyses from BPLTTC,[59] which included 38 RCTs with first incident IHD as an outcome, estimated that a 10 mmHg reduction in SBP was associated with a 10% reduction in IHD risk (HR: 0.90, 95% CI: 0.82-0.98). The meta-analysis conducted by Law et al. in 2009[36] involving 26 primary prevention trials showed that a 10 mmHg reduction in SBP was associated with a 21% reduction in the risk of IHD (RR: 0.79, 95% CI: 0.72-0.86).

The RRs from different sources of evidence for the association between SBP and IHD are presented in ***Figure 2.3***. Overall, for studies conducted in predominantly European populations, observational results showed that a 10 mmHg lower SBP was associated with a 25% (RR: 0.75, 95% CI: 0.75 – 0.76) lower risk of MI, compared

with a 33% lower risk (RR: 0.67, 95% CI: 0.61 – 0.72) in the most comparable MR study. The observational results from Asian populations were more comparable to those in European populations, but the MR results from Asia showed more moderate effect sizes (22% lower risk in CKB and 20% lower odds in Biobank Japan). The most recent randomised evidence (BPLTTC)[59] showed that SBP reduction by BP-lowering medications was associated with a less extreme risk reduction of MI (10% per 10 mmHg lower SBP) compared to observational and genetic studies.

Figure 2.3. Risk of IHD per 10 mmHg lower SBP in East Asian and European populations



Abbreviations. **CKB**: China Kadoorie Biobank; **PSC**: Prospective Studies Collaboration; **BPLTTC**: Blood Pressure Lowering Treatment Trialists' Collaboration; **ICBP**: International Consortium for Blood Pressure; **UKB**: UK Biobank; **RR**: Rate Ratio; **CI**: Confidence Interval; **SBP**: Systolic Blood Pressure; **IHD**: Ischaemic Heart Disease.

Table 2.6. Observational studies of the associations of SBP with risk of IHD, by ancestry

Data Source	Method	Outcome	N Studies	N Participants	N Events	Median Follow-up yrs	Mean age yrs	Female %	Reported SBP unit mmHg*	Reported RR (95% CI)	RR per 10 mmHg lower SBP (95% CI)	Degree of adjustment†
Predominantly European Population												
PSC Europe[52]	IPD	Fatal IHD	33	676,213	28,797	12.0	55	NR	-20✓	0.57 (0.56-0.58)	0.75 (0.75-0.76)	+
Predominantly Asian Population												
PSC Asia[52]	IPD	Fatal IHD	10	88,724	334	8.0	56	NR	-20✓	0.55 (0.48-0.63)	0.74 (0.69-0.79)	+
EPOCH-Japan[54]	IPD	Fatal IHD	11	73,916	770	15.0	58	58.9	+20	1.36 (1.27-1.45)	0.86 (0.83-0.89) ‡	++
APCSC[53]	IPD	Fatal IHD	37	425,325	2,149	7.1	47	43.0	-10✓	0.74 (0.71-0.76)	0.74 (0.71-0.76)	+
CKB[55]	Biobank	IHD	1	489,125	4,450	9.0	51	59.1	+10✓	1.31 (1.28-1.34)	0.76 (0.75-0.78)	++

*If the analysis corrected for regression dilution bias, the reported SBP unit is denoted by ✓✓.

† Degree of adjustment was recorded and denoted as + for age- and sex-adjusted, ++ for plus lifestyle factors (smoking and/or alcohol use) and/or BMI.

‡The regression dilution ratio in CKB was 0.6 and if EPOCH-Japan corrected for regression dilution bias with the same regression dilution ratio as in the CKB analysis, the RR per 10 mmHg lower SBP became 1/sqrt(exp(ln(1.36)/0.6))=0.77, 95% CI: 0.73-0.82.

Abbreviations: **SBP**: systolic blood pressure; **IHD**: ischaemic heart disease; **IPD**: individual participant data; **PSC**: Prospective Studies Collaboration; **EPOCH-JAPAN**: Evidence for Cardiovascular Prevention from Observational Cohorts in Japan; **APCSC**: Asia Pacific Cohort Studies Collaboration; **CKB**: China Kadoorie Biobank; **RR**: relative risk; **CI**: confidence interval; **NR**: not reported.

Table 2.7. Genetic studies on the association of SBP with risk of IHD, by ancestry

Data Source	Method	Outcome	N SNPs	N Participants	N Events	Median Follow-up yrs	Mean age yrs	Female %	Reported SBP unit mmHg	Reported RR (95% CI)	RR per 10 mmHg lower SBP (95% CI)	Variance explained %
Predominantly European Population												
UKB/MVP[56]	1-sample	CVD*	272	331,078	15,293	11.1	57	54.0	+2.6	1.07 (1.04-1.09)	0.77 (0.72-0.86)	NR
UKB/ICBP[82]	2-sample	IHD	253	255,714	8,430	NR	56	57.3	+10	1.50 (1.38–1.63)	0.67 (0.61-0.72)	4.8
Predominantly Asian Population												
CKB[57]	1-sample	IHD	460	60,481	3,995	NR	55	59.4	+10	1.29 (1.18-1.42)	0.78 (0.70-0.85)	NR
Biobank Japan[58]	1-sample	IHD	61	212,453	29,319	NR	63	44.9	+19.69	1.56 (1.24-1.96)	0.80 (0.71-0.90)	NR

*CVD was defined as a composite of first incident myocardial infarction, stroke, heart failure, or cardiovascular-related death.[56]

Abbreviations: **SBP**: systolic blood pressure; **IHD**: ischaemic heart disease; **CVD**: cardiovascular disease; **SNP**: single nucleotide polymorphism; **UKB**: UK Biobank; **MVP**: Million Veteran Program; **ICBP**: International Consortium for Blood Pressure; **CKB**: China Kadoorie Biobank; **NR**: not reported.

Table 2.8. RCTs on the association of SBP with risk of IHD

Data Source	Method	Outcome	N Studies	N Participants	N Events	Median Follow-up yrs	Mean age yrs	Female, %	Reported SBP unit mmHg	Reported RR (95% CI)	RR per 10 mmHg SBP reduction (95% CI)
Global Population											
Law 2009[36]	M-A	IHD	26	108,297	3,429	4.5	62	NR	-10	0.79 (0.72-0.86)	0.79 (0.72-0.86)
BPLTTC[59]	IPD	IHD	38	186,128	10,225	4.3	65	48.7	-5	0.95 (0.91-0.99)	0.90 (0.82-0.98)

Abbreviations: **RCT**: randomised controlled trial; **SBP**: systolic blood pressure; **IHD**: ischaemic heart disease; **M-A**: meta-analysis; **IPD**: individual participant data; **BPLTTC**: Blood Pressure Lowering Treatment Trialists' Collaboration; **RR**: relative risk; **CI**: confidence interval; **NR**: not reported.

2.3.3. Associations between LDL-C and IHD

Observational studies: The observational evidence for the associations between plasma cholesterol and the risk of IHD is presented in ***Table 2.9***. Five studies were included, including four IPD meta-analyses (two in Europe[9,60] and two in Asia[54,72]) and one biobank study (CKB[40]). PSC[60] (90% of the participants from western countries) reported that 1 mmol/L lower usual TChol was associated with 34% lower fatal IHD (HR: 0.66, 95% CI: 0.65-0.68) among participants aged 50 to 69 years old at baseline. ERFC[9] combined results from 58 prospective cohort studies, mostly conducted in Europe and North America, and found that the HR for IHD per 1 mmol/L lower non-HDL cholesterol was 0.65 (95% CI: 0.60-0.70). For studies conducted in Asia, APCSC[72] and CKB[40] reported similar estimates to each other. APCSC estimated that each 1 mmol/L lower usual TChol levels was associated with a 30% lower risk of IHD (HR: 0.70, 95% CI: 0.67-0.74).[72] The nested case-control study from CKB observed that the OR for IHD per 1 mmol/L lower usual levels of LDL-C was 0.73 (95% CI: 0.64-0.84). EPOCH-Japan reported a weaker association, observing that the HR for CHD was 0.80 (0.75-0.86) per 1 mmol/L lower measured TChol but the results were not adjusted for regression dilution bias, and, if the estimates were corrected by the same regression dilution ratio as in ERFC, the HR would be 0.69.

Genetic studies*:* Overall, in European populations, 1 mmol/L genetically-predicted lower plasma levels of LDL-C was associated with a lower risk of IHD, with a 31-44% lower risk per 1 mmol/L lower LDL-C. The results of MR studies on the associations of plasma levels of LDL-C with risks of IHD are presented in ***Table 2.10***. White et al.[90] used data from CARDIoGRAM plus Coronary Artery Disease Genetics consortium (CC4D) and GLGC and estimated an OR of 0.60 (95% CI: 0.53-0.66) for

the associations of LDL-C with IHD. Richardson et al. conducted a GWAS in UKB for LDL-C and selected 220 SNPs associated with LDL-C levels. They applied 209 SNPs that are available in CC4D database and found that 1 mmol/L lower genetically-predicted LDL-C was associated with an OR of 0.56 (95% CI: 0.49-0.63) for IHD. Allara et al.[85] derived genetic associations with LDL-C from GLGC and used 184 variants that explain 14.6% of the variation in LDL-C in the UKB to explore the gene-outcome association. They reported that 1 mmol/L lower genetically-predicted LDL-C was associated with an OR of 0.69 (95% CI: 0.64-0.74) for IHD. Holmes et al.[40] used 42 SNPs to construct an allele score for LDL-C that explained 4.3% of its variation. They estimated that a 1 mmol/L lower genetically-predicted level of LDL-C was associated with an OR for IHD of 0.56 (95% CI: 0.50-0.63). A study involving data from the Copenhagen General Population Study (CGPS) and Copenhagen City Heart Study (CCHS) that constructed a weighted genetic score using 11 SNPs reported that the RR for IHD per 1 mmol/L lower LDL-C was 0.69 (95% CI: 0.52-0.93).[96]

Two genetic studies of plasma levels of LDL-C and risk of IHD were conducted in Asian populations.[58,88] A study using gene-lipid association from Korean Genome and Epidemiology Study (KoGES) and gene-IHD association from Korean Cancer Prevention Study-II (KCPS-II) selected 20 SNPs for LDL-C and found that per 1 mmol/L lower genetically-predicted LDL-C was associated with an OR of 0.61 (95% CI: 0.48-0.76) for IHD. The results from Biobank Japan[58] using 29 SNPs reported that a 1 mmol/L lower genetically-predicted level of LDL-C was associated with an OR of 0.56 (95% CI: 0.45-0.70) for IHD. The MR estimates in Asian populations were largely consistent with the estimates obtained in European populations, with

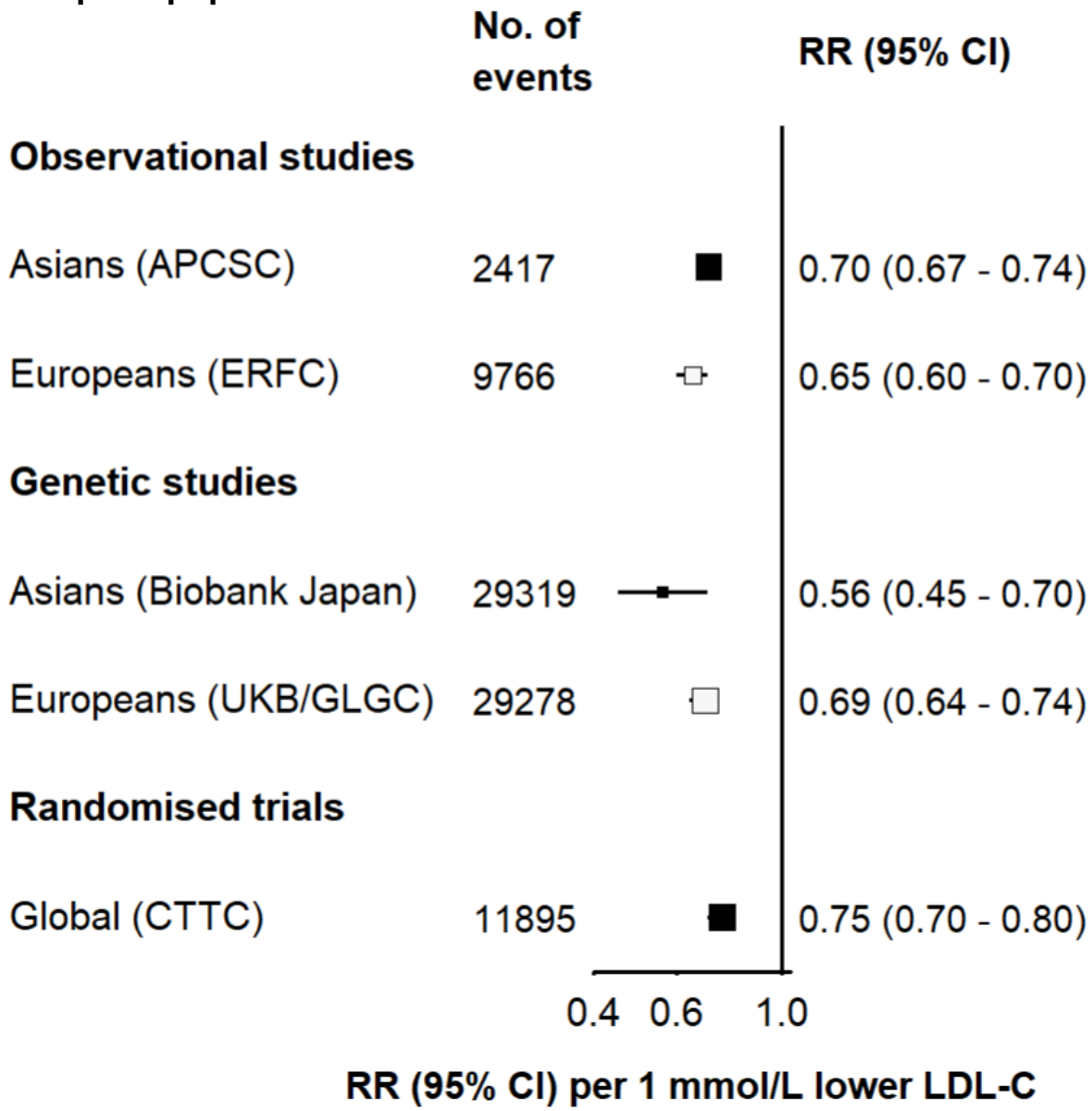
risk reductions of IHD from 39-44% per 1 mmol/L lower genetically-predicted LDL-C levels.

Randomised trials**:* Meta-analyses of RCTs on the effects of LDL-C lowering and the risk of IHD are presented in ***Table 2.11. CTTC,[42] which aggregated IPD from 27 RCTs, reported that among participants without CVD history, a 1 mmol/L lower LDL-C achieved by statin therapy was associated with a 25% lower risk of major vascular events (RR: 0.75, 95% CI: 0.70-0.80). Wang et al. performed a meta-analysis and reported that 1 mmol/L reduction in LDL-C, achieved by statin therapy was associated with a 21% reduction in 10-year risk of MVE (RR: 0.79, 95% CI: 0.58-0.97) in primary prevention trials.[94] The meta-analyses by Briel et al. estimated that 1 mmol/L reduction in LDL-C by drug treatments or dietary interventions was associated with a RR of 0.75 (95% CI: 0.67-0.84).[91] The analyses, however, did not distinguish between primary and secondary prevention trials. Overall, the evidence from randomised trials consistently demonstrated that a 1 mmol/L lower LDL-C was associated with a ~25% reduction in risk of IHD.

Overall, for studies conducted in predominantly European populations, observational results showed ~35% reduction in IHD risk per 1 mmol/L lower LDL-C, compared with a risk reduction of 31-44% per 1 mmol/L genetically-predicted lower LDL-C. The observational results from Asian populations showed a more moderate effect size compared to European populations (~30% lower risk of IHD per 1 mmol/L lower plasma levels of LDL-C) but the MR results were comparable with the studies conducted in European populations (39%-44%). The most reliable randomised evidence on the effects of LDL-lowering (IPD meta-analysis, CTTC) reported that a 1 mmol/L lower level of LDL-C achieved by statin medication was associated with a 25% lower risk of IHD. The effect sizes from observational, genetic, and randomised

evidence for the association between LDL-C and IHD risks are presented in ***Figure 2.4***.

Figure 2.4. Risk of IHD per 1 mmol/L lower LDL-C in Asian and European populations



Abbreviations. **APCSC**: Asia Pacific Cohort Studies Collaboration; **ERFC**: Emerging Risk Factors Collaboration; **CTTC**: Cholesterol Treatment Trialists' Collaboration; **GLGC**: Global Lipids Genetics Consortium; **UKB**: UK Biobank; **RR**: Rate Ratio; **CI**: Confidence Interval; **LDL-C**: Low-Density Lipoprotein Cholesterol; **IHD**: Ischaemic Heart Disease.

Table 2.9. Observational evidence on the association between LDL-C and the risk of IHD

Data Source	Method	Outcome	N Studies	N Participants	N Events	Median Follow-up yrs	Mean age yrs	Female %	Reported C unit mmol/L*	Reported RR (95% CI)	RR per 1 mmol/L lower LDL-C (95% CI)	Degree of adjustment†
Predominantly European Population												
ERFC[9]	IPD	IHD	58	255,989	9766	9.6	58	43.0	+1 non-HDL✓✓	1.53 (1.42-1.66)	0.65 (0.60-0.70)	+++
PSC[60]	IPD	Fatal IHD	61	892,337	15,980	13	50-69	NR	-1 total✓✓	0.66 (0.65-0.68)	0.66 (0.65-0.68)	+
Predominantly Asian Population												
EPOCH-Japan[54]	IPD	Fatal IHD	11	73,916	770	15	58	58.9	+1 total	1.25 (1.16-1.34)	0.80 (0.75-0.86) ‡	++
CKB[78]	Biobank	IHD	1	4,661	912	NR	46	49.8	+0.7 LDL	1.25 (1.13-1.37)	0.73 (0.64-0.84)	++
APCSC[72]	IPD	IHD	30	312,417	2417	4	47	51.6	+1 total✓✓	1.42 (1.35-1.50)	0.70 (0.67-0.74)	+

If the analysis corrected for regression dilution bias, the reported cholesterol unit was denoted by ✓✓

†Degree of adjustment was recorded and denoted as + for age- and sex-adjusted, ++ for plus lifestyle factors (smoking and/or alcohol use) and/or BMI, +++ for plus other vascular risk factors, such as systolic blood pressure and diabetes.

‡The regression dilution ratio in ERFC was 0.6 for non-HDL-C and were EPOCH-Japan corrected for regression dilution bias with the same regression dilution ratio as in the ERFC analysis, the RR per 1 mmol/L lower LDL-C became 1/exp(ln(1.36)/0.6)=0.69, 95% CI: 0.61-0.78.

Abbreviations: **LDL-C**: low-density lipoprotein cholesterol; **IHD**: ischaemic heart disease; **ERFC**: Emerging Risk Factors Collaboration; **PSC**: Prospective Studies Collaboration; **EPOCH-JAPAN**: Evidence for Cardiovascular Prevention from Observational Cohorts in Japan; **CKB**: China Kadoorie Biobank; **APCSC**: Asia Pacific Cohort Studies Collaboration; **IPD**: individual participant data; **RR**: relative risk; **CI**: confidence interval; **NR**: not reported.

Table 2.10. Genetic evidence on the association between LDL-C and the risk of IHD

Data Source	MR Method	Outcome	N SNPs	N Participants	N Events	Median Follow-up yrs	Mean age yrs	Female %	Reported C unit mmol/L	Reported RR (95% CI)	RR per 1 mmol/L lower LDL-C (95% CI)	Variance explained %
Predominantly European Population												
CGPS+CCHS[96]	1-sample	IHD	11	95,490	4,199	NR	58	55.0	+1 LDL	1.45 (1.08-1.93)	0.69 (0.52-0.93)	NR
CplusC4D/UKB[41]	2-sample	IHD	209	184,305	60,810	NR	NR	NR	+0.87 LDL	1.66 (1.49-1.86)	0.56 (0.49-0.63)	NR
Holmes 2015[40]	1-sample	IHD	42	62,199	11,826	NR	59	55.8	+1 LDL	1.78 (1.58-2.01)	0.56 (0.50-0.63)	4.3
CplusC4D/GLGC[90]	2-sample	IHD	130	194,427	63,746	NR	NR	NR	+1 LDL	1.68 (1.51-1.87)	0.60 (0.53-0.66)	7.9
UKB/GLGC[85]	2-sample	IHD	184	367,703	29,278	NR	57	54.1	+1 LDL	1.45 (1.35-1.57)	0.69 (0.64-0.74)	14.6
Predominantly Asian Population												
KCPS-II/KoGES[88]	2-sample	IHD	20	156,701	2932	8	42	39.5	+1mg/dL LDL	1.01 (1.01-1.01)	0.61 (0.48-0.76)	NR
Biobank Japan[58]	1-sample	IHD	29	212,453	29,319	NR	64	42.8	+0.87 LDL	1.65 (1.36, 2.00)	0.56 (0.45-0.70)	NR

Abbreviations: **MR**: Mendelian randomisation; **LDL-C**: low-density lipoprotein cholesterol; **IHD**: ischaemic heart disease; **SNP**: single nucleotide polymorphism; **RR**: relative risk; **CI**: confidence interval; **NR**: not reported; **CGPS**: Copenhagen General Population Study; **CCHS**: Copenhagen City Heart Study; **CplusC4D**: Coronary Artery Disease Genetics Consortium; **UKB**: UK Biobank; **GLGC**: Global Lipids Genetics Consortium; **KCPS-II**: Korean Cancer Prevention Study-II; **KoGES**: Korean Genome and Epidemiology Study.

Table 2.11. RCT evidence on the association between LDL-C and the risk of IHD

Data Source	Method	Outcome	N Studies	N Participants	N Events	Median Follow-up yrs	Mean age yrs	Female, %	Reported C unit mmol/L	Reported RR (95% CI)	RR per 1 mmol/L LDL-C reduction (95% CI)
Global Population											
N Wang 2020[94]	M-A	MVE*	11	81,786	NR	3.7	NR	NR	-1 LDL	0.79 (0.58-0.97)	0.79 (0.58-0.97)
Briel 2009[91]	M-A	IHD	95	299,310	NR	2.8	NR	NR	-0.26 LDL	0.93 (0.90-0.96)	0.75 (0.67-0.84)
CTTC 2012[42]	IPD	IHD	27	174,149	11,895	4.8	63	NR	-1 LDL	0.75 (0.70-0.80)	0.75 (0.70-0.80)

*MVE: Major vascular events, composite of cardiovascular mortality, MI, coronary revascularization, or stroke.

Abbreviations: **RCT**: randomised controlled trial; **LDL-C**: low-density lipoprotein cholesterol; **IHD**: ischaemic heart disease; **M-A**: meta-analysis; **IPD**: individual participant data; **CTTC**: Cholesterol Treatment Trialists' Collaboration; **MVE**: major vascular events; **MI**: myocardial infarction; **RR**: relative risk; **CI**: confidence interval; **NR**: not reported.

2.3.4. Combined effects of BP and LDL-C

Among the included studies, two meta-analyses of observational studies[54,79] and one MR study[74] examined the combined associations of BP and LDL-C with risk of IHD. An IPD meta-analysis[23] of three major polypill trials (TIPS-3, HOPE-3, and PolyIran) showed that after 2 years of follow-up, a fixed-dose combination of at least two BP-lowering agents plus a statin, lowered cholesterol by 0.58 mmol/L (95% CI: 0.55 – 0.62 mmol/L) and SBP by 4.7 mmHg (95% CI: 4.2 – 5.2 mmHg) compared to the control group. The HR for CVD was 0.62 (95% CI: 0.53-0.73) and, assuming a log-linear independent relationship between both SBP and LDL-C and CVD, the HR for CVD per 1 mmol/L lower LDL-C and per 10 mmHg lower SBP would be 0.22 (95% CI: 0.13-0.37). This is consistent with an estimate by Ference et al.[74] that reported that the combined exposure to 1 mmol/L lower genetically-predicted LDL-C and 10 mmHg lower genetically-predicted SBP was associated with an OR of 0.22 (95% CI: 0.17-0.26) of CVD (***Table 2.12***).

The two observational meta-analyses were both conducted in Asian populations. APCSC[79] observed that the effect sizes of the associations between usual SBP and IHD were greater among lower cholesterol groups. They reported that a 10 mmHg lower usual SBP was associated with 25% (HR=0.75, 95% CI: 0.73-0.77), 22% (HR=0.78, 95% CI: 0.74-0.83), 20% (HR=0.80, 95% CI: 0.76-0.85), and 17% (HR=0.83, 95% CI: 0.79-0.88) lower risk of IHD in <4.75, 4.75 to 5.49, 5.50 to 6.24, and ≥6.25 mmol/L cholesterol groups, respectively.

In contrast with APCSC, the results from EPOCH-Japan[54] reported the opposite result for IHD, in that the HR for IHD was higher among participants who had higher levels of TChol. The authors reported that a 10 mmHg lower usual SBP was associated with 10% (HR=0.90, 95% CI: 0.84-0.96), 11% (HR=0.89, 95% CI: 0.83-

0.96), 14% (HR=0.86, 95% CI: 0.80-0.92), and 19% (HR=0.81, 95% CI: 0.76-0.86) lower risk of IHD in <4.7, 4.7 to 5.1, 5.2 to 5.6, and ≥5.7 mmol/L cholesterol groups, respectively. Importantly, EPOCH-Japan did not correct for regression dilution bias in their analyses. However, if EPOCH-Japan had corrected their results for regression dilution bias using the same regression dilution ratios as APCSC, the associations of SBP with risk of IHD within each TChol stratum would have been comparable. Details of the effect sizes are provided in ***Table 2.12***.

MR analyses conducted by Ference et al.[74] utilised data from Liu et al. for the association between genetic variants to construct GRSs. A factorial analysis was conducted and found no evidence for an interaction between the associations of genetic scores for LDL-C and SBP with risk of major coronary events (a composite of coronary death, nonfatal myocardial infarction, or coronary revascularization). The authors concluded that genetically-predicted LDL-C and SBP levels were independently associated with CVD risks.

Table 2.12. Combined effects of cholesterol and SBP on risk of IHD

Lipid levels, mmol/L	RR per 10 mmHg lower SBP (95% CI)
Observational evidence	
APCSC	
<4.75	0.75 (0.73-0.77)
4.75-5.49	0.78 (0.74-0.83)
5.50-6.24	0.80 (0.76-0.85)
≥6.25	0.83 (0.79-0.88)
	$P_{Heterogeneity} < 0.05$
EPOCH-Japan*	
<4.7	0.90 (0.84-0.96)
4.7-5.1	0.89 (0.83-0.96)
5.2-5.6	0.86 (0.80-0.92)
≥5.7	0.81 (0.76-0.86)
	$P_{Heterogeneity} = 0.10$
Genetic evidence	
LDL-C genetic score quartile	RR per genetically-predicted 10 mmHg lower SBP (95% CI) for major coronary events
1	0.54 (0.48-0.61)
2	0.53 (0.46-0.60)
3	0.58 (0.50-0.66)
4	0.57 (0.48-0.66)
	$P_{Heterogeneity} = 0.77$

** EPOCH Japan did not correct for regression dilution and the outcomes were fatal IHD and stroke.*
† Major coronary events refer to a composite of coronary death, nonfatal myocardial infarction, or coronary revascularization. RRs were ORs.

2.4. Discussion

This chapter conducted an umbrella review of observational, randomised, and genetic evidence on the associations of SBP and LDL-C with risk of IHD in European and East Asian populations. The review compared effect sizes across different study designs and ancestry populations, evaluated evidence for the combined effects of SBP and LDL-C on risk of IHD, and identified key gaps in the literature.

Across different study designs, there was general consistency in the direction of the associations for both SBP and LDL-C with IHD risk, but the magnitudes of the associations varied. For a 10 mmHg lower SBP, observational studies in European populations reported a 25% lower risk of IHD,[52] which was broadly comparable to a 33% lower risk in the most comparable MR study.[82] The moderately larger effect size in MR studies is expected, as genetic variants represent a lifelong exposure to lower SBP, whereas observational studies typically measure BP later in life (middle age) and are susceptible to underestimation from random measurement error and reverse causality.[52] In East Asian populations, observational estimates were similar to those in Europeans,[52] but the MR estimates were more moderate (20-22% lower risk).[52,57,58] This could reflect differences in the genetic architecture of BP or lower statistical power in the underlying GWAS for East Asians. By contrast, RCTs of BP-lowering medications showed a smaller risk reduction (10%) compared to estimates from observational and genetic studies,[59] which is also anticipated given their shorter follow-up duration (~ 5 years).

For LDL-C, a similar pattern was observed. In European populations, the ~35% lower IHD risk per 1 mmol/L lower LDL-C in observational studies[9] was consistent with the 31-44% lower risk from genetic studies.[41,96] In East Asian populations, however, a notable discrepancy was observed: observational studies reported a

more moderate risk reduction (~30%)[78] compared to Europeans, whereas MR studies showed a comparable, if not stronger, effect (39-44%).[58,88] A plausible hypothesis is that the attenuated observational associations in East Asian populations is due to confounding from lifestyle or dietary factors and reverse causality, particularly given that baseline LDL-C levels are, on average, lower in East Asian populations. The MR evidence, being less susceptible to these biases, likely reflects a more accurate causal association of LDL-C levels with IHD across different ancestry populations. As with SBP, the effect sizes from RCTs of cholesterol-lowering medication were more modest (25% lower risk),[42] reflecting the shorter-term follow-up from RCTs compared with observational and genetic studies.

Regarding the combined effects of SBP and LDL-C, this review identified conflicting observational evidence from East Asian populations. The APCSC reported that the RRs for the association between SBP and IHD were attenuated at higher cholesterol levels,[79] whereas the EPOCH-Japan study observed a strengthening of these associations.[54] In contrast, factorial MR analyses in European populations have consistently concluded that genetically-predicted levels of LDL-C and SBP are independently associated with CVD risk.[74] The conflicting observational evidence may reflect the heterogeneity of the underlying studies, including differences in data collection procedures, assay methods, and definitions of exposures and outcomes. Moreover, the discrepancy between observational and genetic studies could indicate the presence of residual confounding and reverse causality in observational analyses.

As with all umbrella reviews, the findings of this chapter are dependent on the quality of the underlying systematic reviews and primary studies.[68,69] The umbrella review approach may also omit relevant individual studies that have not yet been

incorporated into existing systematic reviews. Moreover, this review was limited to English-language publications, which may have introduced a language bias and excluded potentially relevant evidence published in other languages. In addition, due to the scope of this chapter, no formal quality appraisal of the included studies was conducted, as the primary aim was to identify research gaps and provide a high-level overview of the existing evidence.

Analysis of the results of this umbrella review identified several critical research gaps that the present thesis aims to address.

Firstly, many of the large meta-analyses of observational studies were conducted decades ago. The IPD data from observational studies such as PSC[52,60] and APCSC[55,72,79] were collected from cohorts initiated more than 20–30 years ago. The implication of this is that their risk estimates may not accurately reflect contemporary populations, which have experienced significant shifts in lifestyle, diet, BP, and greater use of lipid-modifying medications in recent decades.[97] The establishment of large-scale, population-based biobank studies, namely UKB[98] and CKB,[99] provides an opportunity to conduct updated analyses. These contemporary cohorts have included very large numbers of participants (approximately half a million each) and have used uniform, standardised data collection and follow-up procedures, ensuring high-quality data with less heterogeneity than is typical in meta-analyses of individual studies. With a wide range of baseline and follow-up variables, these biobanks now enable more detailed analyses, including the investigation of non-linear dose-response associations and the comparison of different lipid fractions (e.g., ApoB, non-HDL-C) to inform clinical guidelines. Furthermore, CKB and UKB both have reliable, direct measurements of LDL-C and apolipoprotein traits, overcoming a

limitation of previous prospective studies that often relied on calculated LDL-C or total cholesterol.

Secondly, there has been a significant lack of MR studies conducted in East Asian populations. The majority of large-scale GWAS and subsequent MR analyses have been conducted in populations of European ancestry, with only two genetic studies of LDL-C and IHD identified in Asian populations.[57,88] The availability of genetic data in CKB should enable the investigation of the causal relevance of genetically-predicted plasma lipid traits and SBP on MI in a large Chinese cohort.[100] As MR studies are less prone to confounding and reverse causality, such analyses can facilitate the triangulation of evidence from observational and randomised studies and provide a more comprehensive assessment of causality, focusing on trans-ethnic comparisons.

Most importantly, few studies have explored the combined effects of plasma lipid traits and SBP on the risk of IHD,[54,79,80] and the limited available evidence is conflicting. The analyses conducted in UKB and CKB, with their large sample sizes, high-quality data, and available genetic information, can clarify the nature of these combined effects using multiple sources of evidence, i.e., observational and factorial MR analyses. Establishing whether these major risk factors act independently is important to inform strategies for primary prevention of MI in diverse ancestry populations.

Chapter 3 Overview of Data Sources and Study Procedures

3.1. Background

Biobanks are large population-based studies that link medical history, physical measures, and biological samples with long-term follow-up health records.[98] By combining detailed measurements including questionnaires, biomarkers, imaging, with genetic data, they enable robust epidemiological studies. Flagship examples include the UK Biobank (UKB) of ~ 500k participants with extensive phenotyping, open access, and centralised quality control in the UK,[98,101] and the China Kadoorie Biobank (CKB) of ~ 500k participants with systematic linkage to mortality/morbidity registries and health-insurance databases in China.[99]

This chapter provides a detailed description of the two large-scale prospective cohorts that were analysed in this thesis, namely, CKB[99] and UKB.[98] It outlines the study design, setting, and populations for each cohort, followed by a description of the procedures for survey data collection, outcome ascertainment, and genetic data assays. The chapter also includes presentation of the baseline characteristics of the participants in each cohort and their relevant subsets, and a comparative summary of the two study populations, with a discussion of key methodological considerations. This chapter focuses on the description of the data sources and study procedures, but the specific statistical methods used for the analyses are provided in the subsequent individual chapters.

3.2. China Kadoorie Biobank (CKB) design and procedures

3.2.1. CKB study design

CKB is a large prospective cohort study that recruited 512,891 middle-aged and older adults in China.[99] Launched in 2004 by the Clinical Trial Service Unit and Epidemiological Studies Unit at the University of Oxford and the China National Centre for Disease Control (CDC), the study's initial objectives were to investigate the causes of common major chronic diseases in China. CKB collected data on both established and novel risk factors, including physical measurements, blood biochemistry, genetic markers, and a wide array of disease outcomes. All participants were followed up for incident disease outcomes. Ethical approval was obtained from the University of Oxford, the China National CDC, and institutional research boards at the local CDCs.

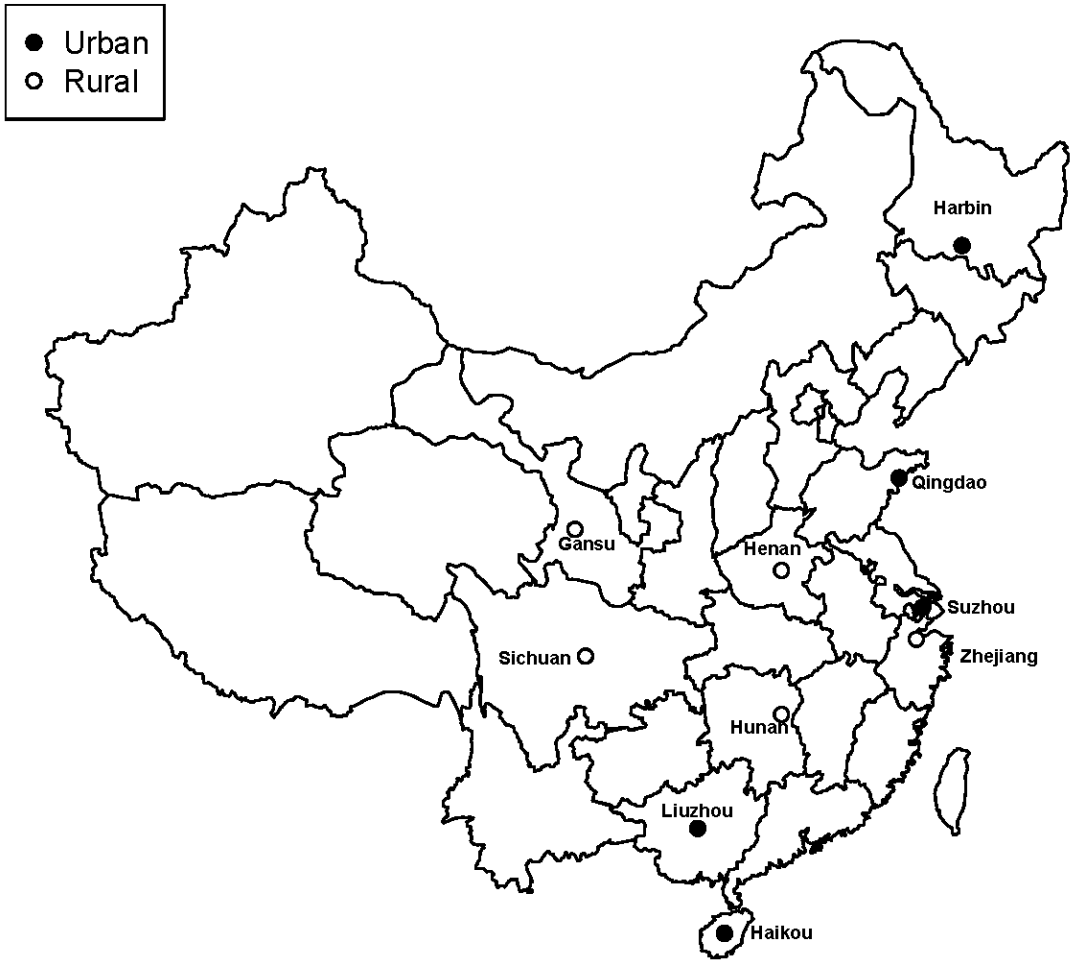
From June 2004 to July 2008, participants were enrolled from 10 diverse regions (five urban and five rural) in China. Eligible individuals were permanent residents aged 35 to 74 years. They were identified using official residential records from each of the 100 to 150 administrative units (rural villages or urban residential committees) within the study regions. After a comprehensive publicity campaign to raise local awareness, invitations and information leaflets were hand-delivered to potential participants. Eligibility was confirmed using unique national identity cards.

To increase study participation, some individuals slightly outside the target age ranges were also included. The recruitment process yielded a response rate of about 30% (ranging from 26–38% in rural regions and 16–50% in urban regions). The main reasons for non-participation were unwillingness to spend time visiting the screening

centre and not being at home. Overall, a total of 512,891 adults were recruited into the study.[100] The analyses used in this thesis are from data release R1801.

The 10 study regions were selected to maximise geographic and socioeconomic diversity, rather than to be representative of the overall Chinese population. Additional selection criteria included the quality of mortality and disease registries, local commitment and capacity, and population stability. Importantly, local disease patterns and exposures to specific risk factors were considered to enhance the study's ability to investigate the determinants of major chronic diseases in China.[99,102] A description of the geographical locations, including rural and urban regions, is shown in ***Figure 3.1***.

Figure 3.1. Map of China showing locations of the 10 CKB study sites and number of participants recruited by each site



Map generated by the author based on internally shared codes from China Kadoorie Biobank

Following approximately 10 years of follow-up, with data censored on 1st January 2014, a nested case-control study was constructed. Around 1,000 MI cases and 6,000 controls were included in the present nested case-control study of MI.[73] The nested case-control subset included measurements of plasma lipid levels. [40,73] In this subset, 17 biochemistry measurements were available for ~1,200 ICH, ~1,200 IS, and ~1,000 MI cases and ~1,200 healthy controls. In addition, a total of 13,000 participants were randomly selected from the remainder of the nested case-control study for biochemistry measurements, yielding a total of 18,253 participants with biochemistry measurements.

3.2.2. CKB survey data collection

At the CKB baseline survey, participants provided written informed consent, which included permission for long-term storage of blood samples and access to their medical records.[99] They completed a computer-assisted, interviewer-administered questionnaire covering demographic information, socioeconomic status, lifestyle factors, and medical history. Trained staff performed standardised physical measurements, including height, weight, waist and hip circumferences, and blood pressure. A 10 ml non-fasting blood sample was collected from each participant, with plasma aliquots stored in liquid nitrogen for long-term analysis.

All participants were followed up for incident disease and mortality through electronic linkage to local and national disease and death registries, in addition to the national health insurance system, which has almost universal coverage. This comprehensive system enabled the accurate and complete ascertainment and coding of health outcomes by trained staff blinded to the baseline information.

Two resurveys were conducted on random subsets of approximately 5% of surviving participants to collect repeat measurements.[100,102] The first resurvey was conducted in 2008 (approximately 4 years after baseline) and the second was conducted between 2013-14 (approximately 8 years after baseline). These resurveys used identical procedures to the baseline assessment, collecting updated questionnaire data, repeat physical measurements, and new blood samples. The repeat measures were used to assess the reliability of baseline measurements and correct regression dilution bias in statistical analyses.[73,76,100]

Study participants from the CKB provided a 10 ml non-fasting blood sample in vacutainers containing EDTA (with time since last meal recorded). Plasma aliquots

were stored in liquid nitrogen. In the CKB biochemistry subset, plasma lipids and apolipoproteins were measured at the Wolfson Laboratory, Nuffield Department of Population Health, University of Oxford, UK. Plasma levels of LDL-C and HDL-C were measured using direct homogeneous assays, while other lipids and apolipoproteins were measured using immunoturbidimetric assays on Beckman Coulter AU680 clinical chemistry analysers.[73,99]

Full details of baseline variable definitions used are provided in ***Table 3.1.*** These include sociodemographic factors (age, sex, education), lifestyle factors (alcohol use, smoking, physical activity), physical and blood measurements (BMI, SBP, diastolic blood pressure), biochemistry (TChol, LDL-C, HDL-C, triglycerides, ApoA1, ApoB, haemoglobin A1c), and medical history and medication (diabetes, hypertension, use of antihypertensive and lipid-modifying drugs).

Table 3.1. Baseline variable definitions in CKB

Variable	Definition
Sociodemographic	
Age	Age at baseline
Sex	Sex of participant
Educational attainment	Self-reported educational attainment, from 0 “No formal school” to 5 “University”.
Lifestyle factors	
Alcohol Use	Self-reported alcohol use with 6 categories, including “weekly” drinkers.
Smoking Status	Self-reported smoking status was categorised into four categories, including “current smoker”.
Physical Activity	MET h/day, including total and work-related physical activity, MET in the analyses was total minus work-related.
Physical and blood measurement	
BMI	BMI values were constructed from height and weight at baseline.
Systolic Blood Pressure	Measured twice after participants had sat for at least 5 mins. A third reading of the participants’ SBP was taken if there was a difference between the first two readings of more than 10 mmHg. SBP in the analysis was mean of last two readings.
Diastolic Blood Pressure	Same as systolic blood pressure.
Biochemistry Measurements	
Total Cholesterol LDL Cholesterol HDL Cholesterol Triglycerides Apolipoprotein B Apolipoprotein A1	Lipids and apolipoproteins (except LDL-cholesterol and HDL-cholesterol) were assayed using Beckman-Coulter AU680 clinical chemistry analysers. LDL-cholesterol and HDL-cholesterol were assayed using N-geneous® reagents, calibrations and settings
HbA1c	Measured by HPLC analysis on a Bio-Rad VARIANT II Turbo
Medical History and Medication	
Diabetes	Self-reported cases and screen-detected cases (a blood glucose level ≥7.0 mmol/L and a fasting time >8 h, a blood glucose level ≥11.1 mmol/L and a fasting time <8 h, or a fasting blood glucose level ≥7.0 mmol/L).
Hypertension	Self-reported hypertension, or having SBP≥140 mmHg or DBP≥90 mmHg, or currently taking antihypertensive medications
Antihypertensive and lipid-modifying medications	Self-reported use of antihypertensive and lipid-modifying medications at baseline

Abbreviations: **UKB**: UK Biobank; **CKB**: China Kadoorie Biobank; **LDL**: Low-Density Lipoprotein; **HDL**: High-Density Lipoprotein; **MET**: Metabolic Equivalent of Task; **BMI**: Body Mass Index; **SBP**: Systolic Blood Pressure; **DBP**: Diastolic Blood Pressure; **CHO-POD**: Cholesterol Oxidase-Peroxidase; **HPLC**: High-Performance Liquid Chromatography; **HbA1c**: Haemoglobin A1c; **GP**: General Practitioner.

Baseline and incident CVD outcomes in CKB are defined in ***Table 3.2***, with prevalent cases including reported CHD, stroke, or TIA at baseline, and incident MI identified via electronic health record linkage using ICD-10 codes (I21.x, I22.x, I23.x) with diagnosis dates after recruitment.

Table 3.2. Definition of prevalent and incident CVD outcomes in CKB

Conditions	CKB
	Electronic healthcare record linkage
Prevalent MI Prevalent Stroke	Baseline prevalent CVD cases include reported coronary heart disease, stroke, or transient Ischaemic attack diagnosis
Incident MI	I21.x, I22.x, I23.x and diagnosis date is later than recruitment date

Abbreviations. **HES**: Hospital Episode Statistics; **MI**: Acute Myocardial Infarction; **UKB**: UK Biobank; **CKB**: China Kadoorie Biobank; **CVD**: Cardiovascular Disease

3.2.3. Genetic study of CKB

Approximately 100,000 participants were genotyped as part of a nested case-control study in CKB.[100] This genetic subset was constructed to be broadly representative of the overall cohort, while also being enriched for disease cases to increase statistical power.

It comprised approximately 20,000 incident cases of major diseases (including stroke and IHD) and a random selection of about 80,000 participants, which included 25,000 individuals who attended the second resurvey. Genotyping was performed using a custom-designed Affymetrix Axiom array panel including approximately 0.8 million SNPs, processed in batches of approximately 50 plates.[100] The quality control of the genetic data was performed by the CKB study team prior to data release. Samples with low call rates (<95%), high missingness (>5%), unusual heterozygosity, sex mismatches, or those identified as population outliers were excluded. Imputation was conducted using the 1000 Genomes Phase 3 East Asian reference panel.[100]

The principal components (PCs) represent axes of genetic variation derived from genome-wide genotype data and are used to capture population structure and ancestry differences within CKB, thereby minimising confounding due to genetic stratification.[103] In the CKB genetic study, principal component analysis (PCA) was performed among unrelated participants to summarise genome-wide allele frequency variation across the 10 geographically diverse recruitment regions. The national PCs reflected major north–south and east–west gradients of genetic variation within China, consistent with known patterns of historical migration. These national PCs were identified as informative for population structure and used as covariates in subsequent genetic analyses.[103]

3.3. UK Biobank (UKB) design and procedures

3.3.1. UKB study design

The UKB is a large-scale prospective cohort study designed to investigate the associations of a wide range of exposures with health outcomes in middle and older age.[98,101] Between 2006 and 2010, over 500,000 participants aged 40 to 70 years were recruited through National Health Service registers and invited to attend one of twenty assessment centres across the UK. The geographical locations of the assessment centres are shown in ***Figure 3.2.***

At these centres, participants completed computer-based questionnaires and underwent physical assessments, including measurements of blood pressure and anthropometry, conducted by trained nurses, with blood samples collected for long-term storage. The study received ethical approval from the NHS North West Multi-Centre Research Ethics Committee and all participants provided written informed consent to participate in the study. UKB provides a robust platform for exploring the determinants and consequences of chronic diseases.[98,101]

The UKB dataset used in this thesis is from Application Number 50474.

Figure 3.2. Locations of UK Biobank assessment centres throughout the United Kingdom



Map generated by the author based on publicly available information from UK Biobank.[98]

3.3.2. UKB survey data collection

During the baseline assessment, all participants provided written informed consent before attending one of 20 assessment centres across the UK. At the centre, each individual completed a self-administered touchscreen questionnaire that collected detailed information on demographics, socioeconomic status, lifestyle behaviours, dietary habits, and medical history. Following the questionnaire, a trained research nurse conducted a personal interview and performed standardised physical measurements including height, weight, waist and hip circumference, and blood pressure. In addition, approximately 45 ml of blood and a urine sample were collected from each participant. These biological samples were processed and stored using automated systems, with serum samples collected in serum separator tubes and plasma samples in EDTA tubes, ensuring that they were maintained under optimal conditions for long-term storage and subsequent biomarker analyses. In UKB, cholesterol was measured using CHO-POD analysis, while apolipoproteins and triglycerides were assessed using immunoturbidimetric assays on a Beckman Coulter AU5800 clinical chemistry analyser at the University of Manchester, UK.[98,103] Further details of the baseline characteristics relevant to this thesis are provided in ***Table 3.3***.

A subset of approximately 20,000 participants was invited to participate in a repeat survey (resurvey) conducted between 2012 and 2013 (a median of 4 years after baseline). This resurvey closely replicated the original baseline procedures, including the completion of the touchscreen questionnaire and the collection of physical measurements and biological samples. The resurvey was designed to update and validate baseline data in addition to recording any changes in health status over time. By employing the identical standardised protocols as the baseline assessment,

the study ensured consistency in data collection and allowed for the adjustment for time-dependent effects in subsequent analyses.[98,101,103]

Following the initial assessments, participants were monitored continuously by regular linkage with national electronic health records, including Hospital Episode Statistics, disease registries, and the Office for National Statistics Causes of Death data. This comprehensive follow-up system enables the accurate capture and coding of incident disease events using standardised diagnostic criteria. In addition, rigorous quality control measures and periodic reviews are implemented to verify data accuracy and consistency.[98,101,103] These follow-up procedures form the foundation for reliable analysis of the long-term associations between different exposures and disease outcomes in the UKB cohort. Details of the definition of CVD outcomes are provided in ***Table 3.4*****.**

Table 3.3. Baseline variable definitions in UKB

Variable	UKB Field	Definition
Sociodemographic		
Age	34; 53	Age when attended assessment centre
Sex	31	Sex of participant
Ethnic background	21000	Ethnicity from baseline questionnaire
Townsend deprivation index	189	Townsend deprivation index was calculated immediately prior to participant joining UK Biobank
Educational attainment	845 6138	Categories (Field ID: 6138) were mapped onto years of education using the ISCED scale. Those with missing data in categorical educational variables were replaced by age completed full-time education (field ID:845)
Lifestyle factors		
Alcohol Use	1558 20117	Self-reported alcohol use was categorised into 7 categories/ It was further categorized as weekly drinker if the participants were in the "Once or twice a week", "Three or four times a week", "Daily or almost daily" categories.
Smoking Status	1239 1249	Self-reported smoking status was categorised into four categories. "Occasional" and "Regular" smokers were further labelled as current smokers.
Physical Activity	402273	Sum of MET minutes per week for walking, moderate activity, and vigorous activity, converted to h/day.
Physical and blood measurement		
BMI	21001	BMI values were constructed from height and weight measured during the initial assessment centre visit.
Systolic Blood Pressure	4080 93	Measured twice after participants had sat for at least 5 mins. If automated reading for SBP was not available, manual reading values were used. SBP in the analysis was mean of first and second readings.
Diastolic Blood Pressure	4079 94	Same as systolic blood pressure.
Biochemistry Measurements		
Total Cholesterol	30690	Measured by CHO-POD analysis on a Beckman Coulter AU5800
LDL Cholesterol	17518	Measured by immunoturbidimetric analysis on a Beckman Coulter AU5800
HDL Cholesterol	30760	
Triglycerides	30870	
Apolipoprotein B	30640	
Apolipoprotein A1	30630	
HbA1c	30750	Measured by HPLC analysis on a Bio-Rad VARIANT II Turbo
Medical History and Medication		
Diabetes	20002 6177 6153 2443	1) Self-report diabetes, diagnosed by doctor (touch screen) 2) Verbal interview: Diabetes was classified as any diabetes (UKB field identifier 1220), type 1 diabetes (1222), type 2 diabetes (1223), diabetic eye disease (1276), diabetic neuropathy/ulcers (1468), and diabetic nephropathy (1607). 3) Self-reported insulin use. 4) screen-detected: HbA1c ≥6.5% (48 mmol/mol)
Hypertension	6177 6153 20002 2443 GP linkage	Hypertension was defined as: Any antihypertensive use as defined by verbal interview, touchscreen, and GP record. SBP ≥ 140 mmHg or DBP ≥ 90 mmHg
Antihypertensive and lipid-modifying medications	20003 6177 6153	Antihypertensive and lipid-modifying medications at baseline were defined using three sources of information: from verbal interview (Field ID: 20003), GP record linkage, and touchscreen questionnaire (Field ID: 6177, 6153).

Abbreviations: **UKB**: UK Biobank; **CKB**: China Kadoorie Biobank; **LDL**: Low-Density Lipoprotein; **HDL**: High-Density Lipoprotein; **MET**: Metabolic Equivalent of Task; **BMI**: Body Mass Index; **SBP**: Systolic Blood Pressure; **DBP**: Diastolic Blood Pressure; **CHO-POD**: Cholesterol Oxidase-Peroxidase; **HPLC**: High-Performance Liquid Chromatography; **HbA1c**: Haemoglobin A1c; **GP**: General Practitioner.

Table 3.4. Definition of prevalent and incident CVD outcomes in UKB

	UKB			
Conditions	**HES ICD-10 Field ID: 41270**	**Verbal Interview non-cancer illness Field ID: 20002**	**Verbal Interview Operations Field ID: 20004**	**Touchscreen Medical conditions Field ID: 6150, 4056 3894**
Prevalent MI	I21.x, I22.x, I23.x and diagnosis date is earlier than recruitment date	Verbal interview: self-reported MI diagnosis (1075)	Verbal interview: self-reported operations: coronary angioplasty (PTCA) +/- stent (1070), other arterial surgery/revascularisation procedures (1071), coronary artery bypass grafts (1095), carotid artery surgery/endarterectomy (1105), carotid artery angioplasty +/- stent (1109), coronary angiogram (1514)	Age heart attack diagnosed has value (Field ID: 3894) Self-reported (Field ID: 6150) heart attack diagnosis
Prevalent Stroke	I60.x, I61.x, I63.x, I64.x and diagnosis date is earlier than recruitment date	Verbal interview: self-reported stroke diagnosis (1081, 1082, 1583, 1491, 1086)		Age stroke diagnosed has value (Field ID: 4056) Self-reported (Field ID: 6150) stroke diagnosis.
Incident MI	I21.x, I22.x, I23.x and diagnosis date is later than recruitment date			

Abbreviations. **HES**: Hospital Episode Statistics; **MI**: Acute Myocardial Infarction; **UKB**: UK Biobank; **CKB**: China Kadoorie Biobank; **CVD**: Cardiovascular Disease

3.3.3. Genetic study of UKB

The UKB genetic study included genotype data for approximately 488,000 participants.[98] DNA was extracted from blood samples collected during assessment centre visits using standardised protocols that ensured high sample quality. Genotyping was performed using two closely related custom-designed arrays. Initially, around 50,000 participants from the UK BiLEVE study were genotyped using the UK BiLEVE Axiom Array, followed by the genotyping of an additional 438,000 participants with the UK Biobank Axiom Array, which shares 95% of its marker content with the UK BiLEVE array.

Genotyping was conducted in 106 sequential batches, each containing approximately 4,700 samples, at the Affymetrix Research Services Laboratory. A custom genotype calling pipeline was developed specifically for biobank-scale studies to optimise data quality across multiple batches. Comprehensive quality control procedures were implemented at both the sample and marker levels. These included assessments of call rates, heterozygosity, sex concordance, and the identification of batch and plate effects.[98,103] Following these quality control steps, high-quality genotype calls were generated for over 489,000 samples at more than 800,000 unique markers. Subsequent genotype imputation, using reference panels such as the Haplotype Reference Consortium and the 1000 Genomes Project, expanded the number of testable variants to around 96 million, thereby providing a powerful resource for genome-wide association studies and other genetic analyses.[98]

In UKB, PCA was performed on a subset of unrelated participants using common autosomal variants that passed quality control filters across all genotyping batches. The analysis identified major axes of genetic variation reflecting population structure

within UKB. The first ten genetic PCs were released as part of the UKB genetic data and were routinely included as covariates in genetic association analyses to account for population stratification.[98,103]

3.4. Baseline characteristics of data sources

3.4.1. Baseline characteristics of CKB participants by study subsets

The baseline characteristics for the full CKB cohort (N = 512,724) are summarised in ***Table 3.5***. The mean age at baseline was 52.0 years. There were marked sex differences in lifestyle factors, and current smoking was common in men (61.1%) but rare in women (2.4%), and weekly alcohol consumption was also much more common among men (33.3%) than women (2.1%). The mean BMI was 23.7 kg/m², and 35.4% of participants were classified as hypertensive.

Table 3.5. Baseline characteristics of participants in CKB cohort

Characteristics Mean (SD) or %	Males (n = 210,205)	Females (n = 302,519)	Overall (n = 512,724)
Demographics			
Age, years	52.9 (10.9)	51.5 (10.5)	52.0 (10.7)
Educated to middle school or above	57.8	43.3	49.2
Lifestyle factors			
Current smoker	61.1	2.4	26.4
Weekly drinker	33.3	2.1	14.9
Total physical activity, MET-h/day*	3.8 (3.7)	8.6 (4.4)	6.6 (4.7)
Medical history and medications			
Diabetes†	5.6	6.2	5.9
Hypertension‡	37.7	33.8	35.4
Antihypertensive use	11.6	12.7	12.3
Physical measurements			
BMI, kg/m²	23.4 (3.3)	23.8 (3.5)	23.7 (3.4)
SBP, mmHg	132.8 (20.0)	129.9 (22.0)	131.1 (21.3)
DBP, mmHg	79.2 (11.4)	76.8 (10.9)	77.8 (11.2)

Abbreviations. **SD**: Standard Deviation, **SBP**: Systolic Blood Pressure, **DBP**: Diastolic Blood Pressure, **BMI**: Body Mass Index, **MET**: Metabolic Equivalent of Task, **CKB**: the China Kadoorie Biobank.

* Total physical activity excludes work-related activity. Original total physical activity variable in CKB includes total and work-related physical activity, MET in the analyses was total minus work-related. This was done so the analyses are comparable across CKB and UKB.

† Diabetes includes self-reported cases and screen-detected cases (a blood glucose level ≥7.0 mmol/L and a fasting time >8 h, a blood glucose level ≥11.1 mmol/L and a fasting time <8 h, or a fasting blood glucose level ≥7.0 mmol/L).

‡ Prevalent hypertension was defined as self-reported hypertension, or having SBP≥140 mmHg or DBP≥90 mmHg, or currently taking antihypertensive medications.

The baseline characteristics for the genotyped subset in CKB (N = 100,640) are presented in ***Table 3.6***. Mean age was 53.7 years (SD 11.0), and 46.9% of participants had completed middle school education or above. Current smoking was reported by 27.5% but the percentage differed substantially between men and women. Weekly alcohol consumption was reported by 15.4% but the percentage also differed substantially between men (33.0%) and women (2.2%). Mean total physical activity was 6.6 MET-h/day (SD 4.8). Diabetes was present in 6.7% of the participants; 40.2% met criteria for hypertension, with 13.7% reporting current use of antihypertensive medications. Compared with the overall CKB cohort (mean age 52.0 years, 35.4% hypertensive, and 5.9% with diabetes), participants in the genotyped subset were slightly older (mean age 53.7 years), had a higher prevalence of hypertension (40.2%) and diabetes (6.7%).

Table 3.6. Baseline characteristics of participants in CKB genetic study

Characteristics Mean (SD) or %	Males (n = 43,015)	Females (n = 57,625)	Overall (n = 100,640)
Demographics			
Age, years	54.8 (11.2)	52.9 (10.8)	53.7 (11.0)
Educated to middle school or above	54.0	41.6	46.9
Lifestyle factors			
Current smoker	60.4	2.9	27.5
Weekly drinker	33.0	2.2	15.4
Total physical activity, MET-h/day*	3.9 (3.8)	8.7 (4.4)	6.6 (4.8)
Medical history and medications			
Diabetes†	6.3	7.0	6.7
Hypertension‡	43.0	38.2	40.2
Antihypertensive use	13.0	14.2	13.7
Physical measurements			
BMI, kg/m²	23.4 (3.3)	23.9 (3.6)	23.7 (3.5)
SBP, mmHg	135.5 (21.5)	132.2 (23.5)	133.6 (22.7)
DBP, mmHg	80.0 (12.0)	77.5 (11.4)	78.6 (11.7)

Abbreviations. **SD**: Standard Deviation, **SBP**: Systolic Blood Pressure, **DBP**: Diastolic Blood Pressure, **BMI**: Body Mass Index, **MET**: Metabolic Equivalent of Task, **CKB**: the China Kadoorie Biobank.
* Total physical activity excludes work-related activity. Original total physical activity variable in CKB includes total and work-related physical activity, MET in the analyses was total minus work-related. This was done so the analyses are comparable across CKB and UKB.
† Diabetes includes self-reported cases and screen-detected cases (a blood glucose level ≥7.0 mmol/L and a fasting time >8 h, a blood glucose level ≥11.1 mmol/L and a fasting time <8 h, or a fasting blood glucose level ≥7.0 mmol/L).
‡ Prevalent hypertension was defined as self-reported hypertension, or having SBP≥140 mmHg or DBP≥90 mmHg, or currently taking antihypertensive medications.

The baseline characteristics of the biochemistry subset (N = 18,253) are presented in ***Table 3.7***. Mean age was 57.0 years (SD 10.5), and 38.3% had completed middle school education or above. Diabetes was present in 8.6% of participants; 55.5% met criteria for hypertension, and 19.0% reported current use of antihypertensive medications. Mean BMI was 23.6 kg/m² (SD 3.5), mean SBP was 142.2 mmHg (SD 25.5). The mean plasma lipid levels in the biochemistry subset were total cholesterol 4.7 mmol/L, LDL-C 2.4 mmol/L, HDL-C 1.2 mmol/L, and triglycerides 2.0 mmol/L. A detailed comparison of the distributions of these lipid traits is presented in Chapter 4. Compared with the overall CKB cohort (mean age 52.0 years, 5.9% with diabetes, and 35.4% hypertensive), participants in the biochemistry subset were older (mean age 57.0 years), had higher prevalences of diabetes (8.6%), and hypertension (55.5%), while having similar BMI (23.6 vs 23.7 kg/m²).

Table 3.7. Baseline characteristics of participants in the CKB biochemistry subset

Characteristics Mean (SD) or %	Males (n = 9,273)	Females (n = 8,980)	Overall (n = 18,253)
Demographics			
Age, years	57.7 (10.6)	56.3 (10.3)	57.0 (10.5)
Educated to middle school or above	46.1	30.4	38.3
Lifestyle factors			
Current smoker	62.8	3.4	33.6
Weekly drinker	30.0	2.2	16.3
Total physical activity, MET-h/day*	4.0 (3.8)	9.0 (4.0)	6.44 (4.6)
Medical history and medications			
Diabetes†	7.6	9.7	8.6
Hypertension‡	55.7	55.3	55.5
Antihypertensive use	17.2	20.9	19.0
Physical measurements			
BMI, kg/m²	23.2 (3.4)	23.9 (3.7)	23.6 (3.5)
SBP, mmHg	142.4 (24.3)	142.0 (26.7)	142.2 (25.5)
DBP, mmHg	82.7 (13.6)	80.8 (13.2)	81.8 (13.4)
Lipid measurements			
Total cholesterol, mmol/L	4.5 (1.0)	4.8 (1.0)	4.7 (1.0)
Apolipoprotein B, g/L	0.8 (0.2)	0.9 (0.2)	0.8 (0.2)
Non-HDL-C, mmol/L	3.3 (0.9)	3.5 (1.0)	3.4 (1.0)
LDL-C, mmol/L	2.3 (0.7)	2.4 (0.7)	2.4 (0.7)
Triglycerides, mmol/L	2.0 (1.7)	2.0 (1.5)	2.0 (1.6)
Remnant cholesterol, mmol/L	1.0 (0.5)	1.1 (0.5)	1.1 (0.5)
Apolipoprotein A1, g/L	1.3 (0.2)	1.4 (0.2)	1.3 (0.2)
HDL-C, mmol/L	1.2 (0.3)	1.3 (0.3)	1.2 (0.3)

Abbreviations. **SD**: Standard Deviation, **SBP**: Systolic Blood Pressure, **DBP**: Diastolic Blood Pressure, **BMI**: Body Mass Index, **HDL-C**: High-Density Lipoprotein Cholesterol, **LDL-C**: Low-Density Lipoprotein Cholesterol, **MET**: Metabolic Equivalent of Task, **UKB**: UK Biobank; **CKB**: the China Kadoorie Biobank.

* Total physical activity excludes work-related activity. Original total physical activity variable in CKB includes total and work-related physical activity, MET in the analyses were total minus work-related. This was done so the analyses were comparable across CKB and UKB.

† Baseline diabetes includes self-reported cases and screen-detected cases (a blood glucose level ≥7.0 mmol/L and a fasting time >8 h, a blood glucose level ≥11.1 mmol/L and a fasting time <8 h, or a fasting blood glucose level ≥7.0 mmol/L).

‡ Prevalent hypertension was defined as self-reported hypertension, or having SBP≥140 mmHg or DBP≥90 mmHg, or currently taking antihypertensive medications.

Table 3.8 shows the heterogeneity in the distribution of plasma lipid levels across the ten CKB regions in the biochemistry subset. All three LDL-related markers, including ApoB, non-HDL-C and LDL-C, showed similar differences between regions, with highest values in the urban region of Haikou (ApoB 1.00 g/L; non-HDL-C 4.08 mmol/L; LDL-C 3.00 mmol/L) and lowest values in the rural region of Gansu (ApoB 0.76 g/L; non-HDL-C 3.02 mmol/L; LDL-C 2.07 mmol/L). Mean levels of triglycerides were highest in Qingdao (2.17 mmol/L) and lowest in Suzhou (1.58 mmol/L), while remnant-C levels were highest in Qingdao (1.25 mmol/L) and lowest in Sichuan (0.91 mmol/L). HDL-C and ApoA1 were highest in Sichuan (HDL-C 1.41 mmol/L; ApoA1 1.49 g/L), and lowest in Henan (HDL-C 1.16 mmol/L; ApoA1 1.27 g/L).

Table 3.8. Mean biochemical plasma lipid traits (SD) by region in CKB biochemistry subset

Lipid traits	Urban					Rural				
	Qingdao	Harbin	Haikou	Suzhou	Liuzhou	Sichuan	Gansu	Henan	Zhejiang	Hunan
ApoB, g/L	0.97 (0.21)	0.90 (0.20)	1.00 (0.18)	0.81 (0.18)	0.91 (0.20)	0.79 (0.19)	0.76 (0.17)	0.81 (0.18)	0.82 (0.19)	0.84 (0.18)
Non-HDL-C, mmol/L	3.99 (0.93)	3.66 (0.84)	4.08 (0.84)	3.24 (0.76)	3.70 (0.88)	3.21 (0.80)	3.02 (0.75)	3.25 (0.78)	3.25 (0.78)	3.39 (0.81)
LDL-C, mmol/L	2.74 (0.76)	2.52 (0.65)	3.00 (0.68)	2.29 (0.57)	2.67 (0.71)	2.30 (0.64)	2.07 (0.55)	2.27 (0.57)	2.27 (0.60)	2.31 (0.63)
Triglycerides, mmol/L	2.17 (1.45)	1.98 (1.25)	1.80 (1.22)	1.58 (1.03)	1.89 (1.23)	1.63 (0.97)	1.89 (1.10)	1.73 (1.08)	1.66 (1.01)	1.91 (1.12)
Remnant-C, mmol/L	1.25 (0.46)	1.14 (0.43)	1.07 (0.38)	0.95 (0.40)	1.03 (0.39)	0.91 (0.35)	0.96 (0.37)	0.98 (0.38)	0.98 (0.37)	1.08 (0.39)
HDL-C, mmol/L	1.26 (0.27)	1.21 (0.27)	1.26 (0.28)	1.29 (0.30)	1.21 (0.27)	1.41 (0.33)	1.24 (0.26)	1.16 (0.25)	1.30 (0.29)	1.28 (0.30)
ApoA1, g/L	1.36 (0.20)	1.29 (0.19)	1.33 (0.19)	1.38 (0.21)	1.31 (0.19)	1.49 (0.25)	1.34 (0.19)	1.27 (0.19)	1.41 (0.21)	1.38 (0.22)

Abbreviations. **SD**: Standard Deviation; **HDL-C**: High-Density Lipoprotein Cholesterol; **LDL-C**: Low-Density Lipoprotein Cholesterol; **ApoB:** Apolipoprotein B, **ApoA1:** Apolipoprotein A1; **Remnant-C:** Remnant Cholesterol.

Table 3.9 shows the mean SBP and plasma lipid traits by age and sex among controls in the CKB biochemistry subset. Mean levels of SBP were higher in older age bands, from 122.52 mmHg in those aged 30 - 39 years to 142.45 mmHg in those aged 70 - 79 years. Among LDL-related markers, mean levels of ApoB and non-HDL-C were higher in older participants, raising from 0.76 g/L and 3.07 mmol/L, respectively, in the youngest group (aged 30 - 39 years) to 0.83 g/L and 3.35 mmol/L in the oldest (aged 70 - 79 years). LDL-C followed a similar pattern, with the lowest mean levels of LDL-C of 2.11 mmol/L at age 30–39 years to the highest mean levels of LDL-C of 2.35 mmol/L at age 70–79 years. In contrast, mean levels of triglycerides and remnant-C showed no clear age trend, with only slight decreases observed. Mean levels of HDL-C and ApoA1 varied little across age groups, with mean HDL-C remaining around 1.22–1.27 mmol/L and ApoA1 ranging from 1.26 to 1.37 g/L. Patterns were broadly similar in men and women, although women tended to have higher levels of HDL-C and ApoA1, and slightly lower levels of triglycerides, across all age groups.

Table 3.9. Mean SBP and biochemical plasma lipid traits (SD) by age and sex in CKB biochemistry subset

Age, years	Sex	SBP, mmHg	ApoB, g/L	Non-HDL-C, mmol/L	LDL-C, mmol/L	Triglycerides, mmol/L	Remnant-C, mmol/L	HDL-C, mmol/L	ApoA1, g/L
30-39	Male	125.61 (13.21)	0.80 (0.19)	3.20 (0.84)	2.19 (0.67)	1.87 (1.12)	1.01 (0.35)	1.17 (0.27)	1.24 (0.17)
	Female	119.37 (13.97)	0.73 (0.16)	2.93 (0.65)	2.02 (0.52)	1.59 (0.81)	0.91 (0.25)	1.26 (0.26)	1.29 (0.16)
	Overall	122.52 (13.93)	0.76 (0.18)	3.07 (0.76)	2.11 (0.61)	1.73 (0.98)	0.96 (0.31)	1.22 (0.26)	1.26 (0.17)
40-49	Male	127.90 (15.84)	0.84 (0.20)	3.36 (0.84)	2.31 (0.63)	1.99 (1.30)	1.06 (0.42)	1.21 (0.27)	1.31 (0.21)
	Female	127.08 (19.32)	0.80 (0.18)	3.22 (0.76)	2.21 (0.60)	1.78 (1.08)	1.01 (0.36)	1.27 (0.27)	1.35 (0.18)
	Overall	127.46 (17.80)	0.82 (0.19)	3.29 (0.80)	2.26 (0.61)	1.88 (1.19)	1.03 (0.39)	1.24 (0.27)	1.33 (0.20)
50-59	Male	132.22 (19.20)	0.80 (0.19)	3.17 (0.82)	2.23 (0.62)	1.78 (1.18)	0.94 (0.39)	1.24 (0.30)	1.35 (0.24)
	Female	134.19 (21.75)	0.86 (0.20)	3.52 (0.86)	2.41 (0.67)	1.99 (1.15)	1.11 (0.41)	1.29 (0.28)	1.39 (0.20)
	Overall	133.23 (20.56)	0.83 (0.20)	3.35 (0.86)	2.32 (0.65)	1.89 (1.17)	1.03 (0.41)	1.27 (0.29)	1.37 (0.22)
60-69	Male	138.47 (21.70)	0.80 (0.19)	3.19 (0.79)	2.26 (0.61)	1.67 (1.04)	0.93 (0.36)	1.24 (0.31)	1.35 (0.24)
	Female	140.09 (22.32)	0.87 (0.20)	3.57 (0.87)	2.47 (0.68)	1.97 (1.12)	1.10 (0.39)	1.31 (0.28)	1.42 (0.21)
	Overall	139.20 (21.99)	0.83 (0.20)	3.36 (0.84)	2.35 (0.65)	1.80 (1.09)	1.00 (0.39)	1.27 (0.30)	1.38 (0.23)
70-79	Male	141.21 (21.01)	0.80 (0.19)	3.18 (0.80)	2.27 (0.63)	1.55 (0.94)	0.91 (0.34)	1.25 (0.30)	1.33 (0.22)
	Female	144.22 (22.54)	0.88 (0.19)	3.59 (0.83)	2.46 (0.64)	1.95 (1.12)	1.13 (0.44)	1.29 (0.30)	1.41 (0.21)
	Overall	142.45 (21.70)	0.83 (0.19)	3.35 (0.84)	2.35 (0.64)	1.71 (1.04)	1.00 (0.40)	1.27 (0.30)	1.37 (0.22)

Abbreviations. **SD**: Standard Deviation, **HDL-C**: High-Density Lipoprotein Cholesterol, **LDL-C**: Low-Density Lipoprotein Cholesterol, **ApoB:** Apolipoprotein B; **ApoA1:** Apolipoprotein A1; **Remnant-C:** Remnant Cholesterol.

3.4.2. Baseline characteristics of UKB participants by study subsets

The key demographics, lifestyle factors, medical history and core physical and lipid measurements for UKB participants (N = 437,545) are provided in ***Table 3.10***. Mean age at baseline was 56.9 years (SD 8.0), and mean duration of education was 14.8 (SD 5.1) years. Current smoking was reported by 10.2% overall, while weekly alcohol consumption was prevalent among 71.5% of participants. Mean total physical activity was 6.7 MET-h/day (SD 7.03). Overall, 5.5% of participants had diabetes at baseline (3.9% of males, 7.4% of females), 56.3% were classified as having prevalent hypertension (50.1% of men, 63.6% of women), and 20.8% reported using antihypertensive medication.

Mean BMI was 27.4 kg/m² (SD 4.8), mean systolic blood pressure 138.2 mmHg (SD 18.6) and mean diastolic blood pressure 82.3 mmHg (SD 10.1). Plasma lipids were characterised by a mean total cholesterol of 5.7 mmol/L (SD 1.1), ApoB of 1.0 g/L (0.2), non–HDL-C of 4.3 mmol/L (1.1), LDL-C of 3.6 mmol/L (0.9), triglycerides of 1.8 mmol/L (1.0), remnant cholesterol of 0.7 mmol/L (0.3), ApoA1 of 1.5 g/L (0.3) and HDL-C of 1.5 mmol/L (0.4). A detailed comparison of the distributions of these lipid traits in UKB is presented in Chapter 4.

Table 3.10. Baseline characteristics of participants in the UKB cohort study

Characteristics Mean (SD) or %	Males (n = 237,475)	Females (n = 200,070)	Overall (n = 437,545)
Demographics			
Age, years	56.7 (7.9)	57.1 (8.1)	56.9 (8.0)
Education years	14.5 (5.1)	15.2 (5.1)	14.8 (5.1)
Lifestyle factors			
Current smoker	8.8	11.9	10.2
Weekly drinker	64.6	79.5	71.5
Total physical activity, MET-h/day*	6.4 (6.5)	6.0 (5.9)	6.7 (7.0)
Medical history and medications			
Diabetes†	3.9	7.4	5.5
Hypertension‡	50.1	63.6	56.3
Antihypertensive use	18.9	23.0	20.8
Physical measurements			
BMI, kg/m²	27.9 (4.2)	27.0 (5.1)	27.4 (4.8)
SBP, mmHg	135.6 (19.2)	141.3 (17.4)	138.2 (18.6)
DBP, mmHg	80.7 (9.9)	84.2 (10.0)	82.3 (10.1)
Lipid measurements			
Total cholesterol, mmol/L	5.9 (1.1)	5.5 (1.1)	5.7 (1.1)
Apolipoprotein B, g/L	1.0 (0.2)	1.0 (0.2)	1.0 (0.2)
Non-HDL-C, mmol/L	4.3 (1.1)	4.2 (1.1)	4.3 (1.1)
LDL-C, mmol/L	3.6 (0.9)	3.5 (0.9)	3.6 (0.9)
Triglycerides, mmol/L	1.6 (0.9)	2.0 (1.2)	1.8 (1.0)
Remnant cholesterol, mmol/L	0.7 (0.3)	0.7 (0.3)	0.7 (0.3)
Apolipoprotein A1, g/L	1.6 (0.3)	1.4 (0.2)	1.5 (0.3)
HDL-C, mmol/L	1.6 (0.4)	1.3 (0.3)	1.5 (0.4)

Abbreviations. **SD**: Standard Deviation, **SBP**: Systolic Blood Pressure, **DBP**: Diastolic Blood Pressure, **BMI**: Body Mass Index, **HDL-C**: High-Density Lipoprotein Cholesterol, **LDL-C**: Low-Density Lipoprotein Cholesterol, **MET**: Metabolic Equivalent of Task, **UKB**: UK Biobank;
* Total physical activity excludes work-related activity.
† Baseline diabetes includes self-reported cases and screen-detected cases (a blood glucose level ≥7.0 mmol/L and a fasting time >8 h, a blood glucose level ≥11.1 mmol/L and a fasting time <8 h, or a fasting blood glucose level ≥7.0 mmol/L).
‡ Prevalent hypertension was defined as self-reported hypertension, or having SBP≥140 mmHg or DBP≥90 mmHg, or currently taking antihypertensive medications.

Table 3.11 shows the mean levels of SBP and plasma lipid traits by age and sex in UKB. Mean levels of SBP were higher in older age groups, increasing from 129.21 mmHg in participants aged 38 - 49 years to 146.42 mmHg in those aged 70 - 79 years. Among LDL-related traits, mean levels of ApoB, non-HDL-C, and LDL-C were higher in older participants, with the lowest values at 1.00 g/L, 4.11 mmol/L, and 3.47 mmol/L, respectively, at age 38–49 years and the highest values of 1.10 g/L, 4.58 mmol/L, and 3.86 mmol/L at age 70–79 years. Mean triglyceride levels showed little variation, ranging between 1.59 and 1.75 mmol/L, and mean levels of remnant-C were similar across age groups (0.65–0.74 mmol/L). Mean levels of HDL-C and ApoA1 also varied only modestly, with HDL-C ranging from 1.42 to 1.51 mmol/L and ApoA1 from 1.50 to 1.59 g/L. Across all age groups, women tended to have higher levels of HDL-C and ApoA1, and lower levels of triglycerides and remnant-C, compared with men.

Table 3.11. Mean SBP and biochemical plasma lipid traits (SD) by age and sex in UKB

Age, years	Sex	SBP, mmHg	ApoB, g/L	Non-HDL-C, mmol/L	LDL-C, mmol/L	Triglycerides, mmol/L	Remnant-C, mmol/L	HDL-C, mmol/L	ApoA1, g/L
38-49	Male	134.56 (14.26)	1.06 (0.23)	4.42 (1.00)	3.66 (0.78)	1.97 (1.16)	0.76 (0.32)	1.27 (0.29)	1.41 (0.22)
	Female	124.91 (15.34)	0.95 (0.21)	3.87 (0.92)	3.31 (0.73)	1.28 (0.72)	0.56 (0.24)	1.54 (0.34)	1.58 (0.26)
	Overall	129.21 (15.63)	1.00 (0.23)	4.11 (0.99)	3.47 (0.77)	1.59 (1.00)	0.65 (0.30)	1.42 (0.35)	1.50 (0.26)
50-59	Male	139.61 (16.51)	1.09 (0.23)	4.52 (0.97)	3.75 (0.76)	1.95 (1.10)	0.77 (0.31)	1.31 (0.30)	1.44 (0.23)
	Female	133.26 (17.82)	1.06 (0.22)	4.44 (1.00)	3.77 (0.80)	1.50 (0.81)	0.67 (0.27)	1.63 (0.36)	1.65 (0.27)
	Overall	135.91 (17.57)	1.07 (0.23)	4.47 (0.99)	3.76 (0.78)	1.69 (0.97)	0.71 (0.29)	1.50 (0.37)	1.57 (0.27)
60-69	Male	145.44 (18.18)	1.08 (0.22)	4.46 (0.93)	3.72 (0.74)	1.88 (1.00)	0.74 (0.28)	1.34 (0.31)	1.47 (0.23)
	Female	141.82 (19.15)	1.12 (0.23)	4.71 (0.99)	3.98 (0.79)	1.66 (0.83)	0.73 (0.27)	1.64 (0.36)	1.67 (0.26)
	Overall	143.35 (18.83)	1.10 (0.22)	4.61 (0.97)	3.87 (0.78)	1.75 (0.91)	0.74 (0.28)	1.51 (0.37)	1.59 (0.27)
70-79	Male	146.95 (17.66)	1.06 (0.20)	4.36 (0.91)	3.66 (0.72)	1.76 (0.94)	0.70 (0.27)	1.35 (0.31)	1.47 (0.23)
	Female	146.02 (19.73)	1.14 (0.24)	4.74 (1.04)	4.01 (0.82)	1.68 (0.81)	0.74 (0.28)	1.62 (0.36)	1.68 (0.26)
	Overall	146.42 (18.86)	1.10 (0.23)	4.58 (1.00)	3.86 (0.80)	1.71 (0.86)	0.72 (0.28)	1.50 (0.36)	1.59 (0.27)

Abbreviations. **SD**: Standard Deviation, **HDL-C**: High-Density Lipoprotein Cholesterol, **LDL-C**: Low-Density Lipoprotein Cholesterol, **ApoB:** Apolipoprotein B; **ApoA1:** Apolipoprotein A1; **Remnant-C:** Remnant Cholesterol, **UKB**: UK Biobank.

3.5. Summary

This chapter described the study design, data collection procedures, and baseline characteristics of study participants in the CKB and UKB cohorts, which are used in the analyses in this thesis.

Overall, both cohorts are large-scale, population-based studies with detailed survey data, physical and biochemical measurements, and long-term follow-up for morbidity and mortality through electronic health records.[99,101] However, there are important differences in their study designs and populations. CKB recruited adults aged 35 - 74 years from ten diverse regions of China, with substantial regional heterogeneity in socioeconomic, lifestyle, and regional exposures. In contrast, UKB recruited participants aged 40-70 years across the UK, who were generally more educated, had higher levels of BMI and plasma LDL-C, and reported a higher use of lipid-lowering medications at baseline. Smoking prevalence was higher in men in CKB compared to the overall UKB cohort, while alcohol consumption was more common in UKB. These differences in baseline distributions of exposures and confounders are important factors for the analyses presented in subsequent chapters of this thesis. The study designs also differ for the datasets used in this thesis. For plasma lipid analyses, CKB relied on a nested case-control subset with biochemical assays, whereas UKB analyses were conducted in the overall prospective cohort.[104]

Using these contemporary biobanks, this thesis is able to use methods from observational and genetic epidemiology to provide updated analyses of plasma lipid traits and BP and MI. The following chapters present findings from observational and MR analyses across these biobanks. Chapter 4 begins with the observational analysis of the association of plasma lipid traits with risk of MI.

Chapter 4 Observational associations of plasma lipid traits and MI

4.1. Background

This chapter compares observational associations including the shape and strength of the associations of plasma lipid traits with MI risk in Chinese and European populations. The analyses were conducted in statin-free adults in the CKB nested case-control subset and the UKB prospective cohort. The analyses compared the associations of LDL-, triglycerides-, and HDL-related traits with risk of MI in diverse ancestry populations, including Chinese and Europeans, and compared associations of conventional lipid traits (LDL-C, TG, HDL-C) with emerging lipid traits (ApoB, remnant-C, and ApoA1).

Although mean LDL-C levels are about one-third lower in Chinese than in European adults in population-based studies of untreated individuals, it remains uncertain whether these cross-ancestry differences influence the strength or shape of associations of cholesterol fractions, apolipoproteins, remnant-C, or lipid ratios with MI.[6] Since most MI outcomes occur in individuals wil moderately elevated levels of LDL-C, and many patients experience recurrent MI events despite treatment with standard doses of LDL-lowering medications, this raises the question about the optimum levels of LDL-C for lowest risk of MI in diverse populations.[11,12,17,18] Recognition of the importance of “residual” lipid-related risk beyond lowering LDL-C levels has also prompted the need for more research on comparative studies of the associations with MI with other lipids (apolipoproteins, non-HDL-C, remnant-C and lipid ratios) beyond conventional lipids (LDL-C, HDL-C and triglycerides) in Chinese and in European populations.

As summarised in Chapter 2, higher plasma levels of LDL-C show a positive, approximately linear association with higher risk of MI in Europeans. Plasma levels of HDL-C are inversely associated with MI risk in both ancestry groups.[9] ApoB, which is highly correlated with LDL-C, reflects the number of circulating atherogenic particles, since each LDL particle carries one ApoB molecule. Hence, plasma ApoB provides a direct measure of atherogenic particle burden.[105] The ERFC reported that additional LDL-related measures, such as ApoB and non-HDL-C, had associations with MI risk similar to those of LDL-C.[9] However, subsequent studies have indicated that ApoB may have a stronger association with risk of MI than LDL-C,[10-20] but no prior research has directly compared these associations in Chinese and European populations, who differ in diet, lifestyle, and distribution of plasma lipid traits.

Recent work has also highlighted remnant-C as a potential additional therapeutic target beyond LDL-C.[44] Lipid ratios such as LDL-C/ApoB (an indicator of LDL particle size) and HDL-C/ApoA1 (reflecting HDL particle size) have also been proposed as independent predictors of MI risk beyond LDL-C alone.[106,107] While small-scale studies suggested broadly similar associations of TChol with risk of MI between Chinese and European populations,[54,72,78] limited evidence exists on the relative importance of these relatively novel lipid measures for risk of MI in these populations.[6,73]

This chapter aims to update the evidence on the associations of conventional lipid traits (LDL-C, TG, and HDL-C) with risk of MI in Chinese and European populations, and to compare these with relatively novel lipid measures, including non-HDL-C, ApoB, remnant-C, ApoA1, and lipid ratios (LDL-C/ApoB and HDL-C/ApoA1).

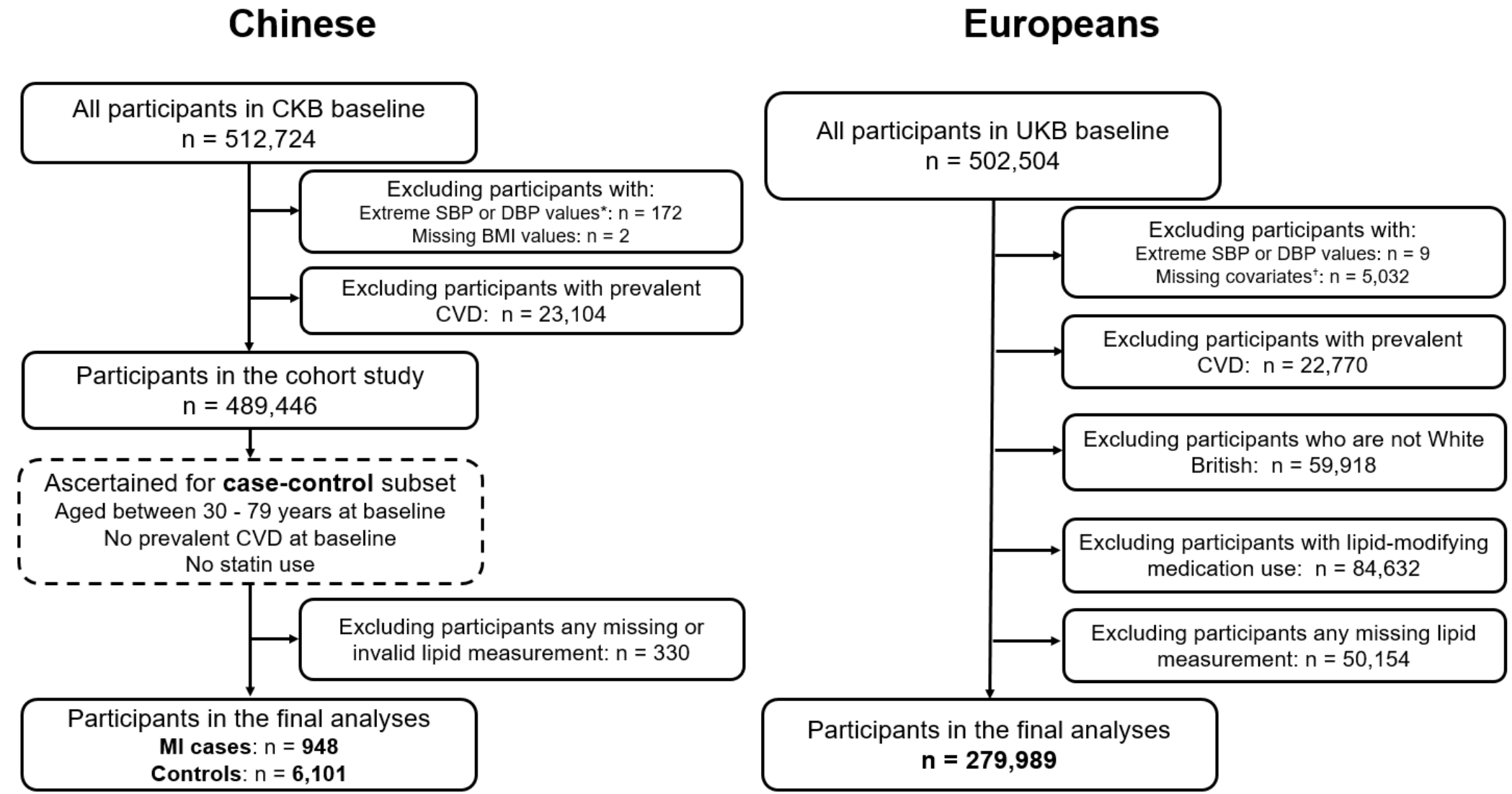
4.2. Methods

4.2.1. Study populations and settings

This chapter used lipid measurements from stored blood samples collected at enrolment for Chinese adults in CKB and Europeans in UKB. The overall study design of CKB and UKB was described in Chapter 3. The present chapter includes the nested case-control study in CKB. The nested case-control study selected 948 first incident non-fatal or fatal MI cases (ICD-10 codes I21 - I23) and 6,101 controls. Participants who reported statin use at baseline were excluded from the nested case-control study design.

For this chapter, the participants in UKB with a history of MI or stroke (n = 22,770) or use of statin therapy at baseline (n = 84,632) were excluded. The analyses were restricted to White British participants to allow for clear cross-ancestry comparison and evidence comparison for genetic analyses in Chapter 5. This includes participants who reported as non-white British in the baseline questionnaire and those who are not in the White British ancestry subset in the genetic subset (excluding n = 59,918 non-White participants). After further excluding participants with missing lipid values (n = 50,154), a total of 279,989 participants were included in the present analyses. Details of the inclusion criteria for the CKB and UKB participants in this chapter are provided in ***Figure 4.1***.

Figure 4.1. Flowchart for the inclusion of study participants in observational analysis



Abbreviations. **CKB**: China Kadoorie Biobank; **SBP**: Systolic Blood Pressure; **DBP**: Diastolic Blood Pressure; **MI**: Myocardial Infarction; **CVD**: Cardiovascular Disease; **BMI**: Body Mass Index.
MI cases in the Chinese population were defined at the time of ascertainment.
*Extreme SBP: SBP ≥ 250 or < 80 mmHg and extreme DBP: DBP ≥ 140 or < 40 mmHg.
† Missing covariates in UKB include: deprivation index (n = 527); BMI (n = 1,937); SBP (n = 797), educational attainment (n = 1,771).

4.2.2. Statistical methods

Baseline characteristics were summarised as means (SDs) for continuous variables and as numbers and percentages for categorical variables. Age- and sex-adjusted Spearman correlation coefficients were calculated to examine the associations among individual lipid traits. For each lipid trait, linear regression models were fitted with age and sex as covariates, and the residuals from these models were obtained. Correlations were then calculated between the residuals for the different lipid traits, providing estimates of the associations between lipid traits that were independent of age and sex. Plasma levels of triglycerides and remnant-C were log-transformed prior to analysis to reduce skewness, stabilise variance, and improve the linearity of associations with the logit and log-hazard functions in logistic and Cox regression models.[9,108] RRs were estimated using logistic regression in CKB and Cox regression in UKB, given both ORs and HRs approximate RRs when the outcomes were relatively rare (< 10%).[109]

In CKB, logistic regression was used to estimate the RRs of MI for fifths of lipid traits after adjustment for age (continuous), sex, region, BMI, diabetes, alcohol use, smoking status, physical activity, education, SBP, ambient temperature, and ambient temperature-squared. These variables were included as covariates because sex, region, BMI, diabetes, alcohol use, smoking, physical activity, education, and SBP are well-established risk factors for MI and are also correlated with lipid profiles, thereby acting as potential confounders.[9,110] In addition, previous analyses in CKB have shown that ambient temperature influences metabolic traits and was therefore considered an additional confounding factor.[111-113]

In UKB, follow-up time was split into 5-year age-at-risk bands using the Lexis expansion method, which restructures each individual's record into multiple intervals

defined by attained age. Stratifying by these 5-year age bands allows each stratum to have its own baseline hazard function, thereby helping to satisfy the proportional hazards assumption.[114] Cox regression with time in study as timescale, stratified by 5-year age-at-risk groups and sex, and adjusted for BMI, diabetes, alcohol use, smoking status, physical activity, ethnicity, Townsend Deprivation Index, years of education, and antihypertensive medications, was used to estimate RRs for the associations of fifths of specific blood lipids with MI. Age-at-risk refers to the participant's age during follow-up rather than age at baseline. The 95% CIs were estimated using group-specific variances (or the floating absolute risk method), which derive confidence limits from the variance of the log risk for each group.[115] This approach enables direct comparisons between any two exposure groups and avoids dependence on an arbitrarily chosen reference category, as in conventional methods.

To assess the incremental contribution of individual lipid measures, likelihood ratio (LR) tests were used to compare nested models before and after adding each lipid variable.[116] Unlike arbitrary thresholds such as a 10% change in the log point estimate, the LR test provides a formal statistical framework grounded in maximum likelihood theory to evaluate whether including an additional variable improves the overall fit of the model to the data. The LR test compares the goodness-of-fit between nested models by evaluating whether the inclusion of an additional variable significantly improves model fit. The test statistic follows a χ^2 distribution, with higher values indicating stronger evidence that the added lipid measure provides incremental explanatory value beyond the baseline model. This approach appropriately accounts for overall model fit and parameter uncertainty, offering a

rigorous and objective assessment of the additional information provided by each lipid measure.[116]

Ridge regression models were used to assess the associations of LDL-C and ApoB with MI risk, and of HDL-C and ApoA1 with MI risk. Ridge regression is a regularisation method that addresses multicollinearity by shrinking regression coefficients towards zero through the addition of an L2 penalty, defined as the sum of squared coefficients ($\lambda\sum\beta^2$). This penalty term is added to the residual sum of squares in the estimation procedure. The effect of the L2 penalty is to distribute influence across correlated predictors, reduce variance, and stabilise coefficient estimates without forcing any coefficient exactly to zero.[117]

Repeat measures of all lipid traits (except lipid ratios) were used to correct for regression dilution bias using the MacMahon-Peto method,[76] based on data from the second resurvey in both CKB and UKB as outlined in Chapter 3. This approach was chosen because it provides a simple empirical estimate of the regression dilution ratio using repeat measurements, without requiring strong distributional assumptions.[118] The second resurvey data allow assessment of long-term within-person variability, ensuring that corrections reflect “usual” lipid levels over extended follow-up. The correction of regression dilution is essential because baseline lipid concentrations are subject to both random measurement error and within-person variability over time, which causes underestimation of the true associations with disease risk. The use of resurvey data enables estimation of the long-term “usual” levels of exposures, thereby reducing attenuation of risk estimates and providing more reliable estimates of associations with risk of disease compared to uncorrected estimates.[76] Sensitivity analyses excluded participants with prevalent diabetes in CKB to reduce residual confounding by diabetes and explored associations at low

LDL-C levels in UKB to facilitate comparisons with CKB. All analyses were conducted in R version 4.2.2.

4.3. Results

4.3.1. Baseline characteristics

Table 4.1 presents the baseline characteristics for MI cases and controls in CKB and for all participants in UKB. Mean levels of age were higher among MI cases in CKB (58.3 years) compared with controls in CKB (52.4 years) and UKB (55.5 years). UKB included a higher proportion of females than CKB. Participants in UKB had higher levels of full-time education than CKB. Among lifestyle factors, a higher percentage of Chinese (CKB controls) than European men was current smokers (61.2% vs 11.9%), but a higher proportion of European than Chinese men were current drinkers (80.6% vs 28.2%). Few Chinese women were current smokers or current drinkers. Both populations had similar levels of physical activity. A higher proportion of Chinese than Europeans had diabetes at baseline (4.8% vs 1.6%). Use of antihypertensive medication was similar in CKB and UKB (11.5% vs 12.8%). Mean levels of BMI in CKB participants were lower than in UKB (23.0 kg/m² vs 27.0 kg/m²), but mean levels of SBP were comparable in CKB and UKB (135.0 mmHg vs 137.0 mmHg). Mean levels of LDL-C were about 30% lower in CKB than in UKB (2.3 mmol/L vs 3.7 mmol/L), as were mean levels of non-HDL-C (3.3 mmol/L vs 4.4 mmol/L).

Table 4.1. Baseline characteristics of participants in CKB and UKB studies

Characteristics Mean (SD) or %	CKB		UKB
	MI cases (n = 948)	Controls (n = 6,101)	(n = 279,989)
Demographics			
Age, years	52.4 (7.1)	58.3 (10.9)	55.5 (8.0)
Female	39.6	47.9	57.1
Educated to middle school or above*	46.1	32.8	62.0
Lifestyle factors			
Male current smoker	75.1	61.2	11.9
Female current smoker	3.5	3.2	8.5
Male weekly drinker	27.0	28.2	80.6
Female weekly drinker	1.3	1.9	66.3
Total physical activity, MET-h/day**	5.5 (4.4)	6.5 (4.5)	6.4 (6.5)
Medical history and medications			
Diabetes†	11.9	4.8	1.6
Hypertension‡	54.2	41.5	49.7
Antihypertensive use	17.7	11.5	12.8
Physical measurements			
BMI, kg/m²	24.0 (3.6)	23.0 (3.3)	27.0 (4.6)
SBP, mmHg	140.8 (25.7)	135.0 (21.3)	137.0 (18.5)
DBP, mmHg	82.7 (13.6)	77.3 (11.0)	82.3 (10.1)
Lipid measurements			
Total cholesterol, mmol/L	4.8 (1.0)	4.6 (0.9)	5.9 (1.0)
Apolipoprotein B, g/L	0.9 (0.2)	0.8 (0.2)	1.1 (0.2)
Non-HDL-C, mmol/L	3.6 (1.1)	3.3 (0.9)	4.4 (1.0)
LDL-C, mmol/L	2.4 (0.8)	2.3 (0.6)	3.7 (0.8)
LDL-C/Apolipoprotein B, mmol/g	2.7 (0.3)	2.8 (0.3)	3.5 (0.2)
Triglycerides, mmol/L	2.0 (1.3)	1.8 (1.1)	1.7 (1.0)
Remnant cholesterol, mmol/L	1.2 (0.5)	1.0 (0.4)	0.7 (0.3)
Apolipoprotein A1, g/L	1.2 (0.2)	1.4 (0.2)	1.6 (0.3)
HDL-C, mmol/L	1.2 (0.3)	1.3 (0.3)	1.5 (0.4)
HDL-C/Apolipoprotein A1, mmol/g	0.9 (0.1)	0.9 (0.1)	0.9 (0.1)

Abbreviations. **SD**: Standard Deviation, **SBP**: Systolic Blood Pressure, **DBP**: Diastolic Blood Pressure, **BMI**: Body Mass Index, **HDL-C**: High-Density Lipoprotein Cholesterol, **LDL-C**: Low-Density Lipoprotein Cholesterol, **MET**: Metabolic Equivalent of Task, **UKB**: UK Biobank; **CKB**: the China Kadoorie Biobank.

* In UKB, this category includes those who were educated to above A levels/AS levels or equivalent.

** Total physical activity excludes work-related activity. UKB has around 20% missing data in self-reported physical activity.

† In UKB, baseline diabetes includes self-reported cases or those participants with a HbA1c level above 48 mmol/mol; In CKB, it includes self-reported cases and screen-detected cases (a blood glucose level ≥7.0 mmol/L and a fasting time >8 h, a blood glucose level ≥11.1 mmol/L and a fasting time <8 h, or a fasting blood glucose level ≥7.0 mmol/L).

‡ In both UKB and CKB, prevalent hypertension was defined as self-reported hypertension, or having SBP≥140 mmHg or DBP≥90 mmHg, or currently taking antihypertensive medications.

Table 4.2 compares mean plasma levels of lipid traits at baseline and at resurvey in CKB and UKB, and provides RDRs for each trait. Mean levels of ApoB, non-HDL-C, and LDL-C were consistently lower in CKB than in UKB, whereas triglyceride and remnant-C levels were slightly higher in CKB. RDRs, which quantify the proportion of long-term variability in baseline measures, were broadly similar between populations, ranging from 0.5 to 0.8 for most traits, with triglycerides and remnant-C showing the highest variabilities.

Table 4.2. Mean plasma levels of lipids at baseline and resurvey

	China Kadoorie Biobank			UK Biobank		
	Baseline	Resurvey	RDR*	Baseline	Resurvey	RDR
ApoB, g/L						
Baseline fifths						
I	0.57	0.69		0.77	0.84	
II	0.71	0.81		0.93	0.98	
III	0.82	0.89		1.05	1.06	
IV	0.94	0.96		1.16	1.13	
V	1.12	1.08		1.38	1.26	
Difference (V - I)	0.55	0.40	***0.72***	0.62	0.42	***0.68***
non-HDL-C, mmol/L						
Baseline fifths						
I	2.27	2.84		3.09	3.42	
II	2.85	3.20		3.82	4.00	
III	3.29	3.57		4.33	4.37	
IV	3.76	3.92		4.86	4.69	
V	4.64	4.42		5.83	5.24	
Difference (V - I)	2.37	1.58	***0.67***	2.73	1.81	***0.66***
LDL-C, mmol/L						
Baseline fifths						
I	1.50	1.97		2.66	2.97	
II	1.94	2.20		3.26	3.43	
III	2.28	2.52		3.66	3.69	
IV	2.65	2.79		4.07	3.98	
V	3.33	3.16		4.84	4.41	
Difference (V - I)	1.84	1.19	***0.65***	2.17	1.44	***0.66***
Triglycerides, mmol/L						
Baseline fifths						
I	0.76	1.25		0.75	1.03	
II	1.17	1.71		1.08	1.31	
III	1.55	1.82		1.41	1.58	
IV	2.14	2.18		1.90	1.86	
V	3.69	2.93		3.10	2.48	
Difference (V - I)	2.92	1.68	***0.58***	2.36	1.46	***0.62***
Remnant-C, mmol/L						
Baseline fifths						
I	0.58	0.79		0.35	0.46	
II	0.79	0.92		0.53	0.57	
III	0.95	1.05		0.65	0.65	
IV	1.16	1.11		0.80	0.71	
V	1.64	1.43		1.12	0.87	
Difference (V - I)	1.06	0.63	***0.60***	0.76	0.41	**0*.53***
ApoA1, g/L						
Baseline fifths						
I	1.09	1.19		1.22	1.33	
II	1.23	1.28		1.40	1.47	
III	1.34	1.37		1.52	1.59	
IV	1.45	1.44		1.66	1.72	
V	1.68	1.59		1.94	1.91	
Difference (V - I)	0.59	0.40	***0.69***	0.73	0.58	***0.79***
HDL-C, mmol/L						
Baseline fifths						
I	0.92	1.02		1.02	1.13	
II	1.09	1.13		1.26	1.35	
III	1.23	1.25		1.44	1.52	
IV	1.39	1.34		1.65	1.73	
V	1.71	1.60		2.04	2.04	
Difference (V - I)	0.79	0.58	***0.73***	1.02	0.90	***0.89***

Abbreviations. RDRs: Regression Dilution Ratios; **LDL-C**: Low-Density Lipoprotein Cholesterol; **HDL-C**: High-Density Lipoprotein Cholesterol; **ApoB**: Apolipoprotein B; **ApoA1**: Apolipoprotein A1; **Remnant-C**: Remnant Cholesterol.
*RDRs were calculated by ratios of interquintile ranges of resurvey and baseline (MacMahon-Peto method).

Age- and sex-adjusted correlations of lipid traits are shown in ***Figure 4.2***. In both populations, LDL-C, non-HDL-C, and ApoB were strongly correlated with each other ($r > 0.9$), while triglycerides and remnant-C were also highly correlated ($r \approx 0.7$). HDL-C and ApoA1 showed strong correlations ($r \approx 0.9$). By contrast, correlations between LDL-related and HDL-related traits were weak or inverse, and triglyceride-related measures were positively correlated with LDL-related traits and inversely correlated with HDL-related traits. Overall, the correlation patterns were consistent across populations, though slightly stronger correlations between LDL-related measures were observed in Europeans compared with Chinese.

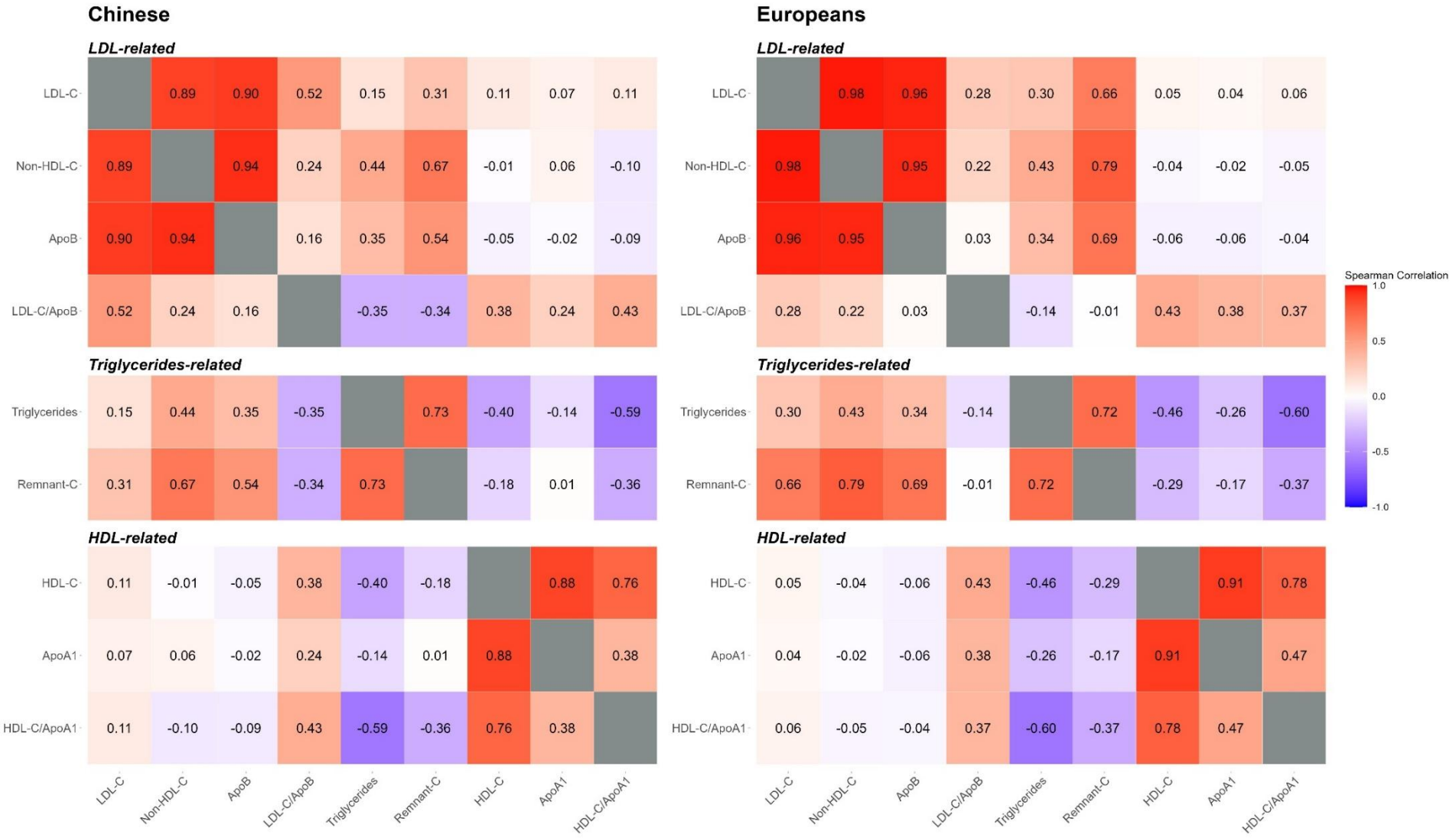
In addition to lower mean values, the distributions of lipid traits also differed between populations (***Figure 4.3***). In CKB, LDL-C, non-HDL-C, and ApoB were not only shifted to lower levels but also showed narrower distributions, with relatively few individuals at the upper end of the range. By contrast, in UKB the distributions were broader, with a substantial rightward tail, particularly for LDL-C. The distribution of the LDL-C/ApoB ratio was similar in these two ancestry populations, but values in UKB were shifted to slightly higher levels. In both populations, LDL-related measures were positively correlated with triglyceride-related measures, while HDL-related measures were inversely correlated with triglyceride-related measures. In UKB, the prevalence of diabetes was inversely associated with fifths of LDL-C/ApoB and HDL-C/ApoA1 ratios (***Table 4.3***). In CKB, the prevalence of diabetes was similar across fifths of LDL-C/ApoB, but was positively associated with fifths of remnant-C, and inversely associated with fifths of HDL-C/ApoA1.

Table 4.3 Prevalence of diabetes, by lipid traits

	China Kadoorie Biobank			UK Biobank		
	Range	N	Diabetes, %	Range	N	Diabetes, %
LDL-C/ApoB, mmol/g						
Baseline fifths						
I	1.21 – 2.57	1,221	6.47	1.85 – 3.33	55,998	3.64
II	2.57 – 2.75	1,220	4.18	3.33 – 3.45	55,998	1.67
III	2.75 – 2.89	1,223	3.76	3.45 – 3.55	55,998	1.13
IV	2.89 – 3.05	1,217	4.35	3.55 – 3.67	55,997	0.86
V	3.05 – 3.77	1,220	5.16	3.67 – 5.37	55,998	0.77
Remnant-C, mmol/L						
Baseline fifths						
I	0.17 – 0.70	1,226	3.10	0.00 – 0.46	56,073	1.40
II	0.70 – 0.86	1,226	3.18	0.46 – 0.59	56,111	1.34
III	0.86 – 1.03	1,209	3.64	0.59 – 0.72	55,882	1.43
IV	1.03 – 1.27	1,222	5.16	0.72 – 0.90	55,933	1.73
V	1.27 – 3.88	1,218	8.87	0.90 – 4.18	55,990	2.17
HDL-C/ApoA1, mmol/g						
Baseline fifths						
I	0.45 – 0.83	1,221	8.11	0.49 – 0.86	55,998	3.71
II	0.83 – 0.89	1,221	6.63	0.86 – 0.92	55,999	1.77
III	0.89 – 0.94	1,219	4.10	0.92 – 0.97	55,996	1.18
IV	0.94 – 1.00	1,220	3.03	0.97 – 1.03	55,997	0.80
V	1.00 – 1.39	1,220	2.05	1.03 – 1.44	55,998	0.62

Abbreviations. RRs: Risk Ratios; **LDL-C**: Low-Density Lipoprotein Cholesterol; **HDL-C**: High-Density Lipoprotein Cholesterol; **ApoB**: Apolipoprotein B; **ApoA1**: Apolipoprotein A1; **Remnant-C:** Remnant Cholesterol.
In the Chinese population, the table was restricted to controls (participants without CVD) only.

Figure 4.2. Correlations between different lipid measures in Chinese and Europeans



Abbreviations. **ApoB**: Apolipoprotein B; **LDL-C**: Low-Density Lipoprotein Cholesterol; **HDL-C**: High-Density Lipoprotein Cholesterol; **ApoA1**: Apolipoprotein A1; **Remnant-C**: Remnant Cholesterol
Correlations adjusted for age and sex. Correlations in the Chinese population were calculated in controls only.

4.3.2. LDL-C-related lipid traits

The shape of the association of usual plasma levels of LDL-C with MI was non-linear in Chinese, with weak evidence of an association with risk of MI at LDL-C levels below 3 mmol/L (***Figure 4.3***). By contrast, in Europeans the association was log-linear across the full range of LDL-C values. The associations of usual plasma levels of non-HDL-C with risk of MI were broadly log-linear in both Chinese and European populations, with similar RR across the exposure range. The RRs (95% CI) for risk of MI per 1-SD higher non-HDL-C were 1.47 (1.30–1.66) in Chinese and 1.59 (1.53–1.65) in Europeans.

For ApoB, the association with risk of MI was also approximately log-linear in both populations, with stronger RR for risk of MI per 1-SD higher levels than for LDL-C in the Chinese population, and no evidence of a threshold effect at lower concentrations. The RRs (95% CI) for risk of MI per 1-SD higher ApoB were 1.45 (1.29–1.62) in Chinese and 1.56 (1.51–1.62) in Europeans.

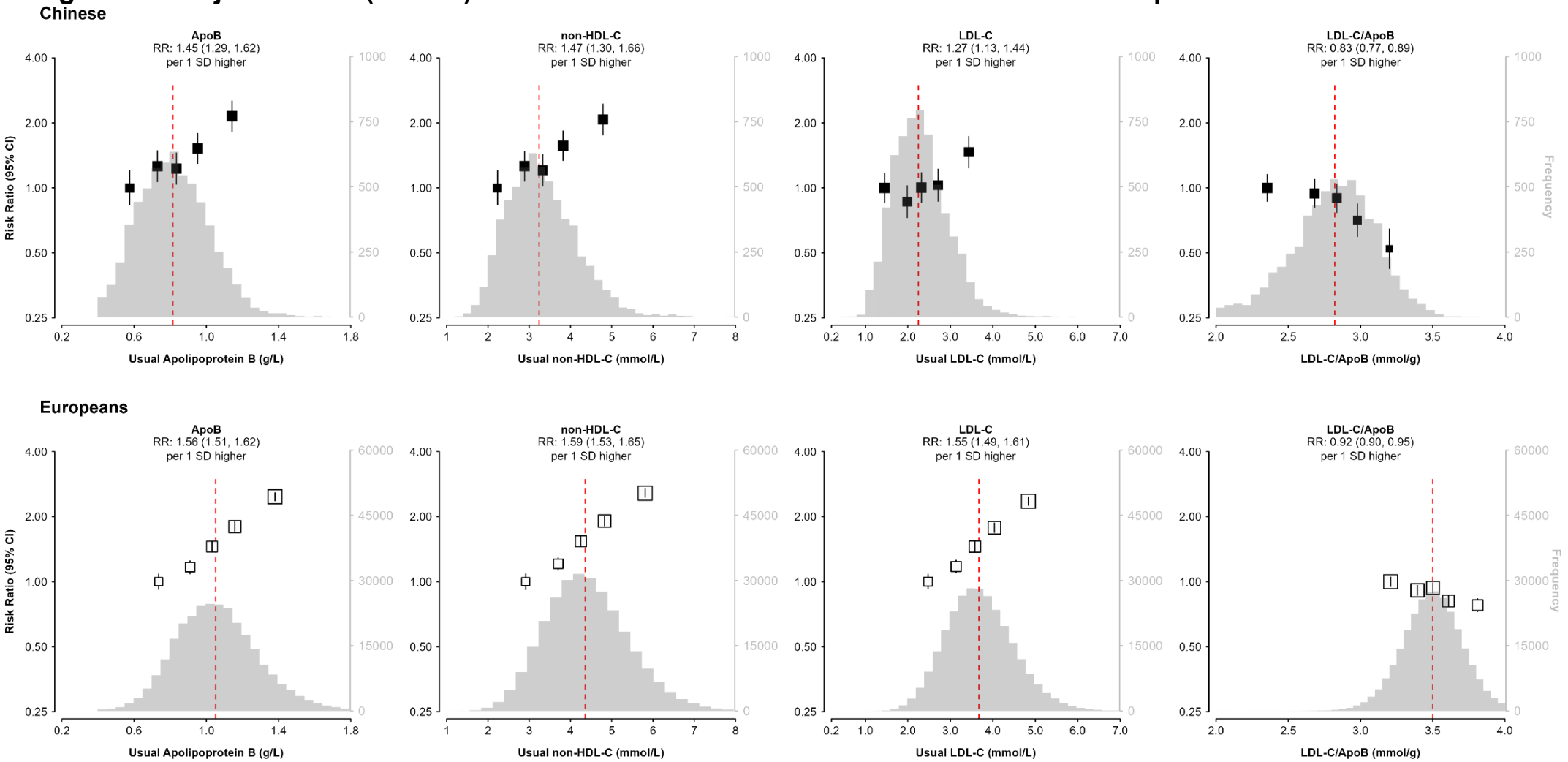
Higher levels of LDL-C/ApoB ratios were inversely associated with risk of MI in both populations, with broadly linear associations across the range of observed values. The RRs (95% CI) per 1-SD higher LDL-C/ApoB ratio were 0.83 (0.77–0.89) in Chinese and 0.92 (0.90–0.95) in Europeans.

The non-linear association of plasma levels of LDL-C and risk of MI in the Chinese population still existed after excluding participants with diabetes at baseline (***Figure 4.4***).

Consistent with the pattern observed in the Chinese population at low plasma levels of LDL-C, the strength of association of usual plasma levels of LDL-C with MI was greatly attenuated below 3 mmol/L LDL-C in Europeans (***Figure 4.5***).

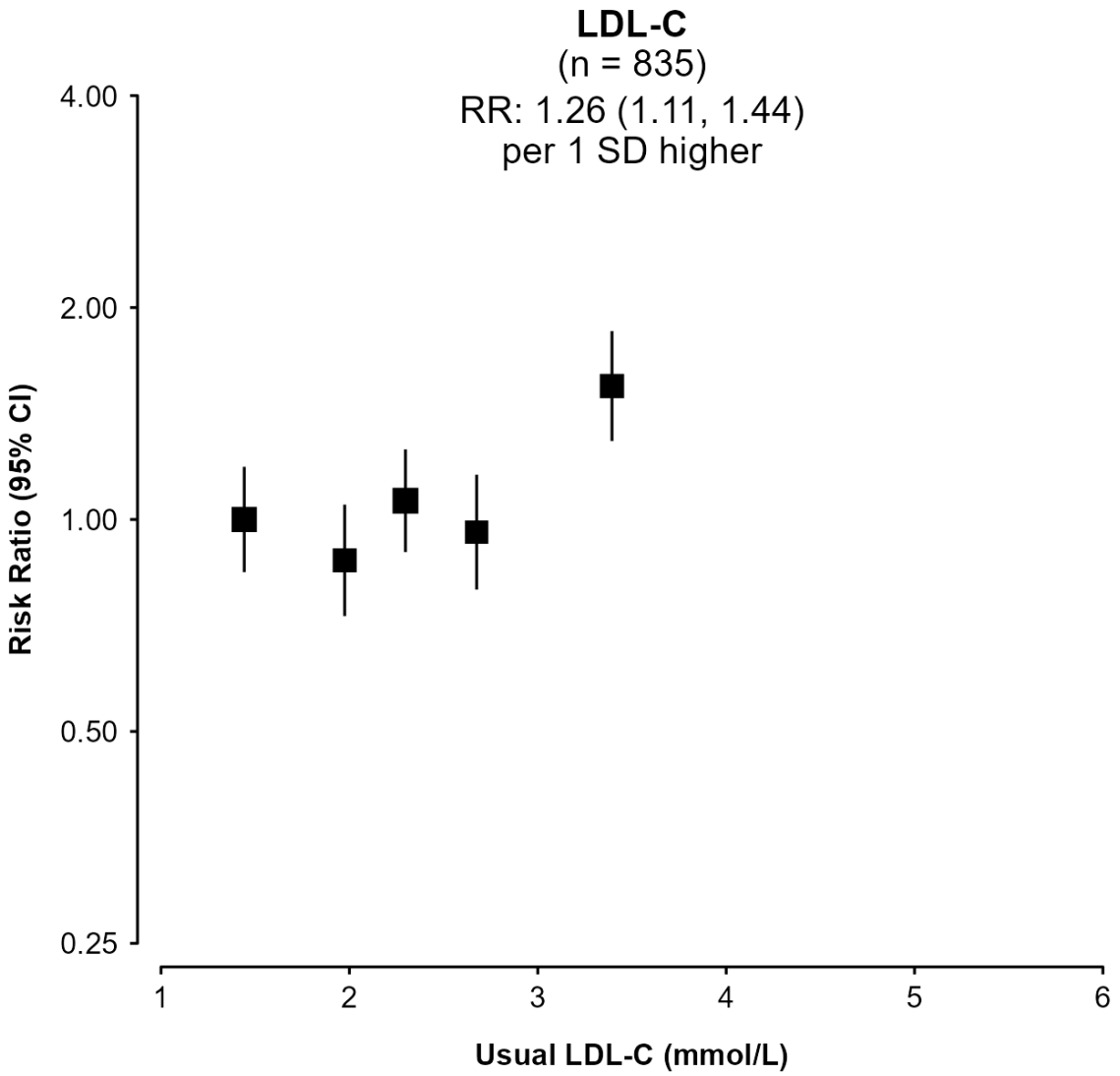
Overall, the strength of association of higher usual levels of LDL-C in Chinese was around half those observed in European populations (RR [95% CI] per 1-SD higher level 1.27 [1.13–1.44] vs 1.55 [1.49–1.61]) after adjustment for conventional risk factors. In contrast with LDL-C, both non-HDL-C and ApoB were equally strongly associated with risk of MI in both Chinese and European populations. The χ^2 values for adding ApoB, non-HDL-C, and LDL-C in Chinese were 41.1, 49.0, and 17.3, respectively, while the corresponding χ^2 values in Europeans were 563.7, 556.5, and 496.2 (***Figure 4.6***). After further lipid adjustment, the associations of ApoB and non-HDL-C with risk of MI remained strong in both populations (χ^2 values of 45.8 and 61.7 in Chinese, and 431.4 and 447.7 in Europeans, respectively). The associations of LDL-C with risk of MI also remained strong after further adjustment, with χ^2 values of 25.0 in Chinese and 441.1 in Europeans.

Figure 4.3. Adjusted RRs (95% CI) of MI for LDL-related measures in Chinese and Europeans



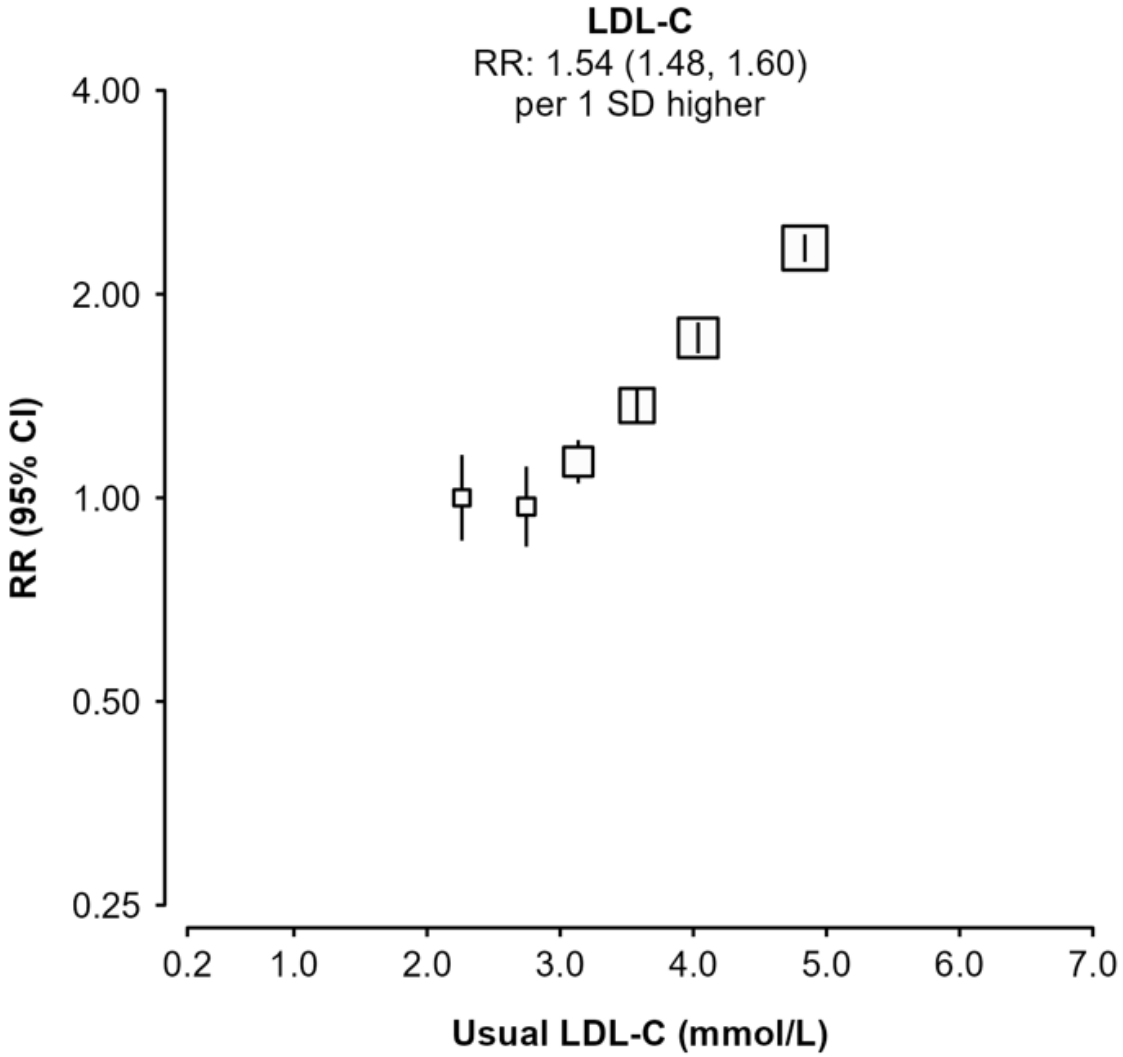
In the Chinese population, the RRs were adjusted for age, sex, region, BMI, diabetes, alcohol use, smoking status, SBP, physical activity, educational attainment, antihypertensive medications, mean temperature, and mean temperature-squared. In European populations, the RRs were stratified by 5-year age-at-risk bands and sex, and adjusted for BMI, diabetes, alcohol use, smoking status, SBP, physical activity, Townsend Deprivation Index, years of education, and antihypertensive medication. The distributions of blood lipids in the Chinese population were obtained from baseline samples in controls only. The red dashed lines indicate median values. SDs are from baseline measurements and can be found in Table 4.2.

Figure 4.4. RR (95% CI) of MI by fifths of LDL-C in CKB, excluding individuals with diabetes



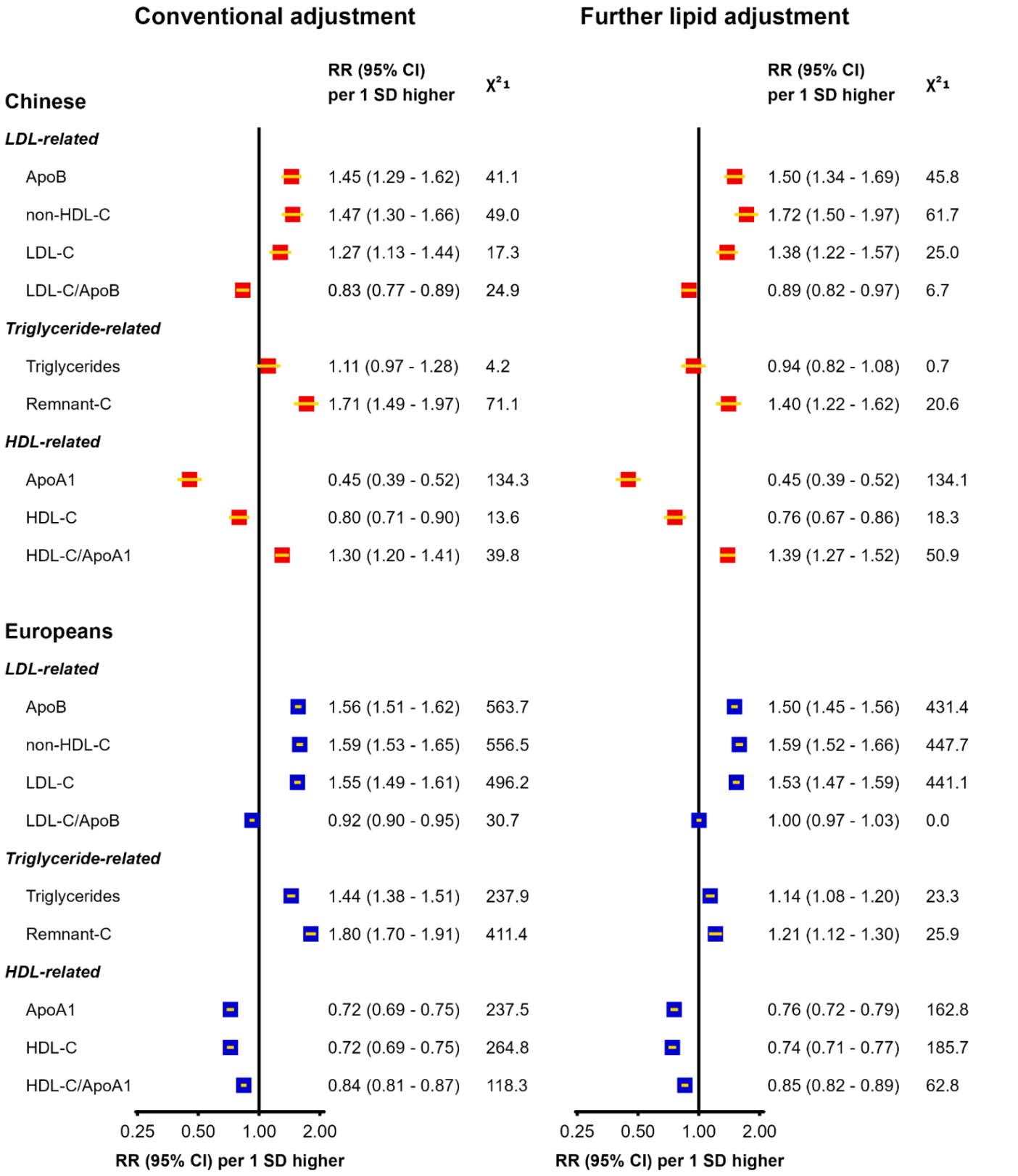
RRs were adjusted for age, sex, region, BMI, alcohol use, smoking status, SBP, physical activity, educational attainment, antihypertensive medications, mean temperature, and mean temperature-squared.

Figure 4.5. RR (95% CI) of MI by fifths of LDL-C in UKB, after dividing the first quintile



RRs were stratified by 5-yrs age-at-risk band and sex, and adjusted for BMI, diabetes, alcohol use, smoking status, SBP, physical activity, Townsend Deprivation Index, years of education, and antihypertensive medications.

Figure 4.6. Risk ratios (95%CI) of MI before and after additional adjustment for lipids in other classes



Abbreviations: RRs: Risk Ratios; **LR**: Likelihood Ratio; **ApoB**: Apolipoprotein B; **LDL-C**: Low-Density Lipoprotein Cholesterol; **non-HDL-C**: non-High-Density Lipoprotein Cholesterol; **Remnant-C**: Remnant Cholesterol**; HDL-C**: High-Density Lipoprotein Cholesterol; **ApoA1**: Apolipoprotein A1; **SD**: Standard Deviation.

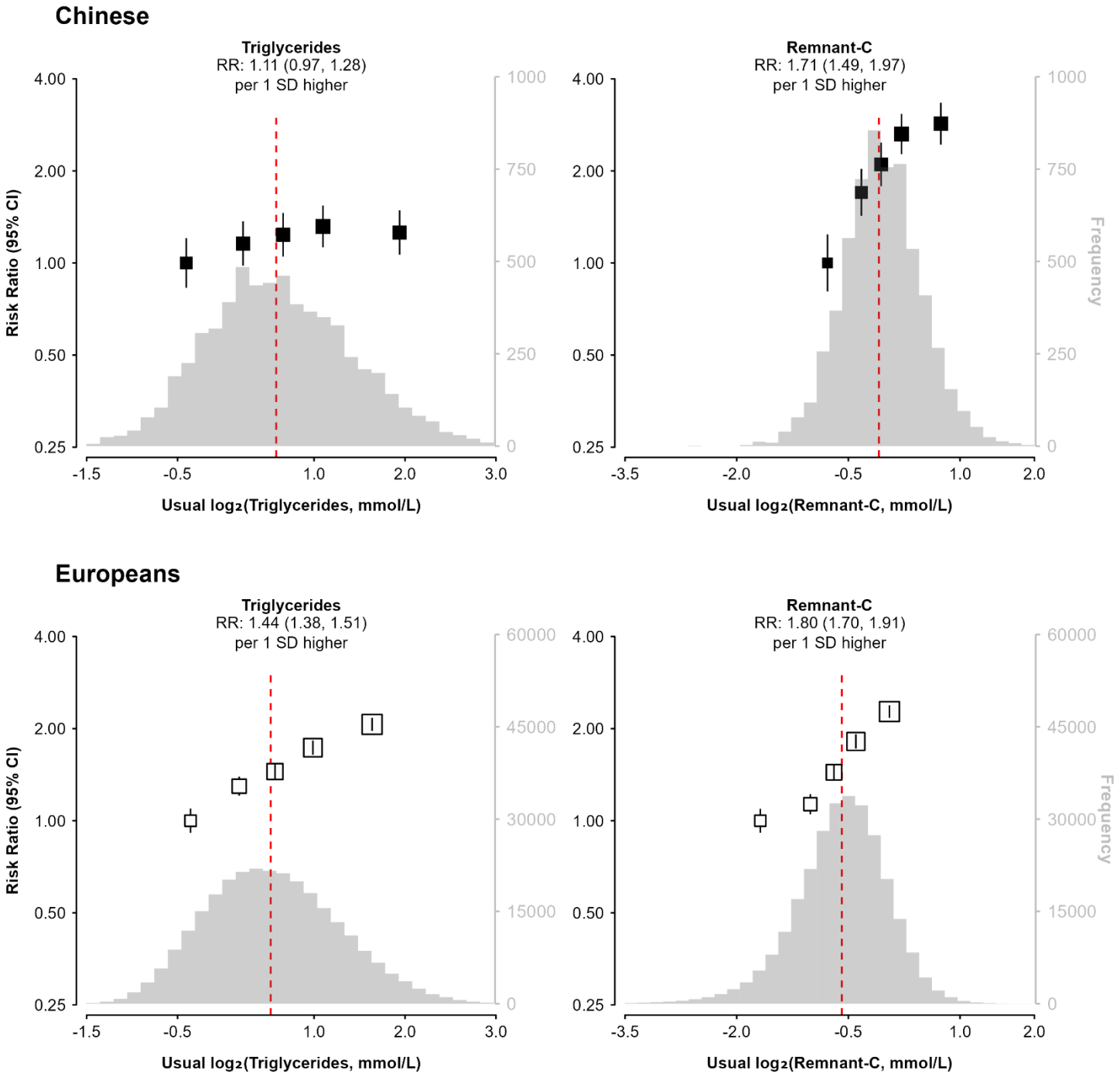
In the Chinese population, basic adjustment models adjusted for age, sex, region, BMI, diabetes, alcohol use, smoking status, SBP, physical activity, educational attainment, antihypertensive medications, mean temperature, and mean temperature-squared; In European populations, basic adjusted models were stratified by 5-yrs age-at-risk band and sex, and adjusted for BMI, diabetes, alcohol use, smoking status, SBP, physical activity, Townsend Deprivation Index, years of education, and antihypertensive medications. The $\chi 2$ values were computed from likelihood ratio tests with 1 degree of freedom. Triglycerides and Remnant-C levels were log-transformed.

***Further lipid adjustments** include adjusting for lipids in different classes: for LDL-related lipids, adjusting for Triglycerides and ApoA1; for Triglycerides-related lipids, adjusting for ApoB and ApoA1; for HDL-related lipids, adjusting for Triglycerides and ApoB.

4.3.3. Triglyceride-related lipid traits

Figure 4.7 shows that higher usual plasma levels of triglycerides were associated with a higher risk of MI in both populations, although the strength of the association appeared weaker in Chinese adults (RR [95% CI]: 1.11 [0.97–1.28] per 1-SD) than in Europeans (1.44 [1.38–1.51]), despite broadly similar mean concentrations in these populations. Remnant-C showed comparable RRs for 1-SD higher levels in Chinese and European participants (1.71 [1.49–1.97] vs 1.80 [1.70–1.91], respectively). In both populations, plasma levels of remnant-C were more strongly associated with risk of MI than plasma levels of triglycerides, with χ^2 values showing a 17-fold difference in Chinese (71.1 vs 4.2 for remnant-C and triglycerides, respectively) and approximately a twofold difference in Europeans (411.4 vs 237.9), as shown in ***Figure 4.6***. The associations of usual plasma levels of triglycerides with risk of MI were no longer significant in Chinese and were greatly attenuated in European populations after additional adjustment for ApoB and ApoA1. The χ^2 values for adding triglycerides and remnant-C under conventional adjustment in Chinese were 4.2 and 71.1, respectively, while the corresponding values in Europeans were 237.9 and 411.4. After further lipid adjustment, the χ^2 values for triglycerides decreased to 0.7 in Chinese and 23.3 in Europeans, while remnant-C remained significantly associated with risk of MI, with χ^2 values of 20.6 in Chinese and 25.9 in Europeans (***Figure 4.6***). Higher usual levels of remnant-C remained associated with risk of MI in both populations although the RRs of MI were attenuated after adjustment for the other lipids.

Figure 4.7. Adjusted RRs (95% CI) of MI for triglyceride-related measures in Chinese and Europeans



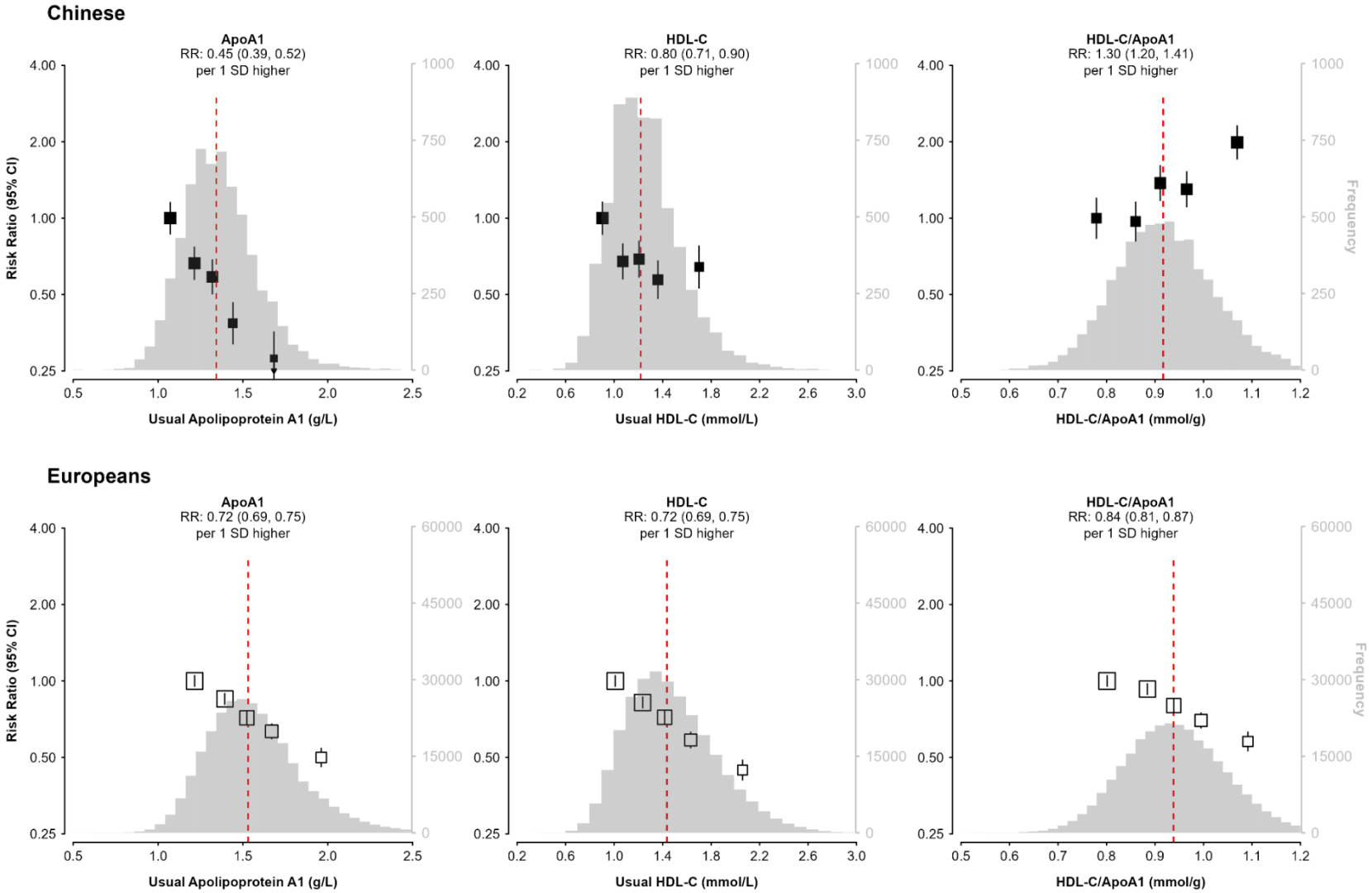
Abbreviations: **RRs**: Risk Ratios; **Remnant-C**: Remnant Cholesterol; **SD**: Standard Deviation.
In the Chinese population, the RRs were adjusted for age, sex, region, BMI, diabetes, alcohol use, smoking status, SBP, physical activity, educational attainment, antihypertensive medications, mean temperature, and mean temperature-squared. In European populations, the RRs were stratified by 5-yrs age-at-risk band and sex, and adjusted for BMI, diabetes, alcohol use, smoking status, SBP, physical activity, Townsend Deprivation Index, years of education, and antihypertensive medications. The distributions of blood lipid levels in CKB were obtained from baseline samples in controls only. The red dashed lines indicate median values. SDs are from baseline measurements and can be referred from Table 4.2.

4.3.4. HDL-C-related lipid traits

Figure 4.8 shows the associations of HDL-C-related lipid traits with risk of MI. Higher usual plasma levels of ApoA1 were inversely associated with risk of MI in both populations. The RRs (95% CI) per 1-SD higher levels were stronger in Chinese than in European populations (0.45; 0.39–0.52 vs 0.72; 0.69–0.75, respectively). In Europeans, the associations of usual plasma levels of HDL-C with risk of MI were comparable to those observed for ApoA1, whereas in Chinese adults, the inverse association between HDL-C and MI risk appeared to weaken at HDL-C concentrations above 1 mmol/L.

After conventional adjustment, the χ^2 values for ApoA1 and HDL-C were 134.3 and 13.6 in Chinese, and 237.5 and 264.8 in Europeans, respectively. After further lipid adjustment, the χ^2 values for ApoA1 and HDL-C were 134.1 and 18.3 in Chinese, and 162.8 and 185.7 in Europeans (***Figure 4.6***).

Figure 4.8. Adjusted RRs (95% CI) of MI for HDL-related measures in Chinese and Europeans



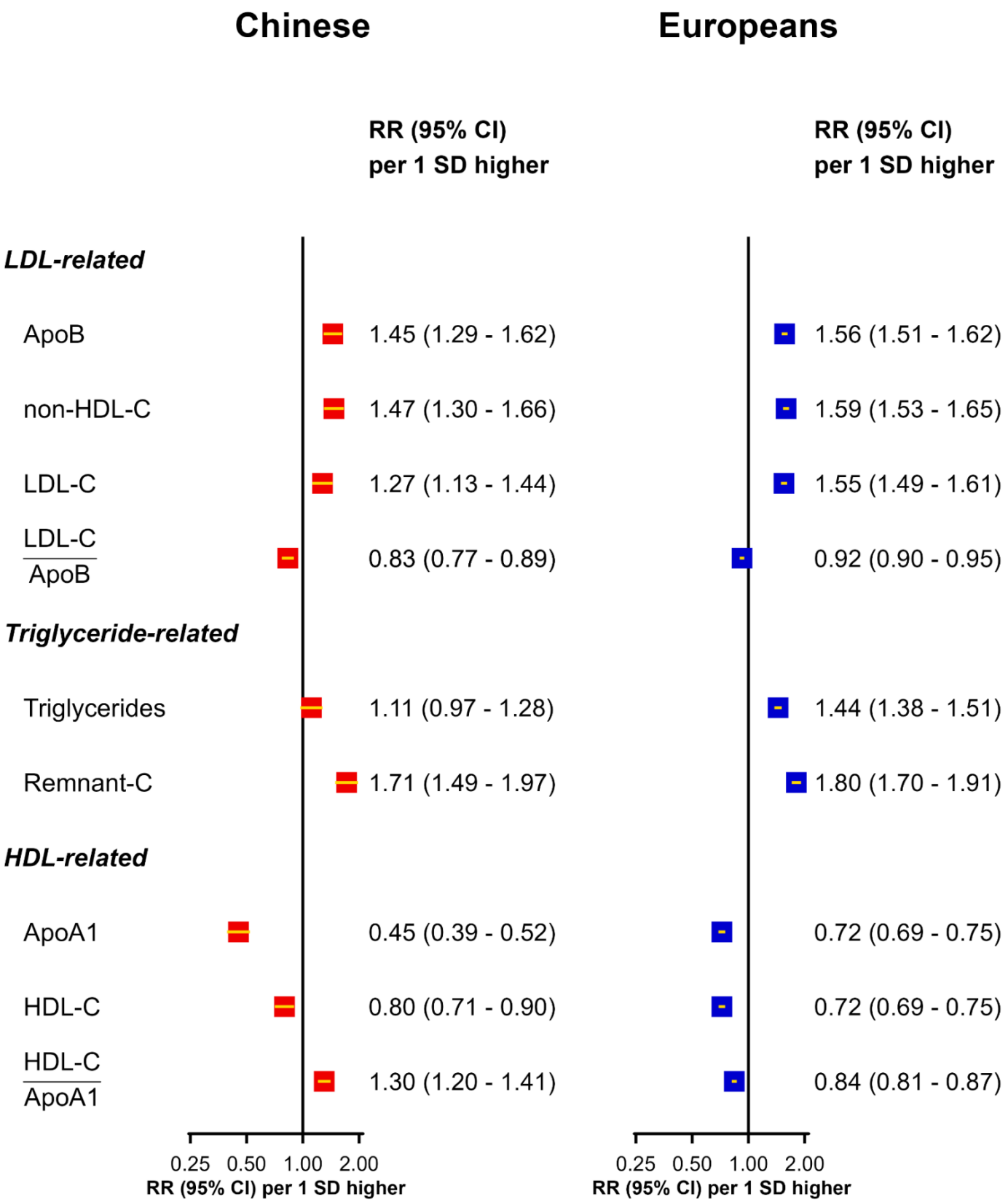
Abbreviations: **RRs**: Risk Ratios; **HDL-C:** High-Density Lipoprotein Cholesterol; **ApoA1**: Apolipoprotein A1; **SD**: Standard Deviation.
In CKB, the RRs were adjusted for age, sex, region, BMI, diabetes, alcohol use, smoking status, SBP, physical activity, educational attainment, antihypertensive medications, mean temperature, and mean temperature-squared. In UKB, the RRs were stratified by 5-yrs age-at-risk band and sex, and adjusted for BMI, diabetes, alcohol use, smoking status, SBP, physical activity, Townsend Deprivation Index, years of education, and antihypertensive medications. The distribution of blood lipids was obtained from measurements in baseline samples in controls only. The red dashed lines indicate the median values. SDs are from baseline measurements and can be found in Table 1.

4.3.5. Lipid ratios

Higher LDL-C/ApoB ratios were inversely associated with risk of MI in both populations, with stronger associations observed in Chinese compared with European participants (RR [95% CI]: 0.83 [0.77–0.89] vs 0.92 [0.90–0.95] per 1-SD higher level of LDL-C/ApoB ratios; ***Figure 4.6***). After additional adjustment for triglycerides and ApoA1, all LDL-related traits remained statistically significantly associated with MI in both populations (***Figure 4.6***). In the Chinese population, multivariable analyses assessing the joint associations of LDL-C and ApoB demonstrated that ApoB, but not LDL-C, was independently associated with risk of MI (***Table 4.4***). The RR per 1-SD higher plasma levels of ApoB was associated with 84% higher risk of MI in Ridge regression model, after adjustment for LDL-C (RR: 1.84, 95% CI: 1.60–2.12). Given plasma levels of ApoB, higher LDL-C concentrations were associated with lower rather than higher risks of MI (RR: 0.73, 95% CI: 0.62–0.86), as shown in ***Table 4.4***.

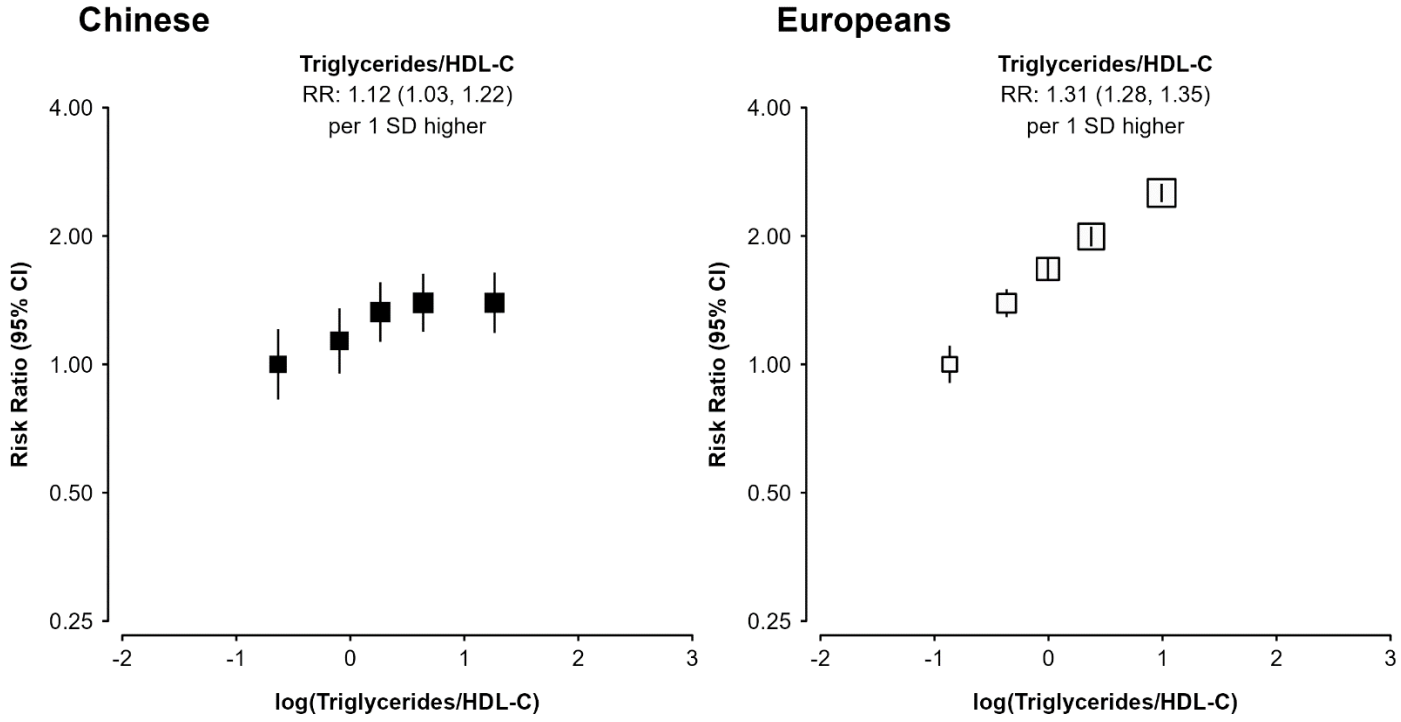
A higher HDL-C/ApoA1 ratio was positively associated with MI risk in the Chinese population (RR: 1.30, 95% CI: 1.20–1.41 per SD increase), whereas an inverse association was observed in Europeans (RR [95% CI]: 0.84 [0.81–0.87] per SD higher levels; ***Figure 4.9***). The associations of all HDL-related traits with risk of MI remained largely unchanged after additional adjustment for ApoB and triglycerides (***Figure 4.6***). Further analyses examining the shape and strength of associations between the triglyceride/HDL-C ratio and MI showed only modest relationships, which were weaker than those for ApoB or remnant-C in both Chinese and European populations (***Figure 4.10***).

Figure 4.9. Adjusted RRs (95% CI) of MI in Chinese and Europeans



Abbreviations: **RRs**: Risk Ratios; **LR**: Likelihood Ratio; **ApoB**: Apolipoprotein B; **LDL-C**: Low-Density Lipoprotein Cholesterol; **non-HDL-C**: non-High-Density Lipoprotein Cholesterol; **Remnant-C**: Remnant Cholesterol; **HDL-C**: High-Density Lipoprotein Cholesterol; **ApoA1**: Apolipoprotein A1; **SD**: Standard Deviation.
In the Chinese population, the RRs adjusted for age, sex, region, BMI, diabetes, alcohol use, smoking status, SBP, physical activity, educational attainment, antihypertensive medications, mean temperature, and mean temperature-squared. In European populations, the RRs were stratified by 5-yrs age-at-risk band and sex, and adjusted for BMI, diabetes, alcohol use, smoking status, SBP, physical activity, Townsend Deprivation Index, years of education, and antihypertensive medications. Triglycerides and Remnant-C levels were log-transformed.

Figure 4.10. Risk ratios (95%CI) of MI for ratio of triglycerides to HDL-C in Chinese and Europeans



Abbreviations. **HDL-C**: High-Density Lipoprotein Cholesterol; **RR**: Risk Ratio; **SD**: Standard Deviation.
In the Chinese population, models adjusted for age, sex, region, BMI, diabetes, alcohol use, smoking status, SBP, physical activity, educational attainment, antihypertensive medications, mean temperature, and mean temperature-squared; In European populations, models were stratified by 5-yrs age-at-risk band and sex, and adjusted for BMI, diabetes, alcohol use, smoking status, SBP, physical activity, Townsend Deprivation Index, years of education, and antihypertensive medications.

Table 4.4. Multivariable risk ratios (95%CI) of MI after adjustment for additional lipids or apolipoproteins

	China Kadoorie Biobank	UK Biobank
Exposure	RR (95% CI) per 1-SD higher	RR (95% CI) per 1-SD higher
ApoB | Adjust for LDL-C	1.84 (1.60 - 2.12)	1.46 (1.37 - 1.55)
LDL-C | Adjust for ApoB	0.73 (0.62 - 0.86)	1.08 (1.01 - 1.15)
ApoA1 | Adjust for HDL-C	0.23 (0.20 - 0.27)	0.89 (0.84 - 0.94)
HDL-C | Adjust for ApoA1	2.31 (1.98 - 2.70)	0.80 (0.76 - 0.84)

Abbreviations. RRs: Risk Ratios; **LDL-C**: Low-Density Lipoprotein Cholesterol; **HDL-C**: High-Density Lipoprotein Cholesterol; **ApoB**: Apolipoprotein B; **ApoA1**: Apolipoprotein A1

Analyses were conducted by Ridge regression. In the Chinese population, RRs were adjusted for age, sex, region, BMI, diabetes, alcohol use, smoking status, SBP, physical activity, educational attainment, antihypertensive medications, mean temperature, and mean temperature-squared; In European populations, RRs adjusted for age, sex, BMI, diabetes, alcohol use, smoking status, SBP, physical activity, Townsend Deprivation Index, years of education, and antihypertensive medications. | refers to further adjustment.

4.4. Discussion

This chapter showed that mean plasma concentrations of LDL-C-related traits were approximately one-third lower in Chinese participants compared with Europeans. Among these traits, the association between LDL-C and MI risk was about half as strong in Chinese as in European populations. By contrast, the strengths of the associations of ApoB and of non-HDL-C with risk of MI were similar between the two populations. The association of plasma levels of remnant-C with risk of MI was stronger compared with triglycerides in both Chinese and European populations. The association of plasma levels of ApoA1 with risk of MI was stronger compared with HDL-C in the Chinese population but was of similar strength in European populations. In both Chinese and European populations, there was an inverse association of LDL particle size with risk of MI, and a positive association of small, dense LDL particles with risk of MI. By contrast, the association of HDL particle size with risk of MI differed by population, being positive in Chinese participants and inverse in Europeans.

Mean concentrations of LDL-C, non-HDL-C, and ApoB were approximately 1-SD or 30% lower in Chinese than in Europeans. Although LDL-C levels were lower and its association with risk of MI was weaker in Chinese adults, the shape and strengths of the associations of non-HDL-C and ApoB with risk of MI were broadly comparable between the two populations. The positive log-linear associations of non-HDL-C with risk of MI, but not those of LDL-C, are consistent with earlier reports based on a smaller number of MI cases in Chinese participants with low LDL-C concentrations.[119] These findings are also concordant with the CTTC meta-analysis, which showed that the proportional reduction in cardiovascular risk from LDL-C-lowering therapy was independent of baseline LDL-C levels.[120] The CTTC meta-

analysis estimated that a 1 mmol/L lower LDL-C was associated with about 24% lower risk of major coronary events, which is concordant with the findings of the present study of a 24% lower MI risk per 1 mmol/L lower usual plasma levels of LDL-C in Chinese adults. The present study demonstrated comparable strengths of associations of ApoB and of non-HDL-C with risk of MI in Chinese and Europeans, which reinforce the trans-ethnic relevance of lowering LDL-C for prevention of MI.[121,122] As more than 80% of participants in CTTC were of European ancestry and fewer than 10% were of Asian ancestry, direct trial evidence in East Asian populations remains limited. Nonetheless, the concordant associations of ApoB and non-HDL-C with risk of MI in UKB and CKB suggest that findings from trials conducted predominantly in European populations are broadly applicable to East Asian populations, although the present chapter is based on observational evidence and cannot establish causality. The weaker association of plasma levels of LDL-C with risk of MI in Chinese may reflect greater residual confounding or reverse causality, indicating that non-HDL-C and ApoB are more reliable markers of atherogenic risk than LDL-C.[13,107,120,123]

The stronger associations of remnant-C with risk of MI compared with triglycerides in both populations are consistent with previous studies in European populations indicating that remnant-C may represent a more important target for residual lipid-related risk beyond LDL-C than triglycerides.[124,125] Moreover, the stronger association of plasma levels of ApoA1 with risk of MI compared with HDL-C in Chinese suggests that apolipoproteins or non-HDL-C may provide more informative measures than cholesterol fractions.

The results indicate that LDL particles containing less cholesterol, corresponding to smaller and denser LDL particles, were more strongly associated with risk of MI than larger, cholesterol-rich LDL particles. These observations align with prior evidence showing that, once ApoB is accounted for, the associations of LDL-C with risk of MI can become inverse.[106] In the Chinese population, a higher HDL-C to ApoA1 ratio was associated with a higher risk of MI, implying that HDL particles high in cholesterol or of larger size may contribute to increased MI risk.[106] This finding is supported by a prospective study in Korea showing a similar positive association of HDL-C/ApoA1 with risk of MI.[126] The divergent associations of HDL-C/ApoA1 between Chinese and Europeans may be partly explained by differences in alcohol intake or diabetes prevalence, which varied across tertiles of LDL-C/ApoB and HDL-C/ApoA1 ratios in UKB.

In summary, the findings of the present study underscore the importance of considering lipid measures beyond LDL-C, HDL-C, and triglycerides when assessing risk of MI in diverse populations where baseline lipid profile may differ. In particular, inclusion of ApoB, ApoA1, and remnant-C may better inform treatment decisions for lipid-modifying therapy and improve estimation of 10-year absolute risk of ASCVD, extending beyond the use of conventional cholesterol fractions alone.[127] The similar strengths of the associations of ApoB and non-HDL-C with risk of MI in the Chinese population compared with these associations in European populations provide support to existing ASCVD prevention guidelines that advocate statin initiation based on overall 10-year absolute cardiovascular risk rather than on levels of lipids.[35] These findings further emphasise the potential utility of ApoB as an alternative marker of atherogenic burden to LDL-C for risk assessment for MI, given that ApoB directly quantifies the number of atherogenic lipoprotein particles, whereas

cholesterol-based measures reflect the cholesterol content within particles classified by density and can be more affected by analytical variability.[35]

The present chapter also had several limitations. Despite the prospective design in UKB and exclusion of prior MI, stroke, and statin users in CKB, the possibility of reverse causality cannot be excluded. UKB and CKB collected extensive information on participants' characteristics as provided in Chapter 3. However, the possibility of residual confounding, or bias from unmeasured factors cannot be excluded.

To strengthen causal inference and reduce concerns about unmeasured confounding and reverse causality, complementary epidemiological approaches such as MR are required to triangulate the associations of plasma lipid traits with risk of MI as observed in this chapter. Accordingly, the following chapter (Chapter 5) will use an MR approach to investigate the causal relevance of major plasma lipid traits for risk of MI in Chinese and European populations to provide complementary evidence to the observational analyses.

Chapter 5 Causal relevance of plasma lipid traits with risk of MI

5.1. Background

The causal relevance of LDL-C in primary prevention of MI is well-established in European populations, and is supported by extensive evidence from large-scale observational studies and RCTs of lipid-modifying therapy.[9,42] Meta-analyses by the CTT Collaboration have demonstrated that each 1 mmol/L reduction in LDL-C was associated with an approximate 25% lower risk of major vascular events.[42] However, most of the included RCTs from our umbrella review were conducted in predominantly European-ancestry populations, and their generalisability to the Chinese population that has about 30% lower mean pre-treatment levels of LDL-C as shown in Chapter 4, remains uncertain.

This uncertainty is further demonstrated by the observational findings presented in Chapter 4 of this thesis, which found that the shape of the positive association of plasma LDL-C levels with risk of MI was non-linear in Chinese at lower plasma levels of LDL-C. The strength of the association of plasma levels of LDL-C with risk of MI was weaker in Chinese compared to European populations. However, the strength of associations for other atherogenic lipid measures, including ApoB and non-HDL-C, were comparable between both populations. These discrepancies raise questions about the selection of the most reliable lipid markers for MI risk assessment in Chinese adults. Given the lack of large-scale RCTs of statins for primary prevention of CVD in China, there is a clear evidence gap to guide clinical practice regarding the use of lipid-modifying medications for primary prevention of CVD.

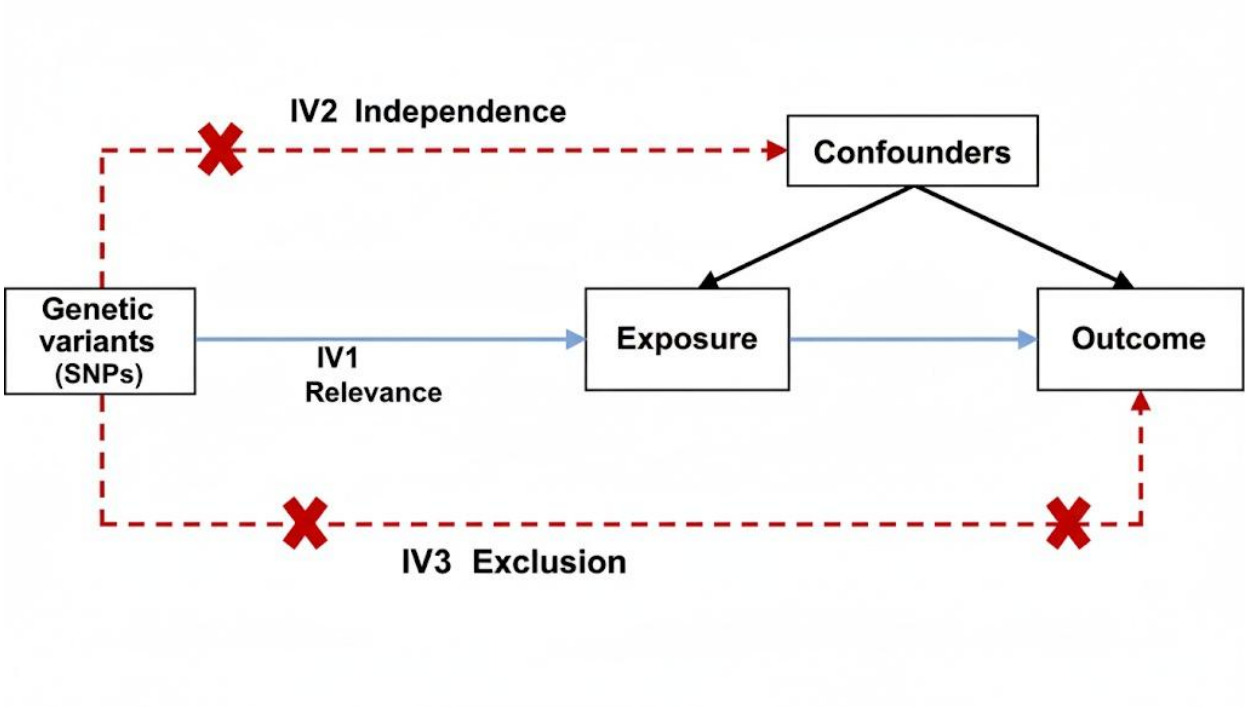
Beyond the traditional focus on LDL-C, significant residual risk for MI persists even in patients achieving target LDL-C levels. This has prompted interest in exploring the effects of other lipid fractions on risk of CVD.[44,106] In particular, elevated triglycerides, transported in triglyceride-rich lipoproteins, are associated with higher risk of MI and have emerged as important contributors to atherogenesis.[39,124] The independent causal role of plasma levels of triglycerides from LDL-C for risk of CVD has been debated. However, epidemiological and genetic studies increasingly suggest that triglyceride-rich lipoproteins contribute to the atherosclerotic process by promoting inflammation and foam cell formation within the arterial wall, and thus increase the risk of CVD.[39,44,124] The observational analyses in Chapter 4 demonstrated log-linear positive associations for higher plasma levels of triglycerides with higher risk of MI, highlighting the need to assess the causal relevance of plasma levels of triglycerides for risk of MI using genetic methods.

Observational studies have consistently demonstrated strong inverse associations between HDL-C levels and MI risk. However, the causal relevance of plasma levels of HDL-C with risk of MI has been challenged by previous MR studies and by the null results of trials of several HDL-C-raising treatments to reduce risk of cardiovascular events.[46,128] The evidence increasingly suggests that HDL-related particles and physiological processes, such as the reverse cholesterol transportation mechanism, as the causal protective factor, rather than HDL-C levels per se.[46,128] The inverse association of plasma levels of HDL-C with risk of MI observed in Chapter 4, which appeared to differ between Chinese and European populations, further underscores the need for more reliable analysis of the causal relevance of plasma levels of HDL-C for risk of MI in a trans-ethnic context.

MR offers a powerful approach to address challenges in observational epidemiology.[26] By using randomly allocated genetic variants as unconfounded proxies for lifelong differences in exposures, MR studies can provide support for causal relevance of these associations and overcome some of the limitations inherent in observational studies. As genetic variants are determined at conception, MR is less prone to reverse causality and can provide robust estimates of the long-term effects of lipid modification, serving to triangulate the observational evidence. The statistical framework of MR is based on instrumental variable (IV) analysis. An IV is a variable that is robustly associated with the exposure but not with the outcome, except through that exposure, and independent of confounders. In MR, genetic variants such as single-nucleotide polymorphisms (SNPs) can serve as IVs, providing a natural experiment that mimics randomisation in trials.

MR requires three core assumptions for the IVs (***Figure 5.1***): the genetic instrument must be relevant to the exposure (IV assumption 1: relevance), independent of confounders (IV assumption 2: independence), and affect the outcome only through the exposure (IV assumption 3: exclusion).[26,129] The relevance assumption means that the selected genetic variants are robustly associated with the exposure of interest, ensuring sufficient instrument strength. The independence assumption requires that the genetic variants are not associated with any confounding factors that influence both the exposure and the outcome, analogous to randomisation in a clinical trial. The exclusion restriction assumption requires that the genetic variants influence the outcome solely through their effect on the exposure and not through any alternative pathways, which would otherwise lead to horizontal pleiotropy.[26,129]

Figure 5.1. Mendelian randomisation and instrument variables



Abbreviations: **MR**, Mendelian randomisation; **IV**, instrumental variable; **SNP**, single-nucleotide polymorphism.

In MR studies, IV assumption 1 is examined using F statistics and R^2, which tests for instrument strength and the proportion of variance in the exposure explained by the IV, respectively.[130] A key challenge in studying plasma lipid traits and risk of MI is that genetic variants influencing one fraction (e.g., HDL-C) often have pleiotropic effects on others (e.g., LDL-C and triglycerides). Such pleiotropy can be vertical, where a variant affects multiple traits along the same biological pathway (e.g., via a causal chain), which is consistent with MR assumptions, or horizontal, where a variant influences the outcome via an independent pathway, violating the exclusion restriction (IV assumption 3) and biasing the results.[131]

One strategy to control for pleiotropy is to prune variants from a genetic risk score (GRS) that are associated with potential confounding pathways. A GRS is a composite measure that sums the effects of multiple genetic variants (SNPs), each weighted by their estimated association with a particular phenotype, to represent an individual's overall genetic predisposition to certain traits. It aggregates the cumulative contribution of effect alleles across loci into a single quantitative value. For instance, in investigating the causal role of HDL-C in CHD, an instrument can be "cleaned" by excluding SNPs that are also associated with LDL-C or triglycerides. This can be implemented by constructing both a "unrestricted" GRS, which includes all SNPs associated with the target lipid, and a "restricted" GRS, which removes SNPs that also show associations with other lipid fractions. The restricted approach increases the specificity of the instrument for the target lipid, though at the cost of reduced power.[40,129] The restricted GRS helps to mitigate both horizontal and vertical pleiotropy, as this approach effectively removes variants associated with non-target lipid traits. However, the primary aim is to address horizontal pleiotropy, since its presence violates the exclusion restriction assumption (IV3) of MR.

This chapter uses an MR approach to assess and compare the causal relevance of genetically-predicted levels of plasma lipid traits, including LDL-C, triglycerides, and HDL-C for risk of MI in Chinese and European populations.

5.2. Methods

5.2.1. Study populations and genotyping

Detailed description of the study cohorts, including the genetic subsets of CKB and UKB, have been previously reported in Chapter 3. The UKB study recruited about half a million individuals from 22 assessment centres throughout the UK for a prospective study and participants completed a baseline survey between 2006 and 2010.[101]

In UKB, analyses were restricted to White British participants (excluding n = 92,905 participants). This includes participants who reported as non-white British in the baseline questionnaire and those who are not in the White British ancestry subset in the genetic subset. We applied this restriction due to the ethnic differences in the SNPs associated with plasma lipid traits. Participants with a prior history of MI or stroke (n = 21,473) and those who reported use of statin therapy at baseline (n = 79,056) were excluded from the analyses. To make the analysis comparable across different plasma lipid traits, analyses were restricted to a subset with complete lipid measurements (excluding n = 43,386 participants with any missing lipid measurements).

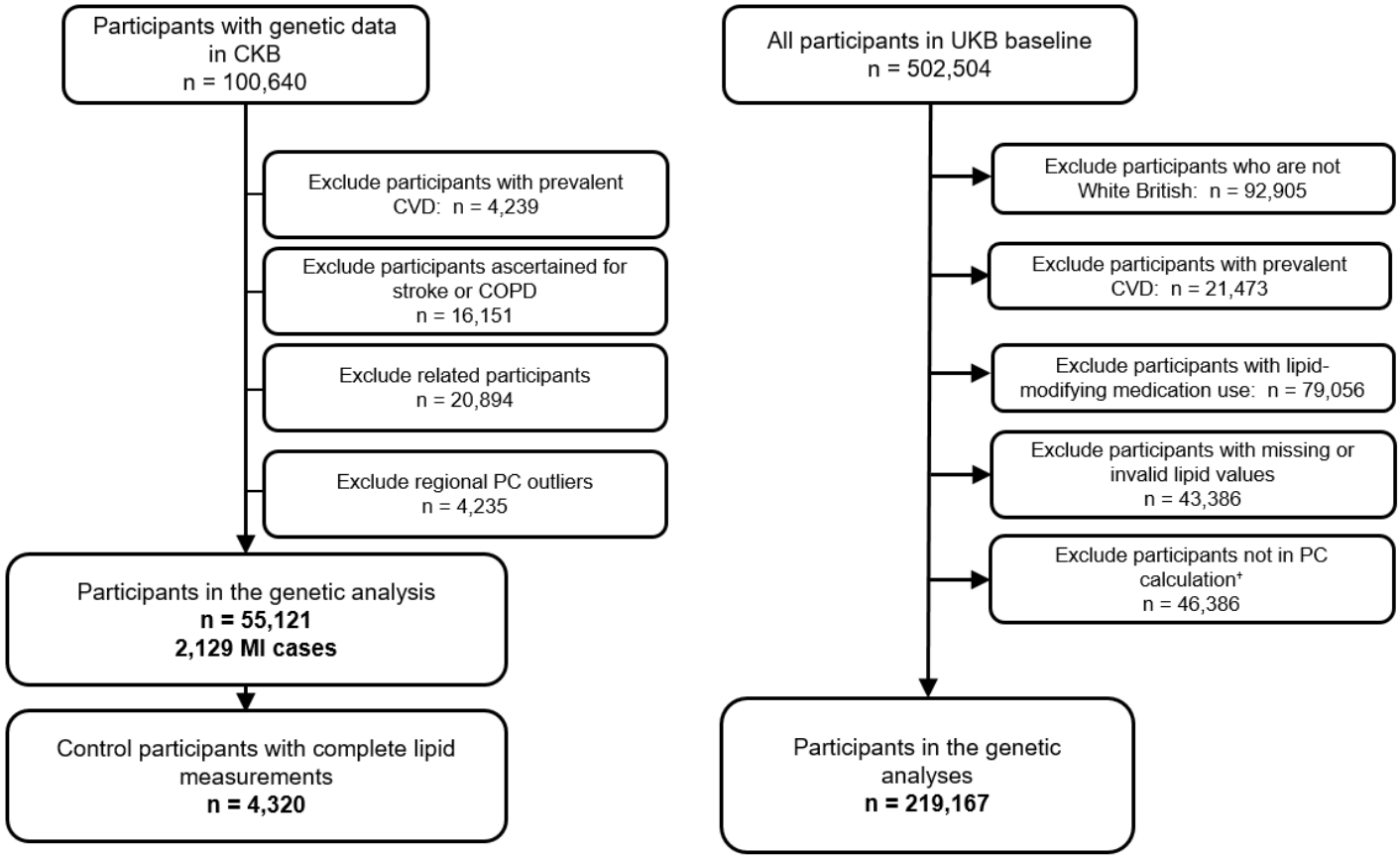
UKB performed internal quality control checks for inclusion in the calculation of PC and only those who had missing rate in autosomes ≤ 0.02, were not outliers for missingness or heterozygosity, in a maximal set of unrelated individuals, and not sex-discordant were included in the PC analyses.[103] Those who were not in the PC analyses were excluded (n = 46,386). Overall, a total of 219,167 participants were included in the genetic analyses in UKB.

The CKB study recruited 512,724 adults from 10 diverse areas (5 urban and 5 rural) in China between 2004 and 2008. A subset of the participants was selected for

genotyping (n = 100,640).[100] Participants with prevalent CVD (n = 4,329) or ascertained for stroke or COPD (n = 16,151) were excluded. Related participants (n = 20,894), defined as individuals with a second-degree or closer genetic relationship, and regional PC outliers (n = 4,235) were also excluded from the analysis. Overall, a total of 55,121 participants with 2,129 MI cases were included in the present analyses. Analyses of genetic association with plasma lipid traits were conducted in a nested case-control study with biochemical measurements (i.e., on 4,320 control individuals with complete data on plasma lipid measurements).

Further details and a flow-chart of the inclusion criteria are described in ***Figure 5.2.***

Figure 5.2. Inclusion criteria for CKB and UKB



†In UKB, participants in the PCA include those who have a missing rate in autosomes of ≤ 0.02; are not outliers for missingness or heterozygosity; are in a maximal set of unrelated individuals; are not sex-discordant.
Abbreviations. **MI**: Myocardial Infarction; **CVD**: Cardiovascular Disease; **BMI**: Body Mass Index; **PCA**: Principal Component Analysis; **CKB**: China Kadoorie Biobank; **UKB**: UK Biobank.

5.2.2. Instrumental variable selection and restriction

Variants related to LDL-C, triglycerides, and HDL-C were extracted from the Global Lipids Genetics Consortium (GLGC),[132] which included ancestry-specific data. GWAS data were downloaded from the GLGC website.[133] For European ancestry, GLGC provided data excluding UKB. Clumping was performed using PLINK LD clumping method through the TwoSampleMR package.[134] Within a specified window, SNPs in LD above an r^2 threshold are pruned, retaining the SNP with the smallest p value as the index SNP. The clumping window was set to 10,000 kb and the r^2 threshold to 0.001 to ensure high independence of instruments by capturing only the lead SNPs while pruning correlated variants over large genomic distances and minimising residual LD.[134,135]

For East Asian ancestry variant selection, GLGC included data from East Asian ancestry but did not provide a subset excluding CKB. Clumping was performed using PLINK LD clumping method, with the clumping window set to 10,000 kb, r-squared set to 0.001, and reference super-population panel set as East Asians.[134]

To minimise pleiotropic effects of the instruments, a restricted set of SNPs was selected. The restricted GRS mitigates both horizontal and vertical pleiotropy by removing variants associated with non-target lipid traits, but primarily addressing horizontal pleiotropy that violates the exclusion restriction assumption (IV3) of MR. The original, unrestricted SNPs were retrieved from GLGC. The restricted SNPs were chosen based on the following criteria:

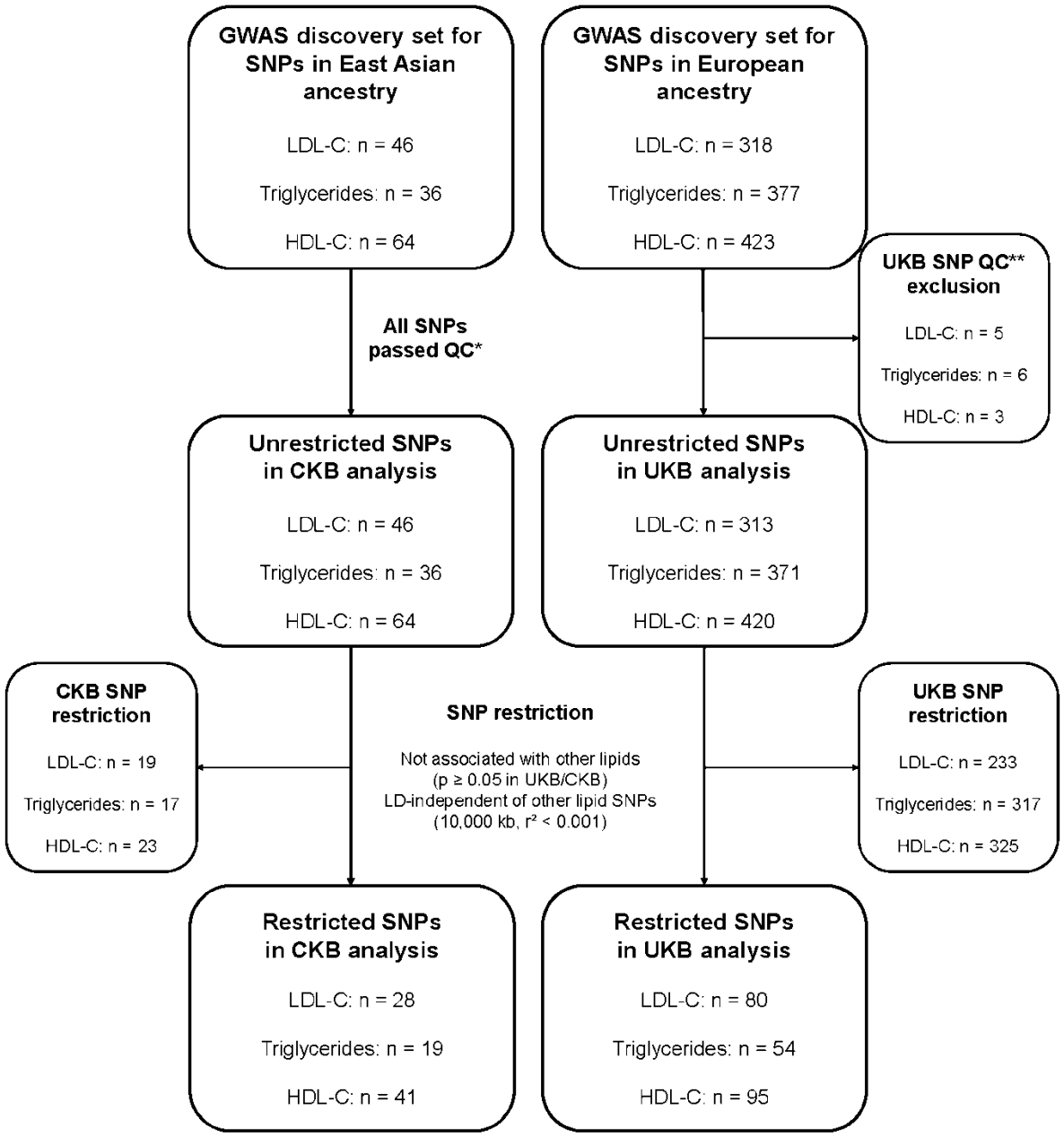
(1) they were associated with the target lipid trait at the GWAS significance level ($p < 5 \times 10^{-8}$).

(2) they showed no association with the other two lipid traits in either UKB or CKB datasets ($p \geq 0.05$).

(3) they were not in LD with SNPs significantly associated with the other two lipid traits at GWAS level (using a clumping window of 10,000 kb and an r^2 threshold of 0.001).

Variant selection, including the number of SNPs for each lipid trait, is presented in ***Figure 5.3.*** The list of SNPs used for each lipid trait is presented in *Supplementary materials:* ***Tables S1*** to ***S6***.

Figure 5.3. Instrument selection in Chinese and European populations



* CKB SNP QC excludes missing SNPs in CKB and minor allele frequency less than 0.005.
** UKB SNP QC excludes SNPs with imputation information score less than 0.4 and minor allele frequency less than 0.005.

5.2.3. Statistical methods

Before constructing the GRS, allele alignment was harmonised to ensure consistency between the effect and reference alleles in the gene-exposure and gene-outcome datasets. Specifically, effect alleles from the exposure GWAS were harmonised so that their direction corresponded to the same effect alleles in the gene-outcome database.

GRS was calculated with the formula as:

$$GRS = \sum_{1}^{n} \beta_i \times dosage_i$$

where β_i was the beta value from GWAS in GLGC.[132] $Dosage_i$ is the imputed dosage (ranging from 0 - 2) from UKB or CKB. Observed plasma levels were standardised (mean = 0 and SD = 1). A two-stage regression method was applied to estimate the causal effects. In UKB, the observed plasma lipid levels were regressed on the GRS to obtain genetically-predicted lipid levels, with adjustment for the ten PCs, age, age-squared, sex and genotyping array. The PCs represent axes of genetic variation derived from genome-wide data, capturing population structure and ancestry differences within the cohort to minimise confounding by genetic stratification.[103]

First stage:

$$Observed\ lipid\ levels = \beta_0 + \beta_1 GRS + \beta_2 Age + \beta_3 Age^2 + \beta_4 Sex + \beta_5 Array$$

$$+ \beta_{6-15} PC_{1-10} + \in$$

Where observed lipid levels are standardised lipid levels from biochemistry measurement.

The calculated βs were then used to calculate the genetically-predicted lipid levels for each individual ($\widehat{Lipid\ levels}$).

In the second stage, the predicted lipid levels were used as the exposure with MI as the outcome in a logistic regression model:

$$MI = \partial_0 + \partial_1 \widehat{Lipid\ levels} + \partial_2 Age + \partial_3 Age^2 + \partial_4 Sex + \partial_5 Array$$

$$+ \partial_{6-15} PC_{1-10} + e$$

The exponentiated causal estimate ∂_1 is interpreted as the OR for MI per 1-SD genetically-predicted lipid traits.

F statistics and R^2, which tests for instrument strength and the proportion of variance in the exposure explained by the instrumental variable, respectively, were calculated.[130] F statistics were derived from the first stage regression model of observed plasma lipid levels on GRS. For R^2 calculation, models were fitted with GRS, the first ten PCs, age, age-squared and sex (original model), and then a reduced model without GRS, and

$$R^2 = (R^2_{orginal\ model} - R^2_{reduced\ model})/(1 - R^2_{reduced\ model})$$

In CKB, a similar approach was used. Biochemistry measurements are only available in a nested case-control subset. Observed plasma lipid levels were regressed in this control subset on the GRS to obtain γ_{5-15} from the below formula.

First stage (in nested case-control subset):

$$Observed\ lipid\ levels = \gamma_0 + \gamma_1 GRS + \gamma_2 Age + \gamma_3 Age^2 + \gamma_4 Sex$$

$$+ \gamma_{5-15} PC_{1-11} + \in$$

The obtained γs were used to calculate genetically-predicted lipid levels in the CKB genetic subset. Genetically-predicted lipid levels were used as exposures with MI as the outcome in a logistic regression model, with adjustment for age, age-squared, sex, study regions, and 11 national PCs. The national PCs capture large-scale genetic structure across the ten recruitment regions of CKB, reflecting differences in ancestry and genetic variation that could otherwise confound genetic associations.[100]

Second stage (in CKB genetic subset):

$$MI = \delta_0 + \delta_1 \widehat{Lipid\ levels} + \delta_2 Age + \delta_3 Age^2 + \delta_4 Sex + \delta_{5-15} PC_{1-11} + e$$

The exponentiated causal estimate δ_1 is interpreted as the OR for MI per 1-SD genetically-predicted lipid traits in CKB.

All statistical analyses were performed using R version 4.2.2.

5.2.4. Sensitivity analyses

To assess the robustness of the primary MR findings and to test for potential violations of the MR assumptions, particularly from horizontal pleiotropy, a series of sensitivity analyses were performed.

First, Multivariable MR (MVMR) was conducted as a supporting analysis rather than the main approach. The primary strategy in this thesis relied on the restricted GRS framework to address horizontal pleiotropy, and MVMR therefore provides complementary evidence by modelling correlated lipid traits directly rather than relying solely on pruning SNPs. MVMR can estimate the effect of each exposure conditioned on the other exposures in the model.[136] With MVMR, the estimand is the direct effect without confounding or mediation. In an MVMR model, SNPs are associated with at least one of the exposures in the model. For an MVMR model involving LDL-C, HDL-C, and triglycerides, SNPs associated with any of the exposures were selected and further clumped.[136]

Second, MR-Egger regression was conducted.[137] The inverse-variance weighted (IVW) method combines the ratio estimates from each genetic variant in a meta-analysis framework, weighting them by the inverse of their variance.[138] The IVW approach constrains the intercept of the regression of the genetic associations with the outcome on the genetic associations with the exposure to zero. Unlike the standard IVW method, MR-Egger allows the intercept to be freely estimated. MR-Egger provides a pleiotropy-adjusted causal estimate under the Instrument Strength Independent of Direct Effect (InSIDE) assumption, which requires that the strength of each instrument's association with the exposure is independent of its direct pleiotropic effect on the outcome. Although MR-Egger is more robust to directional pleiotropy, it typically has lower precision and wider confidence intervals than IVW,

and the interpretation of the intercept is generally more informative than that of the slope.

Third, the weighted median method was applied.[138] This approach provides a consistent causal estimate even if up to 50% of the weight in the analysis comes from invalid (i.e., pleiotropic) genetic instruments. It is therefore more robust to the influence of outlying or pleiotropic SNPs than the standard IVW model. This estimator complements MR-Egger by not relying on the InSIDE assumption, and is robust in scenarios where pleiotropy is widespread.

Fourth, the mode-based estimator (MR-Mode) was used.[139] This method is robust to pleiotropy by identifying the most frequent or modal causal effect estimate among the individual instruments, thereby reducing the influence of SNPs that may have pleiotropic effects different from the mode. The mode-based estimator provides an additional check, as it performs well when the largest cluster of SNPs reflects the true causal effect.

In addition to these methods, an analysis using an unweighted GRS was performed, where all risk-factor-increasing alleles were assigned an equal weight of 1. This approach serves as a sensitivity analysis to assess the robustness of findings to potential bias from external GWAS effect estimates, ensuring that results are not driven by weighting.[140]

Finally, the Wald ratio for each individual SNP was calculated, and these were combined in a fixed-effect meta-analysis to provide an overall estimate. This approach pools SNP-specific estimates to obtain an overall effect, assuming a common true effect across instruments under the assumption of no heterogeneity.[26,138]

Consistency of the causal estimates across these different analytical methods could increase confidence in the primary findings. MR-Egger provides a pleiotropy-adjusted estimate under the InSIDE assumption and primarily tests violations of the exclusion restriction assumption (IV assumption 3) by allowing for horizontal pleiotropy. The weighted median estimator yields consistent estimates even when up to half of the instruments are invalid, thereby offering protection against violations of both the independence and exclusion assumptions (IV assumptions 2 and 3). The mode-based estimator identifies the most frequent causal effect across variants, reducing sensitivity to outlying or pleiotropic SNPs. The unweighted GRS and SNP-specific Wald ratio analyses assess robustness to potential bias in instrument weighting and evaluate the contribution of individual variants.[26,140]

Statistical power for the MR analyses was estimated using a noncentral χ^2 framework for binary outcomes, with calculations based on instrument strength (R^2), case fraction (K/N), and a Type-I error rate (α) of 0.05.[141] The odds ratio of the outcome per 1-SD higher exposure was set to the expected MR effect size derived from previous large-scale studies. In East Asian populations, the umbrella review in Chapter 2 reported that the best available OR estimates for the association per 1-SD higher genetically-predicted levels of exposure with IHD were approximately 1.77 for LDL-C, 0.89 for HDL-C, and 1.44 for triglycerides.[58] In European populations, corresponding OR estimates for the association per 1-SD higher genetically-predicted levels of exposure with IHD from UKB and GLGC data were approximately 1.45 for LDL-C, 1.25 for triglycerides, and 0.91 for HDL-C.[85] These values were used to inform the expected effect sizes in the power calculations. The power estimation aimed to assess whether the MR analyses were adequately powered to detect plausible causal effects. This approach helps interpret non-significant findings by

distinguishing whether they are more likely due to limited statistical power or a true absence of causal effect. Power calculations were implemented through mRnd: Power calculations for Mendelian Randomization.[141]

5.3. Results

5.3.1. Genetic risk score and lipid traits

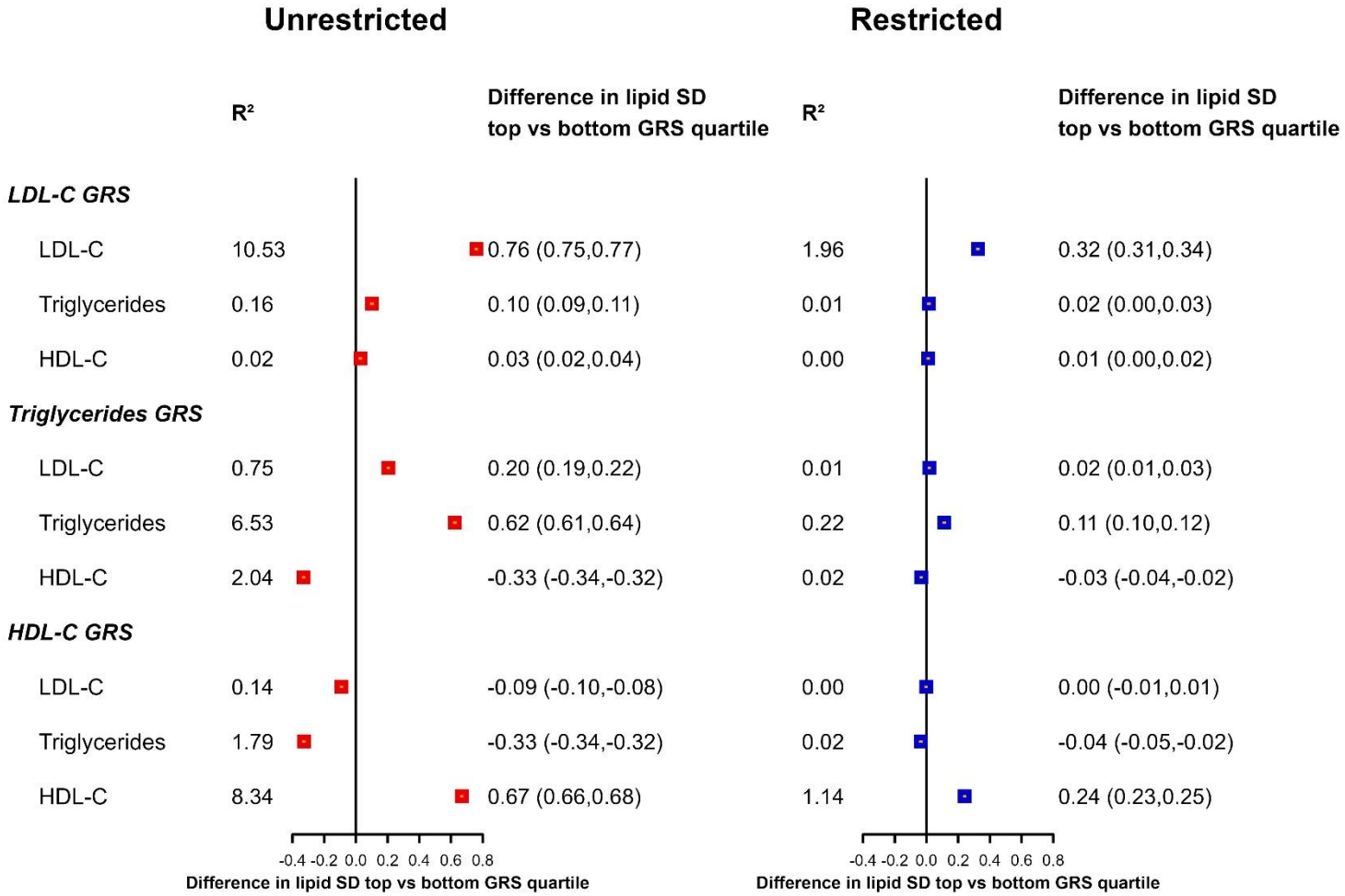
The GRS constructed for each lipid trait demonstrated strong associations with their corresponding observed lipid measures in both the CKB and UKB cohorts, as shown in ***Figures 5.4*** and ***5.5***. The use of both unrestricted and restricted GRS allowed for an assessment of the impact of potential pleiotropy. The restricted scores, designed to be more specific to a single lipid trait, explained a smaller proportion of the variance but showed substantially weaker associations with the other lipid traits as required.

In the European population (***Figure 5.4***), the unrestricted LDL-C GRS explained 10.53% of the variance in plasma levels of LDL-C, with individuals in the top quartile of the GRS having 0.76 SD (95% CI 0.75 – 0.77) higher mean plasma levels of LDL-C than those in the bottom quartile. However, this GRS was also associated with plasma levels of triglycerides and HDL-C. In contrast, the restricted LDL-C GRS explained only 1.96% of the variance in plasma levels of LDL-C (0.32 SD difference, 95% CI 0.31 – 0.34) but showed minimal association with plasma levels of triglycerides or HDL-C. Similarly, the unrestricted HDL-C GRS explained 8.34% of the variance in plasma levels of HDL-C but was also strongly associated with plasma levels of triglycerides (-0.33 SD). In contrast, the restricted HDL-C GRS explained only 1.14% of the variance in plasma levels of HDL-C but largely eliminated the pleiotropic effect on plasma levels of triglycerides (-0.04 SD).

In the Chinese population (***Figure 5.5***), a similar pattern was observed. The unrestricted LDL-C GRS explained 9.17% of the variance in plasma levels of LDL-C (0.70 SD difference, 95% CI 0.62–0.78), while the restricted score explained 6.55% of the variance in plasma levels of LDL-C (0.60 SD difference, 95% CI 0.53–0.68)

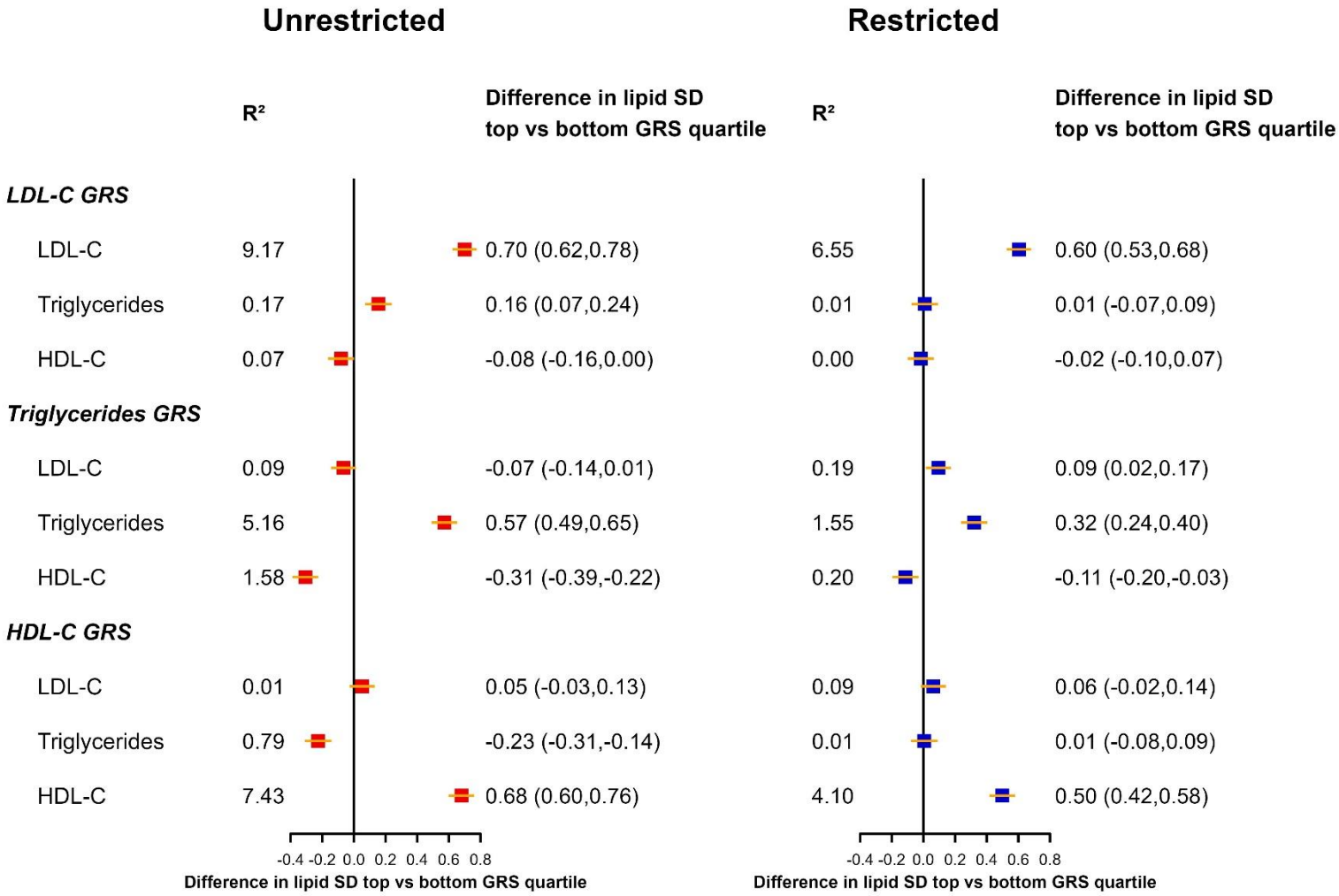
but minimised associations with plasma levels of triglycerides and HDL-C. For triglycerides, the unrestricted GRS explained 5.16% of the variance (0.57 SD difference, 95% CI 0.49–0.65) in plasma levels of triglycerides and was strongly associated with plasma levels of HDL-C (-0.31-SD). The restricted triglyceride GRS explained 1.55% of the variance in plasma levels of triglycerides (0.32 SD difference, 95% CI 0.24–0.40) and substantially reduced the pleiotropic associations with plasma levels of HDL-C (-0.11 SD).

Figure 5.4 Association of unrestricted and restricted GRS with lipid traits in European populations



Differences refer to the mean differences in each lipid trait between the top and bottom GRS quartiles, expressed in standard deviation (SD) units of the respective lipid distribution.
Analyses were adjusted for age, age-squared, sex, genotyping array, and the first ten principal components.
Abbreviations. **HDL-C**: High-Density Lipoprotein Cholesterol; **LDL-C**: Low-Density Lipoprotein Cholesterol; **SD**: Standard Deviation; **GRS**: Genetic Risk Score

Figure 5.5. Association of unrestricted and restricted GRS with lipid traits in the Chinese population



Differences refer to the mean differences in each lipid trait between the top and bottom GRS quartiles, expressed in standard deviation (SD) units of the respective lipid distribution.
Analyses were adjusted for age, age-squared, sex, study regions, and 11 national PCs.
Abbreviations. **HDL-C**: High-Density Lipoprotein Cholesterol; **LDL-C**: Low-Density Lipoprotein Cholesterol; **SD**: Standard Deviation; **GRS**: Genetic Risk Score.

Power calculations for the MR analyses are summarised in ***Table 5.1***. Statistical power was high for analyses using unrestricted GRS, approaching 100% for LDL-C and triglycerides in both Chinese and European populations. Analyses for LDL-C using restricted GRS also retained power above 90% in both populations. In contrast, restricted analyses showed lower power for triglycerides, particularly in the Chinese population (38%) but also in Europeans (49%). Power was low for HDL-C in both populations, which could be attributed to the predicted OR being close to the null.

Table 5.1. Power calculation for unrestricted and restricted MR analyses

Exposures	R2	Predicted OR	Power
Chinese			
LDL-C			
Unrestricted	9.17	1.77	1.00
Restricted	6.18		1.00
Triglycerides			
Unrestricted	5.16	1.44	0.99
Restricted	0.74		0.38
HDL-C			
Unrestricted	7.43	0.89	0.28
Restricted	3.95		0.17
Europeans			
LDL-C			
Unrestricted	10.53	1.45	1.00
Restricted	1.60		0.94
Triglycerides			
Unrestricted	6.53	1.25	0.98
Restricted	1.54		0.49
HDL-C			
Unrestricted	8.34	0.91	0.37
Restricted	1.14		0.09

In the Chinese population, the sample size is 55,121 with 2,129 MI cases. In European populations, the sample size is 219,167, with 3,979 MI cases.
Type-I error rate (α) = 0.05.
Abbreviations. **HDL-C**: High-Density Lipoprotein Cholesterol; **LDL-C**: Low-Density Lipoprotein Cholesterol; **MR**: Mendelian Randomisation**.**

5.3.2. Genetically-predicted LDL-C and MI risk

In the primary MR analyses, higher levels of genetically-predicted LDL-related traits were associated with a higher risk of MI (***Figure 5.6***). A 1-SD higher genetically-predicted plasma levels of LDL-C using the unrestricted GRS was associated with a 34% higher risk of MI (RR: 1.34, 95% CI 1.14 – 1.57) in the Chinese population and a 40% higher risk of MI (RR: 1.40, 95% CI 1.27 – 1.55) in European populations. Using the restricted GRS, a 1-SD higher genetically-predicted plasma level of LDL-C was associated with a 23% higher risk of MI (RR: 1.23, 95% CI 1.01 – 1.50) in the Chinese population and a 37% higher risk of MI (RR: 1.37, 95% CI 1.04 – 1.79) in European populations.

5.3.3. Genetically-predicted triglycerides and MI risk

Higher levels of genetically-predicted triglycerides were associated with a higher risk of MI (***Figure 5.6***). A 1-SD higher level of genetically-predicted triglycerides using the unrestricted GRS was associated with a 41% higher risk of MI (RR:1.41, 95% CI: 1.16–1.71) in the Chinese population and a 40% higher risk of MI (RR:1.41, 95% CI 1.24–1.60) in European populations. Using the restricted GRS, a 1-SD higher level of genetically-predicted triglycerides was associated with a 46% higher risk of MI (RR:1.46, 95% CI 0.86 – 2.46) in the Chinese population and an 87% higher risk of MI (RR:1.87, 95% CI 0.75 – 4.68) in European populations.

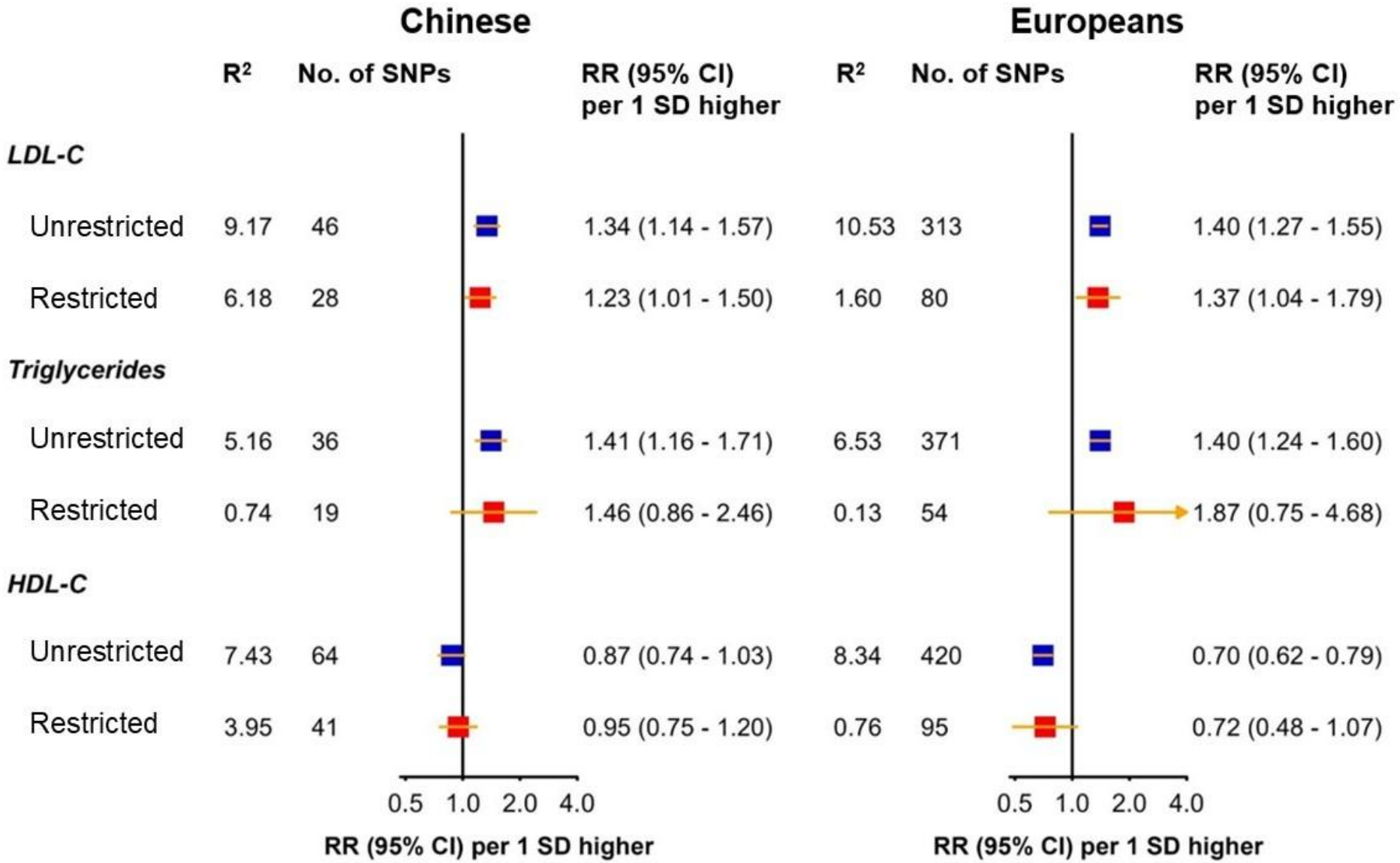
5.3.4. Genetically-predicted HDL-C and MI risk

Using unrestricted GRS for HDL-C, a 1-SD higher level of genetically-predicted HDL-C was associated with a 13% lower risk of MI (RR: 0.87, 95% CI 0.74 – 1.03) in the

Chinese population and a 30% lower risk of MI (RR: 0.70, 95% CI 0.62 – 0.79) in European populations (***Figure 5.6***).

Using restricted GRS for HDL-C, levels of genetically-predicted HDL-C were not significantly associated with risk of MI in either population. A 1-SD higher level of genetically-predicted HDL-C was associated with a 5% lower risk of MI (RR: 0.95, 95% CI 0.75 – 1.20) in the Chinese population and a 28% lower risk of MI (RR: 0.72, 95% CI 0.48 – 1.07) in European populations (***Figure 5.6***).

Figure 5.6. Genetically-predicted lipid traits and MI in Chinese and European populations



Analyses in European populations adjusted for age, age-squared, sex, genotyping array, the 10 principal components. Analyses in Chinese populations adjusted for age, age-squared, sex, study regions, and 11 national PCs.

Abbreviations. **LDL-C**: Low-Density Lipoprotein Cholesterol; **HDL-C**: High-Density Lipoprotein Cholesterol; **MI**: Myocardial Infarction; **RR**: Risk Ratio; **CI**: Confidence Interval; **SD**: Standard Deviation; **SNP**: Single Nucleotide Polymorphism; **R²**: Proportion of variance explained; **PCs**: Principal Components.

5.3.5. Sensitivity Analyses

The results from the MVMR analyses are provided in ***Table 5.2***. In the Chinese population, the MVMR-IVW OR was 1.40 (95% CI 1.15–1.70) for the association of a 1-SD higher level of genetically-predicted LDL-C and risk of MI. For triglycerides, the MVMR-IVW OR was 1.40 (95% CI 1.14–1.72) for the association of a 1-SD higher level of genetically-predicted triglycerides and risk of MI, while for HDL-C the MVMR-IVW OR was 0.97 (95% CI 0.79–1.18). In the European population, the corresponding MVMR-IVW ORs for risk of MI were 1.42 (95% CI 1.22–1.65) for LDL-C, 1.20 (95% CI 1.01–1.42) for triglycerides, and 0.89 (95% CI 0.75–1.05) for HDL-C.

The results from the primary MR-IVW analyses were compared with those from other MR methods to assess the robustness of the findings to potential pleiotropy, with the results provided in ***Table 5.3***.

In the Chinese population, for the unrestricted LDL-C GRS, the MR-IVW OR was 1.34 (95% CI 1.14–1.57) for the association of a 1-SD higher level of genetically-predicted LDL-C and risk of MI. The MR-Mode estimate for the same association was OR: 1.23 (95% CI 0.97–1.56), the MR-Median estimate for the association of 1-SD higher level of genetically-predicted LDL-C and risk of MI was OR: 1.24 (95% CI 0.97–1.60), and the MR-Egger estimate for the same association was OR: 1.38 (95% CI 1.06–1.78), with MR-Egger intercept significantly different from zero (p-value = 0.02). For the restricted LDL-C GRS, the MR-IVW OR for the association of 1-SD higher level of genetically-predicted LDL-C and risk of MI was 1.25 (95% CI 1.02–1.53), and the MR-Egger intercept p-value was not significantly different from zero (p-value = 0.58). For triglycerides, the unrestricted GRS yielded an MR-IVW OR of 1.41 (95% CI 1.16–1.71) for the association of a 1-SD higher level of genetically-predicted triglycerides and risk of MI, with an MR-Egger intercept was not significantly different from zero (p-value = 0.31). For HDL-C, the unrestricted GRS resulted in an MR-IVW OR of 0.87 (95% CI 0.74–1.03) for a 1-SD higher level of

genetically-predicted LDL-C and risk of MI, with an MR-Egger intercept not significantly different from zero (p-value = 0.51).

In European populations, for the unrestricted LDL-C GRS, the MR-IVW OR was 1.47 (95% CI 1.32–1.64) for the association of 1-SD higher level of genetically-predicted LDL-C and risk of MI. The MR-Mode estimate for the same association was OR: 1.32 (95% CI 1.09–1.60), the MR-Median estimate was OR: 1.27 (95% CI 1.05–1.53), and the MR-Egger estimate was OR: 1.61 (95% CI 1.32–1.96), with an MR-Egger intercept not significantly different from zero (p-value = 0.23). For the restricted LDL-C GRS, the MR-IVW OR was 1.38 (95% CI 0.99–1.93) for the association of 1-SD higher level of genetically-predicted LDL-C and risk of MI, and the MR-Egger intercept was not significantly different from zero (p-value = 0.22). For triglycerides, the unrestricted GRS yielded an MR-IVW OR of 1.38 (95% CI 1.22–1.56) for the association of 1-SD higher level of genetically-predicted triglycerides and risk of MI, with an MR-Egger intercept not significantly different from zero (p-value = 0.80). For HDL-C, the unrestricted GRS resulted in an MR-IVW OR of 0.73 (95% CI 0.65–0.81) for the association of 1-SD higher level of genetically-predicted HDL-C and risk of MI, with an MR-Egger intercept not significantly different from zero (p-value = 0.16).

Table 5.2. Multivariable MR results for the association of lipid traits with MI in European and Chinese populations

	Europeans		**Chinese**	
Exposures	MVMR-IVW OR (95% CI)	P value	MVMR-IVW OR (95% CI)	P value
MVMR: LDL-C and Triglycerides, and HDL-C (503 SNPs)			98 SNPs	
LDL-C	1.42 (1.22 – 1.65)	< 0.01	1.40 (1.15 – 1.70)	< 0.01
Triglycerides	1.20 (1.01 – 1.42)	0.04	1.40 (1.14 – 1.72)	< 0.01
HDL-C	0.89 (0.75 – 1.05)	0.16	0.97 (0.79 – 1.18)	0.74

Analyses in European populations adjusted for age, age-squared, sex, genotyping array, the first ten principal components. Analyses in Chinese populations adjusted for age, age-squared, sex, study regions, and 11 national PCs.
Abbreviations. **SNPs**: Single Nucleotide Polymorphisms; **OR**: Odds Ratio; **HDL-C**: High-Density Lipoprotein Cholesterol; **LDL-C**: Low-Density Lipoprotein Cholesterol; **IVW**: Inverse Variance-Weighted.

Table 5.3. MR Sensitivity analyses for lipid traits with MI in Chinese and European populations

Exposures	MR-IVW		MR-Mode (Weighted)		MR-Median (Weighted)		MR-Egger	
	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P Egger intercept
Chinese								
LDL-C								
Unrestricted	1.34 (1.14 – 1.57)	< 0.01	1.23 (0.97 – 1.56)	0.10	1.24 (0.97 – 1.60)	0.09	1.38 (1.06 – 1.78)	0.02
Restricted	1.25 (1.02 – 1.53)	0.03	1.23 (0.95 – 1.59)	0.13	1.21 (0.92 – 1.59)	0.17	1.32 (0.99 – 1.76)	0.58
Triglycerides								
Unrestricted	1.41 (1.16 – 1.71)	< 0.01	1.27 (1.02 – 1.58)	0.04	1.17 (0.88 – 1.54)	0.28	1.19 (0.85 – 1.65)	0.31
Restricted	1.37 (0.67 – 2.81)	0.38	0.54 (0.23 – 1.26)	0.18	1.12 (0.53 – 2.37)	0.78	0.34 (0.10 – 1.21)	0.03
HDL-C								
Unrestricted	0.87 (0.74 – 1.03)	0.12	0.99 (0.74 – 1.34)	0.24	1.00 (0.77 – 1.29)	0.98	0.91 (0.68 – 1.21)	0.51
Restricted	0.95 (0.77 – 1.18)	0.65	1.05 (0.77 – 1.42)	0.78	1.01 (0.74 – 1.39)	0.94	0.95 (0.65 – 1.38)	0.79
Europeans								
LDL-C								
Unrestricted	1.47 (1.32 – 1.64)	< 0.01	1.32 (1.09 – 1.60)	< 0.01	1.27 (1.05 – 1.53)	< 0.01	1.61 (1.32 – 1.96)	0.23
Restricted	1.38 (0.99 – 1.93)	0.06	1.48 (0.96 – 2.27)	0.08	1.47 (0.95 – 2.25)	0.08	1.79 (1.06 – 3.03)	0.22
Triglycerides								
Unrestricted	1.38 (1.22 – 1.56)	< 0.01	1.41 (1.14 – 1.74)	< 0.01	1.34 (1.08 – 1.67)	< 0.01	1.36 (1.11 – 1.66)	0.80
Restricted	1.54 (0.76 – 3.13)	0.23	0.88 (0.16 – 4.92)	0.89	1.68 (0.67 – 4.22)	0.27	2.25 (0.49 – 10.28)	0.58
HDL-C								
Unrestricted	0.73 (0.65 – 0.81)	< 0.01	0.81 (0.63 – 1.03)	0.09	0.79 (0.66 – 0.95)	0.01	0.80 (0.66 – 0.97)	0.16
Restricted	0.75 (0.50 – 1.14)	0.18	0.90 (0.43 – 1.88)	0.78	0.74 (0.40 – 1.37)	0.33	0.97 (0.50 – 1.87)	0.34

Analyses in European populations adjusted for age, age-squared, sex, genotyping array, the 10 principal components. Analyses in Chinese populations adjusted for age, age-squared, sex, study regions, and 11 national PCs.

Abbreviations. **SNPs**: Single Nucleotide Polymorphisms; **OR**: Odds Ratio; **HDL-C**: High-Density Lipoprotein Cholesterol; **LDL-C**: Low-Density Lipoprotein Cholesterol; **IVW**: Inverse Variance-Weighted.

5.4. Discussion

In this chapter, MR was used to assess the causal relevance of several plasma lipid traits for risk of MI in both Chinese and European ancestry populations. The primary findings from these analyses confirmed that higher levels of genetically-predicted LDL-C were causally associated with a higher risk of MI in Chinese as in European ancestry populations. Moreover, the strength of associations of higher levels of genetically-predicted LDL-C with risk of MI were similar in both ancestry populations. In contrast, there was no strong evidence of a causal protective association for higher levels of genetically-predicted HDL-C with MI risk in either population. Using a restricted GRS, levels of genetically-predicted triglyceride were not significantly associated with risk of MI in either ancestry population, although the analyses lacked power. These results were broadly consistent across a range of sensitivity analyses designed to account for potential pleiotropy and strengthen the causal relevance of the findings.

The present MR findings provide support for the causal relevance of LDL-C that could help to interpret the observational associations reported in Chapter 4. For LDL-C, the causal estimates from the MR analyses were consistent and significant in both Chinese and European ancestry populations, irrespective of whether an unrestricted or a more specific restricted GRS was used. This consistency strengthens the inference that plasma levels of LDL-C are causally associated with risk of MI in both ancestry populations. Furthermore, it suggests that the weaker observational associations from Chapter 4 for measured LDL-C with MI risk in Chinese compared to European populations is likely attributable to the biases inherent in observational studies, such as residual confounding or reverse causality, rather than a true ancestry difference in the underlying causal effect. The results provide support for

public health strategies and clinical practice on the need to initiate drug treatment based on overall CVD risk profile, rather than on levels of LDL-C,[121,142] as the association of LDL-C with risk of MI remained strong among Chinese with on average much lower pre-treatment levels of LDL-C compared with Europeans. The CTTC meta-analysis demonstrated that there is a consistent risk reduction in major vascular events with statin therapy per change in LDL-C in participants with plasma LDL-C levels as low as 1.6 mmol/L.[142] There is a marked disparity in statin prescription for primary prevention of CVD in China and Europe, due to concerns of lower average pre-treatment levels of LDL-C in China and the increased risk of haemorrhagic stroke with LDL-C lowering.[66,143,144] Previous analyses from CKB showed that although lower plasma levels of LDL-C were associated with an increased risk of intracerebral haemorrhage, the absolute number of ischaemic strokes prevented outweighed the excess haemorrhagic strokes, demonstrating a net absolute benefit of LDL-C–lowering therapy for stroke prevention, even in China where the incidence of haemorrhagic stroke is higher than in Western ancestry populations.[73] Given the consistent causal effect sizes for LDL-C from the current analyses, the findings support the potential value of wider use of statin and other LDL-lowering therapies in China for primary prevention of MI, although confirmation from RCT evidence in Chinese populations would be required before making clinical recommendations.

A key challenge in assessing the causal relevance of individual lipid fractions is the high correlation between them, which can result in horizontal pleiotropy in genetic analyses.[40,78] This was particularly evident in the analysis of plasma levels of triglycerides. Using the unrestricted GRS, which captures a broader genetic influence, higher levels of genetically-predicted triglycerides were significantly

associated with a higher risk of MI. However, this unrestricted score was also associated with plasma levels of LDL-C and HDL-C, making it difficult to isolate the independent effect of plasma levels of triglycerides. To address this, the analysis using the restricted GRS, which was constructed to be more specific to triglyceride pathways and minimise associations with other lipids, showed an attenuated and non-significant association of higher levels of genetically-predicted triglycerides with risk of MI, although the analyses lacked statistical power and thus the estimates should be interpreted with caution. However, the MVMR analyses, which modelled LDL-C, HDL-C, and triglycerides jointly, indicated that triglycerides retain an independent association with risk of MI in Chinese and European populations. It is noteworthy that, in the current analyses, the MVMR effect estimate for the association of genetically-predicted levels of triglycerides was larger in the Chinese population than in the European population. Given that the European population had stronger instruments (higher R^2 and F statistics) for both LDL-C and triglycerides, this pattern could indicate that once the mediation effect captured by atherogenic lipoproteins, such as those reflected by LDL-C, is fully accounted for, the direct effect of triglycerides becomes less apparent. This in turn may suggest that the triglyceride effect on risk of MI is mediated through LDL-C.[40,45,78,145]

The results for plasma levels of HDL-C are consistent with previous MR studies where a null association of levels of genetically-predicted HDL-C and risk of MI was observed.[46,128,145] In Europeans, the higher levels of genetically-predicted HDL-C using unrestricted GRS was associated with a lower risk of MI, but this GRS was also strongly associated with lower triglycerides. When the more specific restricted GRS for HDL-C was used, the association of higher levels of genetically-predicted HDL-C with risk of MI was attenuated to the null. In the Chinese population, there

was no significant association between levels of genetically-predicted HDL-C and risk of MI with either unrestricted or restricted GRSs, but these analyses are mostly underpowered. The MVMR analyses results were consistent with these findings, showing non-significant associations of genetically-predicted levels of HDL-C with risk of MI in both Chinese and European populations in the model including triglycerides, HDL-C, and LDL-C. Taken together, these findings suggest that the apparent protective effect of plasma levels of HDL-C is unlikely to be causal in either European or Chinese populations. This conclusion is consistent with previous large-scale MR studies and the null results of trials of several HDL-C-raising therapies to reduce cardiovascular events.[46]

This chapter also has several limitations. First, while efforts were made to reduce horizontal pleiotropy, it cannot be entirely excluded. The MR-Egger analyses showed some evidence of directional pleiotropy for the unrestricted LDL-C score in the Chinese analysis, although the main findings remained consistent in other sensitivity analyses that were more robust to such effects. Second, the restricted GRS, while more specific, explained less of the variance in their respective lipid traits. This approach led to the lack of statistical power to detect causal effects, particularly for triglycerides. Third, the genetic instruments for the Chinese population were derived from an East Asian GWAS, which included CKB participants, potentially leading to some bias from sample overlap; however, this bias is generally towards the confounded observational estimate in MR settings where the GRS associations with both the exposure and the outcome were conducted in the same database.[146,147] Finally, the MR analyses assess the effect of lifelong exposure to genetically altered lipid levels, and these effects may not be directly equivalent to the effects of pharmacological interventions initiated later in life.[148]

In conclusion, this chapter provides evidence for the causal relevance of plasma levels of LDL-C for the risk of MI in both Chinese and European populations, confirming that the proportional effects of LDL-C lowering are consistent across populations despite differences in baseline lipid levels. These results complement the observational findings presented in Chapter 4.

Having confirmed the causal relevance of LDL-C for MI, the next step in addressing the broader objectives of this thesis is to examine the second major risk factor in the multi-factorial intervention for prevention of MI, e.g., SBP, and its joint effects with LDL-C. Chapter 6 therefore focuses on the observational associations of SBP and LDL-C, both individually and in combination with LDL-C, with risk of MI. These analyses extend the observational findings of Chapter 4 by incorporating SBP. This provides a foundation for the factorial MR analyses presented in Chapter 7, which will assess the causal relevance of the combined effects of LDL-C and SBP for MI prevention, reflecting the overarching goal of this thesis to inform the potential for multi-factorial intervention in MI prevention.

Chapter 6 Observational associations of SBP and LDL-C with MI risk

6.1. Background

Elevated levels of SBP and LDL-C are among the most important modifiable risk factors for MI, and their causal relevance has been established by previous studies, including observational, genetic, and randomised clinical trial evidence,[36,42,51,74,79,87] as summarised in Chapter 2.

The design of multi-factorial interventions to reduce CVD risk, guided by the combination therapy concept, proposes that concurrently lowering levels of SBP and LDL-C could substantially reduce the burden of CVD.[21] One key assumption of this strategy is that the effects of lowering SBP and LDL-C are largely independent of one another, and therefore, the combined effects on CVD risk reduction would be expected to be multiplicative on a relative risk scale.[21]

However, several previous observational studies have questioned this assumption by reporting conflicting results of potential interactions between the effects of SBP and blood lipids on CVD risk. The APCSC,[79] a joint meta-analysis of cohorts from the Asia-Pacific region, reported that the strength of associations of usual SBP with CHD was more extreme in individuals with lower total cholesterol levels. The study reported that each 10 mmHg higher in usual levels of SBP was associated with a 34% higher risk of CHD in the lowest total cholesterol group (< 4.75 mmol/L) and a 21% higher risk of CHD in the highest cholesterol group (≥ 6.25 mmol/L). The observed results suggested an attenuation of the strength of association of LDL-C with risk of CHD at higher plasma cholesterol levels, and the study further reported a significant interaction of linear SBP levels and total cholesterol levels with risk of

CHD (P for the likelihood ratio test of interactions with linear SBP < 0.001). In contrast, analyses from the EPOCH-Japan consortium reported that the association of SBP with risk of CHD was stronger among individuals with high cholesterol levels.[54] They reported that per 1 SD higher levels of SBP (20.0 mmHg) was associated with a 24% higher risk of CHD in the lowest total cholesterol group (< 4.7 mmol/L) and a 52% higher risk of CHD in the highest total cholesterol group (≥ 5.7 mmol/L). They also reported a significant interaction in the strength of the associations of total cholesterol levels and SBP with risk of CHD, although the evidence for interaction is weak (P for interaction = 0.04). The discrepant findings of these studies have resulted in uncertainty about the combined observational associations of SBP and LDL-C with risk of MI, particularly across diverse populations with different mean levels of LDL-C and SBP.

Reliable estimates of the joint associations of SBP and LDL-C, and clarification of any potential interactions, are essential for optimising public health strategies and planning clinical trials on the joint effects of lowering both of these risk factors. Following the detailed analyses of the associations of individual lipid traits with MI in Chapter 4, this chapter aims to investigate the observational association of SBP with MI risk and to explore the combined effects of lowering SBP and LDL-related lipid traits on MI risk in both Chinese and European populations.

6.2. Methods

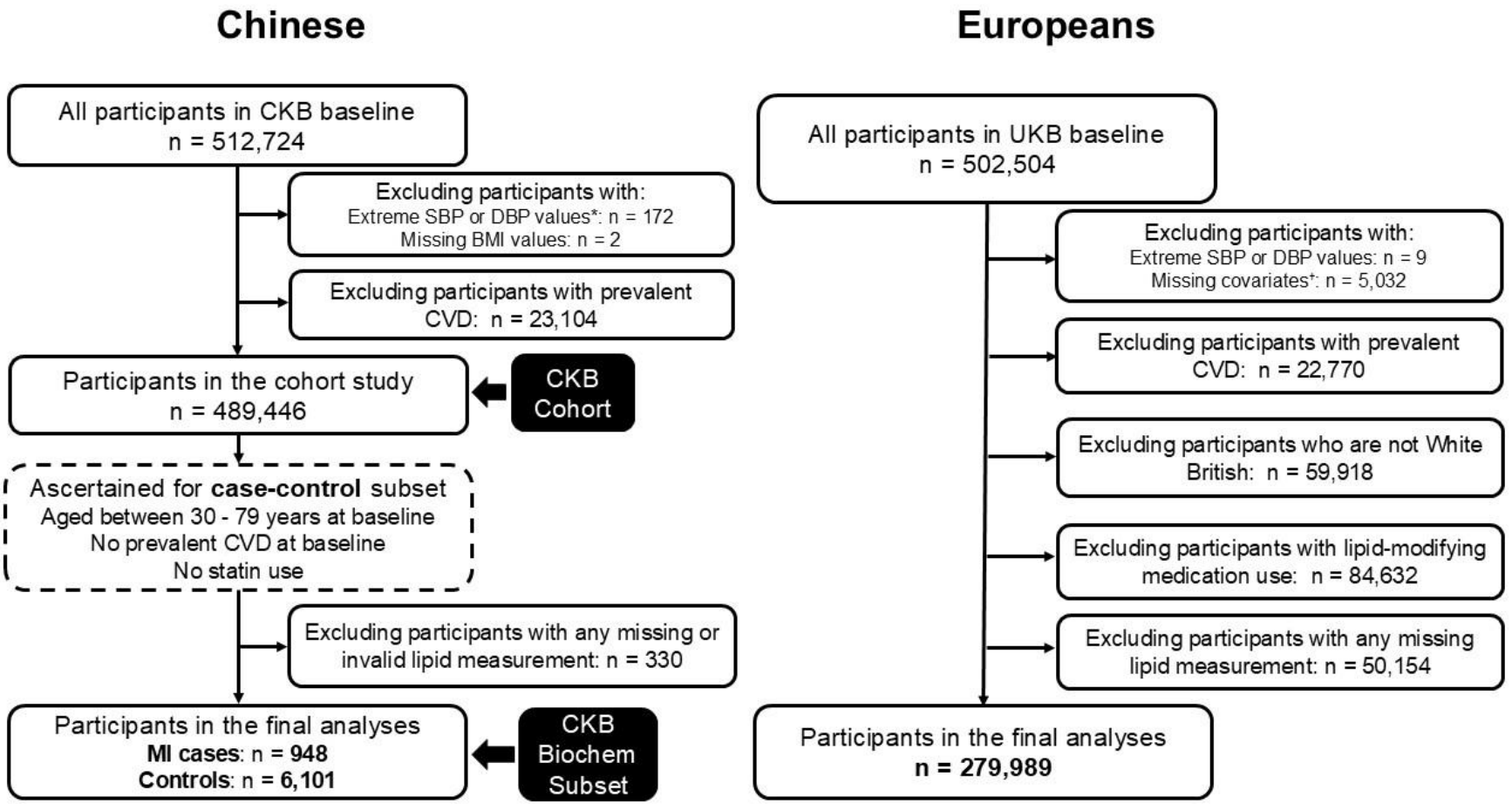
6.2.1. Study populations and settings

The analysis in this chapter involved a prospective cohort study from CKB, with SBP measurements, as outlined in Chapter 3.[73,99] For analyses involving both SBP and lipid traits, the analysis in this chapter involved the nested case-control study subset from CKB outlined in Chapter 4.[73,99] The analyses presented in this chapter on European populations involved the UKB cohort study,[98,101] as outlined in Chapter 3.

For the analysis of the association of SBP with MI risk only, the prospective cohort study from CKB excluded participants with prevalent CVD at baseline (n = 23,104) or with extreme or missing SBP or BMI values (n = 174). In the end, 489,446 participants, among whom 7,617 had first incident MI during follow-up, were included in the analysis for the Chinese population. We excluded individuals with prevalent CVD or on lipid-modifying medication at baseline for the UKB cohort study. Overall, a total of 279,989 participants with 5,341 first incident MI cases were included in the present analysis for European populations. Details of the participant selection process for all analyses are provided in ***Figure 6.1***.

The analysis on the interaction of SBP and plasma lipid traits included 948 MI cases and 6,101 controls from CKB, after excluding individuals with prevalent CVD or current use of lipid-modifying medication at baseline.

Figure 6.1. Flowchart for inclusion of study participants in Chinese and Europeans in observational analysis



Abbreviations. **CKB**: China Kadoorie Biobank; **SBP**: Systolic Blood Pressure; **DBP**: Diastolic Blood Pressure; **MI**: Myocardial Infarction; **CVD**: Cardiovascular Disease; **BMI**: Body Mass Index.
MI cases in the Chinese population were defined at the time of ascertainment.
*Extreme SBP: SBP ≥ 250 or < 80 mmHg and extreme DBP: DBP ≥ 140 or < 40 mmHg.
‡ missing covariates in UKB include: deprivation index (n = 527); BMI (n = 1,937); SBP (n = 797); educational attainment (n = 1,771).

6.2.2. Statistical methods

The baseline characteristics of the study populations were described in Chapter 3. For the present analyses, the age- and sex-adjusted mean SBP and LDL-C levels were estimated.

The analyses examined the association of usual levels of SBP with risk of MI. In UKB and the overall CKB cohort, Cox regression was used to estimate HRs and 95% CIs. In the CKB nested case-control subset, logistic regression was used to estimate ORs and 95% CIs. Both HRs and ORs are referred to as RRs. Specifically, analyses in CKB were adjusted for study region, education, smoking, alcohol consumption, physical activity, BMI, and diabetes. Analyses in UKB were stratified by 5-year age-at-risk groups and sex, and adjusted for Townsend Deprivation Index, education, smoking, alcohol consumption, physical activity, BMI, and diabetes. Adjustment for study region was necessary because the CKB includes participants from 10 study regions across China, which vary in mean SBP levels and incidence of MI.[99] Education, smoking, alcohol consumption, physical activity, BMI, and diabetes were included as covariates because they are established risk factors for MI and are also associated with SBP.[52,53,149] Repeat measurements of SBP from the resurveys in both CKB and UKB were used to correct for regression dilution bias using the MacMahon-Peto method, yielding associations with usual levels of SBP.[76]

To assess for potential interaction between SBP and lipid traits, analyses of the associations of LDL-C with MI risk were stratified by quartiles of baseline SBP. ApoB was also included as an additional analysis although it was not used in the MR analyses in Chapter 5 due to lack of valid instruments. A chi-squared test for heterogeneity was used to assess whether the RRs for LDL-C (or ApoB) differed across the SBP quartiles.[150] Because the regression models are additive on the

logarithmic scale, which corresponds to multiplicativity on the RR scale, the chi-squared test for heterogeneity provides a formal test of interaction on a multiplicative (log-additive) scale. Specifically, the test compared the goodness of fit between models that allowed separate log-RR estimates for each SBP quartile and models that constrained these estimates to be equal. A non-significant P value (e.g., P > 0.05) indicates that the strength of association between LDL-C (or ApoB) and risk of MI does not differ materially across levels of SBP, implying no evidence of interaction on the multiplicative (log-RR) scale. Heterogeneity of the strength of associations of levels of LDL-C with risk of MI across SBP strata was assessed with the estimation of Cochran's Q and I^2 statistics.[151] All statistical analyses were performed using R version 4.2.2.

6.3. Results

6.3.1. Baseline characteristics

The baseline characteristics of the participants included in this chapter's analyses are provided in ***Table 6.1***. Mean age was lower in the overall CKB cohort than in UKB (51.6 vs 55.5 years), while controls in the CKB biochemistry subset were older (58.3 years). UKB included a higher proportion of women and individuals with higher education than both the overall CKB cohort and the CKB biochemistry subset. A much higher proportion of men in the overall CKB cohort were current smokers compared with UKB (61.2% vs 11.9%), whereas weekly alcohol consumption was more common in UKB (80.6% vs 33.9% among male participants in the overall CKB cohort). Few Chinese women were smokers or drinkers. Diabetes prevalence was higher in the overall CKB cohort than in UKB (5.4% vs 1.6%), while antihypertensive use was similar (10.5% vs 12.8%). Mean BMI was lower in the overall CKB cohort than in UKB (23.6 vs 27.0 kg/m²), but mean SBP was comparable (130.6 vs 137.0 mmHg). Mean levels of LDL-C and non-HDL-C were about one-third lower among the controls in the CKB biochemistry subset than in UKB (2.3 vs 3.7 mmol/L and 3.3 vs 4.4 mmol/L, respectively).

Table 6.1. Baseline characteristics of participants

Characteristics Mean (SD) or %	Overall CKB cohort (n = 489,446)	CKB biochemistry subset		UKB (n = 279,989)
		MI cases (n = 948)	Controls (n = 6,101)	
Demographics				
Age, years	51.6 (10.5)	52.4 (7.1)	58.3 (10.9)	55.5 (8.0)
Female	59.1	39.6	47.9	57.1
Educated to middle school or above*	49.2	46.1	32.8	62.0
Lifestyle factors				
Male current smoker	62.0	75.1	61.2	11.9
Female current smoker	2.3	3.5	3.2	8.5
Male weekly drinker	33.9	27.0	28.2	80.6
Female weekly drinker	2.1	1.3	1.9	66.3
Total physical activity, MET-h/day**	6.6 (4.7)	5.5 (4.4)	6.5 (4.5)	6.4 (6.5)
Medical history and medications				
Diabetes†	5.4	11.9	4.8	1.6
Hypertension‡	33.7	54.2	41.5	49.7
Antihypertensive use	10.5	17.7	11.5	12.8
Physical measurements				
BMI, kg/m²	23.6 (3.4)	24.0 (3.6)	23.0 (3.3)	27.0 (4.6)
SBP, mmHg	130.6 (21.0)	140.8 (25.7)	135.0 (21.3)	137.0 (18.5)
DBP, mmHg	77.6 (11.1)	82.7 (13.6)	77.3 (11.0)	82.3 (10.1)
Lipid measurements				
Total cholesterol, mmol/L	-	4.8 (1.0)	4.6 (0.9)	5.9 (1.0)
Apolipoprotein B, g/L	-	0.9 (0.2)	0.8 (0.2)	1.1 (0.2)
Non-HDL-C, mmol/L	-	3.6 (1.1)	3.3 (0.9)	4.4 (1.0)
LDL-C, mmol/L	-	2.4 (0.8)	2.3 (0.6)	3.7 (0.8)
Triglycerides, mmol/L	-	2.0 (1.3)	1.8 (1.1)	1.7 (1.0)
HDL-C, mmol/L	-	1.2 (0.3)	1.3 (0.3)	1.5 (0.4)

Abbreviations. **SD**: Standard Deviation, **SBP**: Systolic Blood Pressure, **DBP**: Diastolic Blood Pressure, **BMI**: Body Mass Index, **HDL-C**: High-Density Lipoprotein Cholesterol, **LDL-C**: Low-Density Lipoprotein Cholesterol, **MET**: Metabolic Equivalent of Task, **UKB**: UK Biobank; **CKB**: the China Kadoorie Biobank.

* In UKB, this category includes those who were educated to above A levels/AS levels or equivalent.

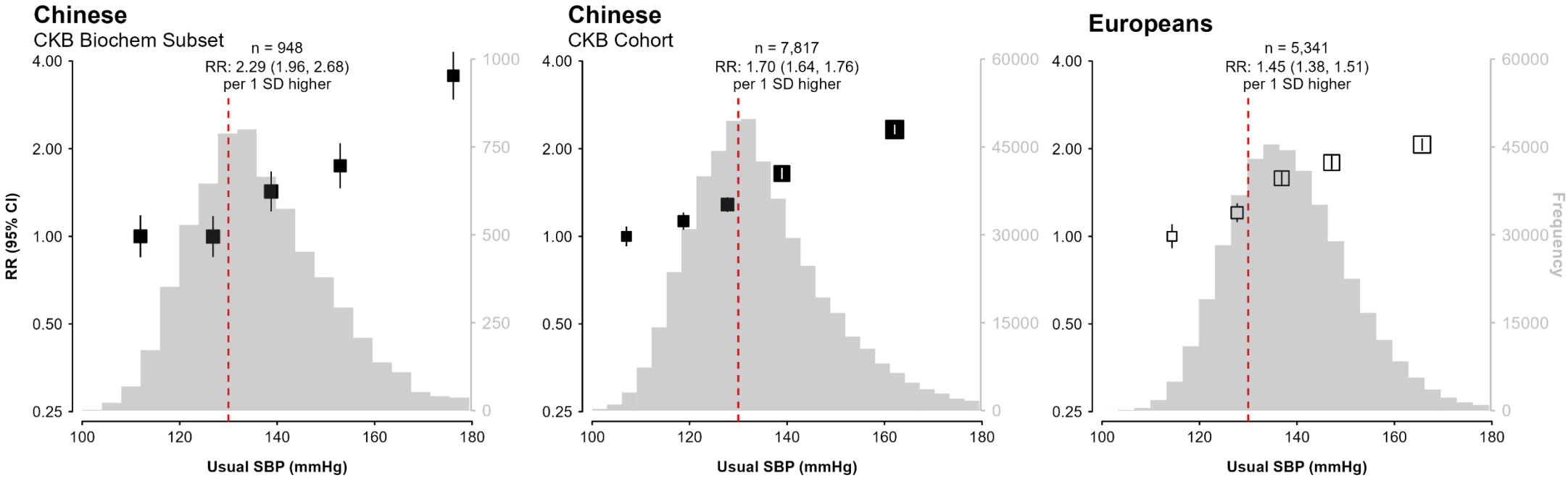
** Total physical activity excludes work-related activity. UKB has around 20% missing data in self-reported physical activity.

† In UKB, baseline diabetes includes self-reported cases or those participants with a HbA1c level above 48mmol/mol; In CKB, it includes self-reported cases and screen-detected cases (a blood glucose level ≥7.0 mmol/L and a fasting time >8 h, a blood glucose level ≥11.1 mmol/L and a fasting time <8 h, or a fasting blood glucose level ≥7.0 mmol/L).

6.3.2. Association of SBP with risk of MI

In both the Chinese and European populations, there were strong, positive, and approximately log-linear associations between higher usual levels of SBP and the risk of MI (***Figure 6.2***). In the overall CKB cohort, each 1-SD higher usual levels of SBP (21.01 mmHg) was associated with a 70% higher risk of MI (RR: 1.70, 95% CI: 1.64 - 1.76). A similar, though more extreme, association was observed in the CKB nested case-control subset, where 1-SD higher usual levels of SBP (25.35 mmHg) were associated with a 129% higher risk of MI (RR: 2.29, 95% CI: 1.96 - 2.68). In the UKB cohort, the association of higher levels of SBP and risk of MI was also log-linear and significant, but the strength of the association was weaker compared to the Chinese population. Each 1-SD higher usual levels of SBP (18.60 mmHg) was associated with a 45% higher risk of MI (RR: 1.45, 95% CI: 1.38 - 1.51).

Figure 6.2. Adjusted RRs for risk of MI by quintiles of SBP in Chinese and European populations



Red dotted lines indicate age-standardised mean SBP levels: 130.0 mmHg in the CKB Biochemistry subset (controls only), 128.9 mmHg in the CKB cohort, and 135.5 mmHg in the UK Biobank.
SDs for SBP: 25.35 mmHg in the CKB Biochemistry subset, 21.01 mmHg in the CKB cohort, and 18.60 mmHg in the UK Biobank.
Grey columns show the distribution of SBP. The distribution in the CKB Biochemistry subset is based solely on control participants.

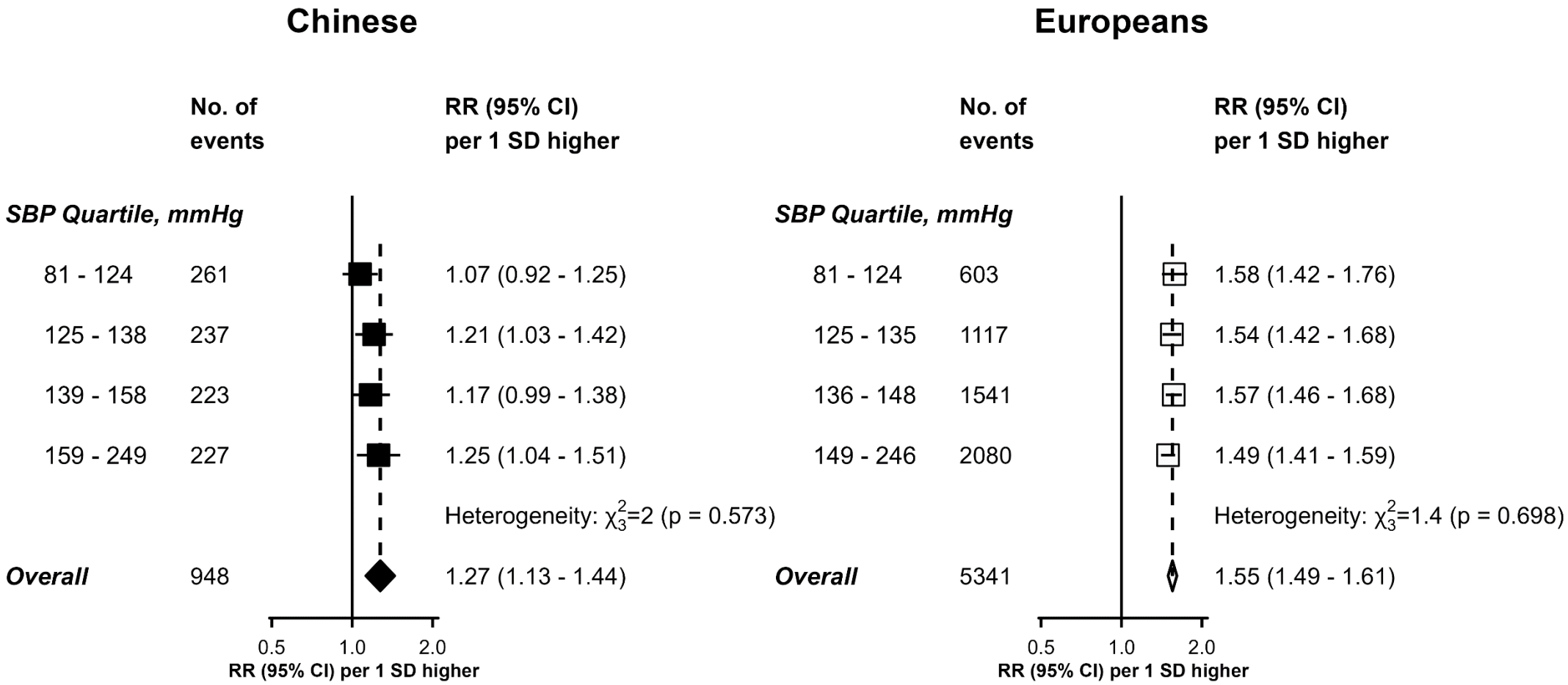
6.3.3. Association of LDL-C with risk of MI stratified by SBP levels.

In both the Chinese and European populations, the positive associations of plasma levels of LDL-C with risk of MI were consistent across all four quartiles of SBP (***Figure 6.3***). In the Chinese population, the RRs (95% CI) for risk of MI per 1-SD higher LDL-C were 1.07 (0.92–1.25) in the lowest SBP quartile (81–124 mmHg), 1.21 (1.03–1.42) in the second quartile (125–138 mmHg), 1.17 (0.99–1.38) in the third quartile (139–158 mmHg), and 1.25 (1.04–1.51) in the highest quartile (159–249 mmHg), with an overall RR of 1.27 (95% CI 1.13–1.44). In Europeans, the corresponding RRs (95% CI) for risk of MI per 1-SD higher LDL-C were 1.58 (1.42–1.76), 1.54 (1.42–1.68), 1.57 (1.46–1.68), and 1.49 (1.41–1.59) across successive SBP quartiles, with an overall RR of 1.55 (95% CI 1.49–1.61). There was no evidence of heterogeneity of the association of LDL-C with risk of MI across SBP quartiles, with P for heterogeneity = 0.573 in Chinese and 0.698 in European populations.

Similar results were observed when using ApoB as the LDL-C-related marker (***Figure 6.4***). In the Chinese population, higher plasma levels of ApoB were associated with a higher risk of MI across all SBP quartiles, with RRs (95% CI) for risk of MI per 1-SD higher ApoB ranging from 1.24 (1.06–1.44) in the lowest SBP quartile (81–124 mmHg) to 1.38 (1.14–1.68) in the highest quartile (159–249 mmHg). The overall RR per 1-SD higher ApoB was 1.45 (95% CI 1.29–1.62). In Europeans, the corresponding RRs for risk of MI per 1-SD higher ApoB were 1.57 (1.42–1.75) for the lowest SBP quartile, 1.59 (1.47–1.72) for the second quartile, 1.57 (1.47–1.68) for the third quartile, and 1.50 (1.42–1.59) for the highest quartile, with an overall RR of 1.56 (95% CI 1.51–1.62). The strength of associations was broadly consistent across quartiles of SBP levels in both Chinese and European

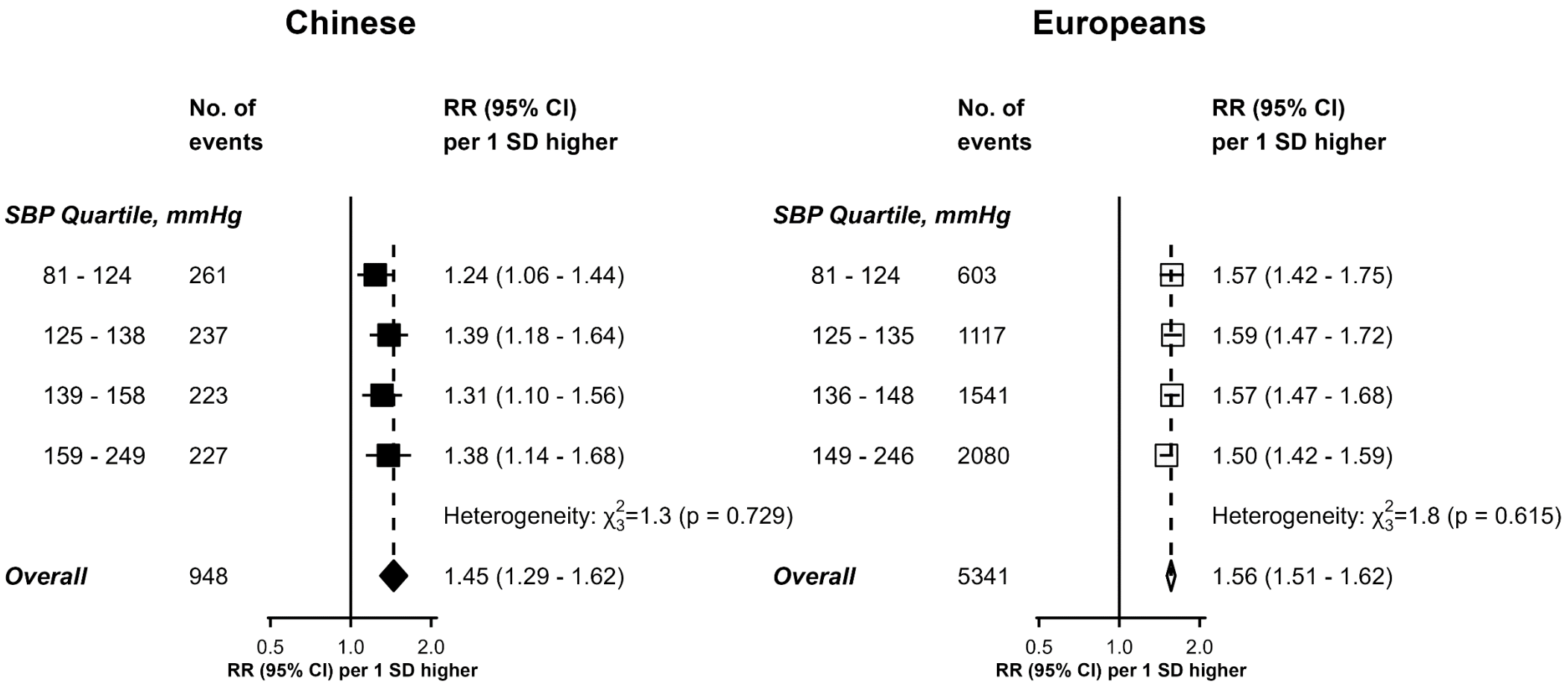
populations, with no evidence of interaction on a multiplicative scale between ApoB and SBP on the risk of MI (P for heterogeneity = 0.729 in Chinese and 0.615 in Europeans).

Figure 6.3. Association of LDL-C with risk of MI stratified by SBP quartiles



Abbreviations. **SBP**: Systolic Blood Pressure; **LDL-C**: Low-Density Lipoprotein Cholesterol; **MI**: Myocardial Infarction; **RR**: Risk Ratio; **CI**: Confidence Interval; **SD**: Standard Deviation.

Figure 6.4. Association of ApoB with risk of MI stratified by SBP quartiles



Abbreviations. SBP: Systolic Blood Pressure**; ApoB:** Apolipoprotein **B; MI:** Myocardial Infarction**; RR:** Risk Ratio**; CI:** Confidence Interval**; SD:** Standard Deviation.

6.4. Discussion

In this chapter, we investigated the joint observational associations of SBP and LDL-related lipid traits with MI risk in large-scale Chinese and European cohorts. The primary findings were twofold: first, higher usual levels of SBP were strongly and log-linearly associated with a higher risk of MI in both populations; second, the strength of the association of plasma levels of LDL-C and ApoB with risk of MI did not differ across levels of SBP. This lack of heterogeneity across SBP strata suggests that the effects of elevated SBP and plasma LDL-C are independent and that their associations with risk of MI act in a multiplicative (log-additive) manner in both Chinese and European populations.

The finding of a strong, positive, and log-linear association of higher usual levels of SBP with risk of MI is highly consistent with previous evidence summarised in the umbrella review of cohort studies outlined in Chapter 2. The findings of the present study were consistent with the previous meta-analyses of cohort studies such as PSC[52] and APCSC.[53] The estimate with the highest number of events from our umbrella review (number of fatal IHD = 28,797) was from PSC Europe,[52] which reported that a 10 mmHg lower usual SBP was associated with a 25% lower risk of fatal IHD (HR: 0.75, 95% CI: 0.75-0.76). The estimates were broadly comparable with those in PSC Asia [52] and APCSC[53] with an estimated ~25% lower risk of incident IHD or IHD mortality per 10 mmHg lower SBP. In summary, the umbrella review demonstrated concordant associations of SBP with risk of IHD between Asian populations (e.g., in APCSC and PSC Asia) and European populations (e.g., in PSC Europe). In the present study, the RR for MI was 1.45 and 1.70 per SD higher SBP (18.5 mmHg in Europeans and 21.0 mmHg in Chinese), corresponding to an 18% and 22% lower risk of MI per 10 mmHg lower SBP, respectively. Thus, the strength

of association in CKB was broadly consistent with previous large-scale cohorts such as APCSC and PSC, whereas the corresponding strength of association in UKB was weaker compared with PSC Europe (18% vs 25% lower risk per 10 mmHg lower SBP). This weaker strength of the association of SBP with risk of MI in UKB could potentially reflect the higher prevalence of antihypertensive treatment in this population compared with studies included in the PSC, and changes in anthropometric and metabolic characteristics in European populations, including higher waist-to-hip ratio and altered dietary patterns, which may modify the observed association between SBP and risk of MI.[152-154]

Importantly, observational studies typically report a more extreme association of levels of SBP and risk of MI than those observed in meta-analyses of BP-lowering RCTs. IPD analyses from BPLTTC,[59] which included 38 primary prevention trial data points, estimated that a 10 mmHg reduction in SBP was associated with a 10% reduction in IHD risk (HR: 0.90, 95% CI: 0.82-0.98). The consistency in the strength of associations of levels of SBP with risk of MI observed in the present analyses and previous studies reinforces the validity of the SBP measurements in both CKB and UKB and confirms that SBP is an important independent risk factor for MI in both Chinese and European populations.

The results of the joint associations of higher plasma levels of LDL-C and levels of SBP with risk of MI clarified the conflicting findings from previous studies. Our findings differ from both the APCSC analysis, which suggested the associations of levels of SBP were weaker among participants with higher levels of cholesterol,[79] and the EPOCH-Japan study,[54] which reported a stronger association of levels of SBP with risk of MI among participants with lower levels of cholesterol. The reasons for these discrepancies may relate to differences in study design, statistical power,

and the range of cholesterol levels in the study populations. By using two large, contemporary biobanks with standardised measurements and rigorous correction for within-person variability (regression dilution bias), the present analysis provides a robust assessment of interaction. The consistent findings obtained in both the Chinese and European populations, which have different ranges of baseline levels of LDL-related traits, strengthen the conclusions that the proportional risks associated with SBP and LDL-related lipids are independent of one another.

The demonstration of independence has important implications for public health and clinical practice. It provides observational evidence to support the rationale behind the combination pill,[21] which is based on the assumption that simultaneously lowering both SBP and LDL-C levels should result in a combined, multiplicative risk reduction. If the risk for MI associated with LDL-C were to be attenuated at higher SBP levels, for example, the relative benefit of LDL-lowering would be diminished in individuals with hypercholesterolemia. The results of the present study indicate that the positive associations of LDL-C with risk of MI were observed across all baseline plasma levels of SBP, suggesting that the strength of association between LDL-C and MI risk did not materially differ by SBP levels.

In conclusion, building on findings from Chapter 5, which established the causal relevance of LDL-C for risk of MI, this chapter provides additional evidence that the observational associations of SBP and LDL-related lipid traits with risk of MI are independent in both Chinese and European populations. These findings further advance the overall aim of this thesis to inform multi-factorial strategies targeting both SBP and LDL-C for the prevention of MI. The results provide observational support for strategies that concurrently address both risk factors for primary prevention of MI. Building on the evidence presented here, the next chapter (Chapter

7) aims to triangulate the observational findings by providing a causal assessment of whether the independent and multiplicative (log-additive) effects observed here persist when examined using MR approaches that better reflect unconfounded, lifelong differences in these traits.

Chapter 7 Causal relevance of combined effects of LDL-C and SBP with MI risk

7.1. Background

The previous chapter (Chapter 6) demonstrated that the observational associations of LDL-C with risk of MI were consistent across SBP quartiles, which suggests that the effects of SBP and LDL-C on risk of MI are multiplicative. This supports the basic premise of the risk calculation in the combination therapy approach proposed by Wald and Law, which assumes independent effects of SBP and LDL-C on risk of CVD.[21]

Some recent combination pill trials showed that the combined effects of lowering SBP and LDL-C simultaneously on risk of CVD were not as strong as previously predicted by the original calculation from Wald and Law.[21,155] Although our observational analyses suggest independence of the associations of levels of SBP and plasma levels of LDL-C with risk of MI, such analyses cannot exclude residual confounding.

A factorial RCT is a study design that simultaneously tests two or more interventions within the same population. In a 2 × 2 factorial RCT, participants are randomly allocated to one of four groups: receiving intervention A only, intervention B only, both interventions together, or neither (placebo).[156] This design enables assessment of both the independent and combined effects of the two interventions. Similar to factorial RCTs, a factorial MR approach can be employed to estimate the combined effects of modifying two exposures.[157] This design uses genetic variants as instruments to independently stratify a population by their predicted levels of two different exposures (e.g., high/low plasma levels of LDL-C and high/low levels of

SBP). This allows for the assessment of the main effects of each exposure and an estimation of their combined effect on a given outcome. By comparing odds ratios across the four genetic groups, this method can test for a multiplicative (log-additive) interaction between the two exposures in their association with the outcome.

Previous factorial MR studies in European populations have suggested that the effects of LDL-C and SBP on risk of CVD are independent. Ference et al. conducted a large factorial MR study in the UK Biobank which showed that genetically lower LDL-C and SBP were associated with independent, log-additive, and dose-dependent reductions in major coronary events.[74] The study highlights the need to extend such evidence to Chinese populations where risk factor profiles and disease patterns differ markedly.[74]

This chapter, therefore, aims to build upon the findings from the observational analysis in Chapter 6 by using a factorial MR design. Furthermore, the restricted instruments for LDL-C generated in Chapter 5 can be used in this chapter to minimise pleiotropic effects. The primary objectives are to assess the combined causal relevance of genetically-predicted plasma levels of LDL-C and levels of SBP on risk of MI in Chinese and European populations, and to test for potential interaction between them. By conducting these analyses in parallel in both CKB and UKB, this chapter seeks to provide reliable, trans-ethnic evidence on the validity of strategies that advocate for the joint lowering of SBP and LDL-C.

7.2. Methods

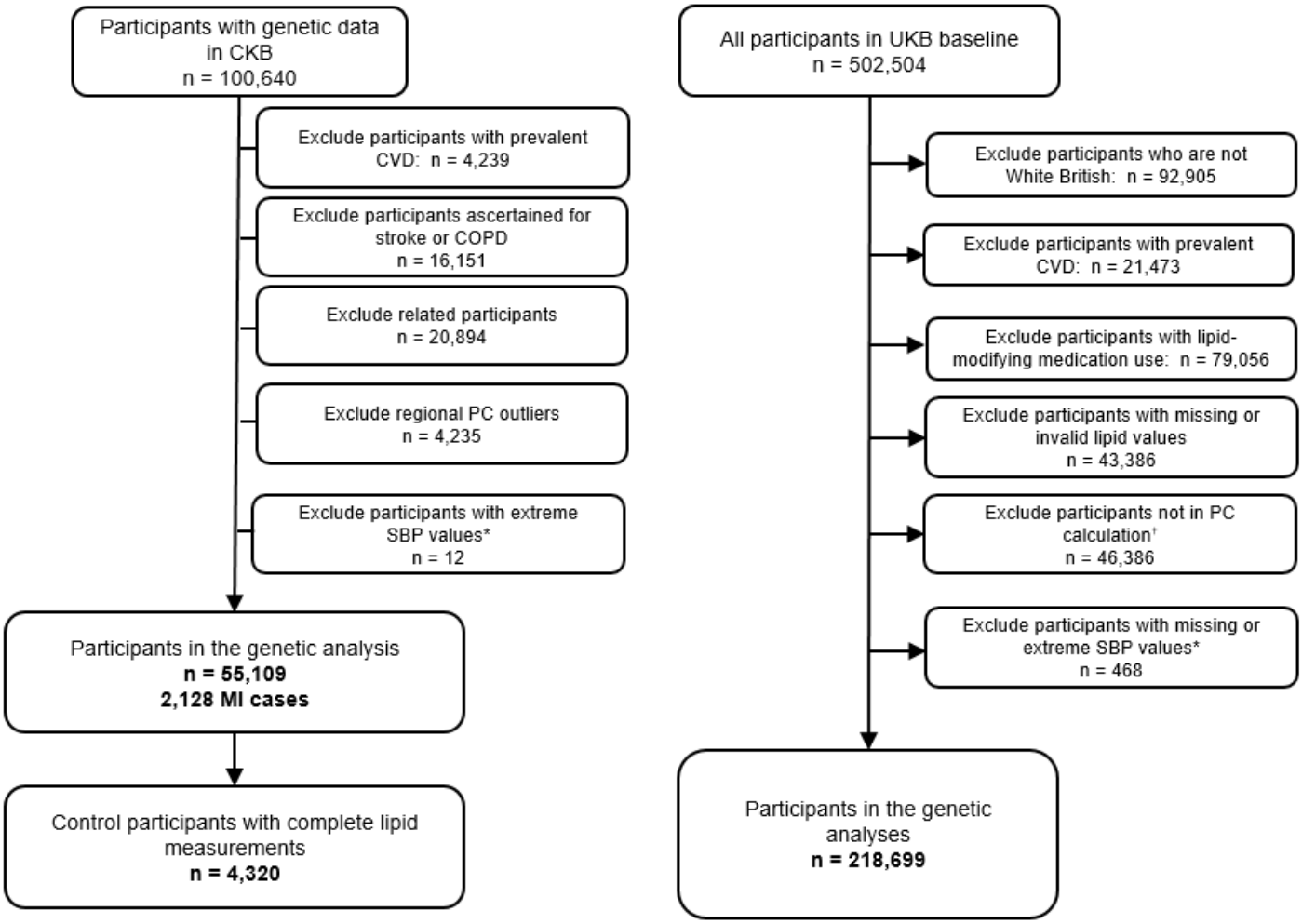
7.2.1. Study sample

The analyses in this chapter used the genetic subsets of CKB and UKB. The selection of participants for analyses in this chapter is displayed in ***Figure 7.1***.

For the Chinese population, CKB included 100,640 participants with genetic data. Exclusions were made for individuals with prevalent CVD (n = 4,239), those ascertained for stroke or COPD (n = 16,151), related individuals (n = 20,894), regional PC outliers (n = 4,235), and those with extreme SBP values (n = 12). This resulted in a final sample of 55,109 participants, including 2,128 MI cases. For analyses requiring lipid measurements, a subset of 4,320 control participants with complete lipid measurements was used.

For European populations, UKB study involved 502,504 participants at baseline. Exclusions included participants who were not of White British ancestry (n = 92,905), had prevalent CVD (n = 21,473), were on lipid-modifying medication (n = 79,056), had missing or invalid lipid values (n = 43,386), were not included in the PC calculation (n = 46,386), or had missing or extreme SBP values (n = 468). The final sample for the factorial MR analysis consisted of 218,699 participants.

Figure 7.1. Flowchart for inclusion in CKB and UKB



*Individuals with extreme values of SBP (SBP <80 mmHg or ≥250 mmHg) or with missing SBP were excluded

†In UKB, participants in the PCA include those who have missing rate in autosomes ≤ 0.02; are not outliers for missingness or heterozygosity; are in a maximal set of unrelated individuals; are not sex-discordant.

Abbreviations. **MI**: Myocardial Infarction; **CVD**: Cardiovascular Disease; **BMI**: Body Mass Index; **PCA**: Principal Component Analysis; **CKB**: China Kadoorie Biobank; **UKB**: UK Biobank.

7.2.2. Instrumental variable selection

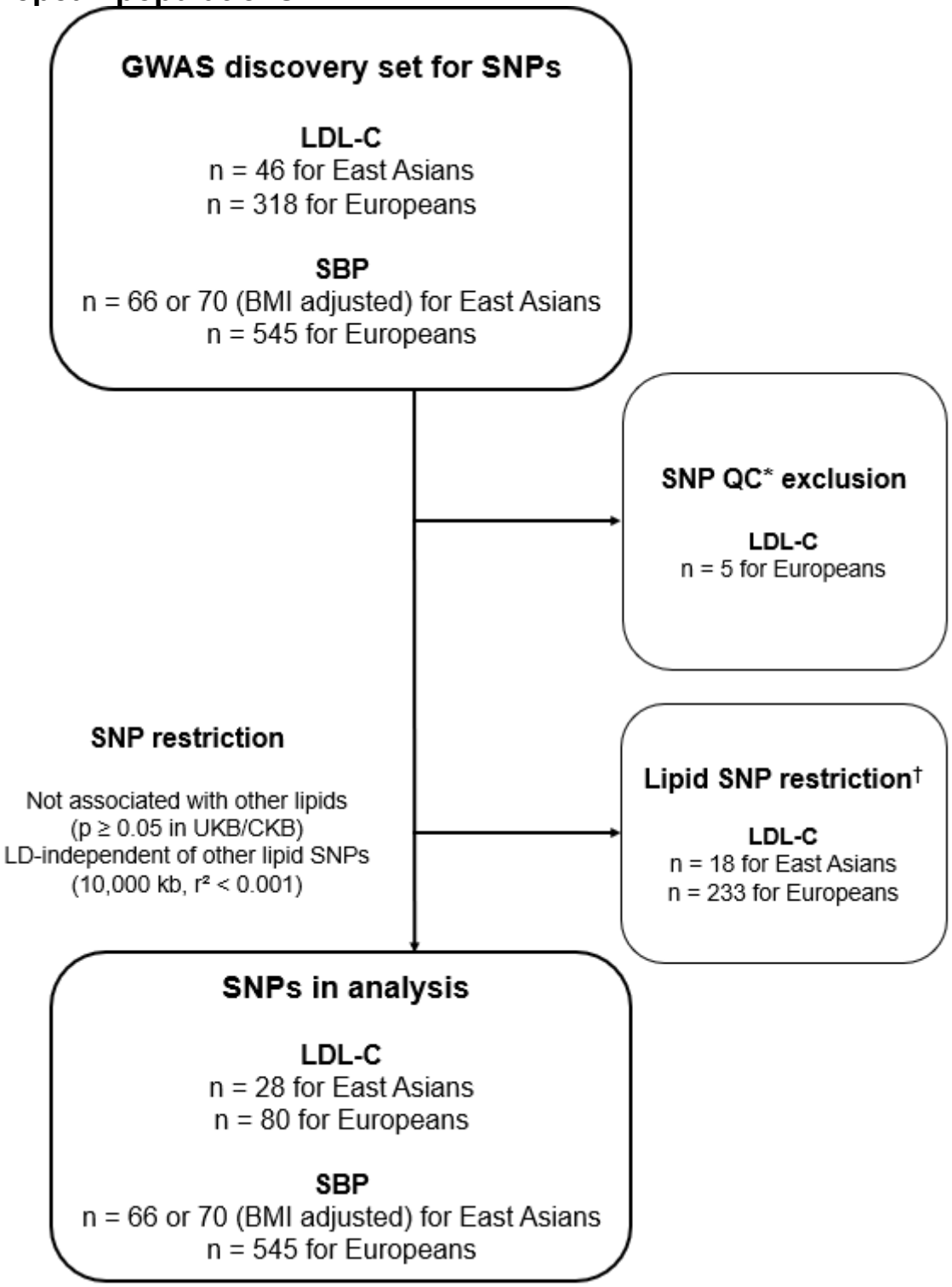
Variants for GRS for SBP were constructed using SNPs from China Kadoorie Biobank (CKB) in the Chinese population,[146] and using 545 SNPs from a large discovery set in European populations.[158] The SNPs from CKB include those with GWAS significant associations unadjusted for BMI (n = 66) and adjusted for BMI (n = 70). The SNPs used in the analyses for the Chinese population are presented in **Tables S7** – **S9**.

Weights (beta values associated with SBP) were derived by jack-knifing to avoid bias towards the observational association due to overfitting of effect size estimates in CKB.[146,159] Effect sizes for each variant were repeatedly re-estimated in 99% of the sample and then applied to the remaining 1% of individuals (100-fold jack-knifing), iterating this process across the entire dataset. This leave-one-out resampling ensures that the genetic weights used to construct the score for each participant are independent of their own observed phenotype, thereby reducing overfitting and bias towards observational correlations.[159,160] The variants for SBP from Biobank Japan[161] were compared with instruments in CKB, as shown in ***Table 7.1***.

Variants related to LDL-C were extracted from the Global Lipids Genetics Consortium,[132] which includes ancestry-specific data. GWAS data were downloaded from the GLGC website.[133] For European ancestry, GLGC provides data excluding UKB. All clumping was performed using PLINK LD clumping method, with the clumping window set to 10,000 kb and r-squared set to 0.001 to balance the need to capture the lead SNP in LD and reduction of noise.[134,135] LDL-C instruments were further restricted based on the following criteria: (1) they showed no association with triglycerides or HDL-C in either UKB or CKB datasets ($p < 0.05$); and (2) they were not in LD with SNPs significantly associated with triglycerides or HDL-C at GWAS

level (using a clumping window of 10,000 kb and an r^2 threshold of 0.001). Overall, 28 SNPs were selected for the Chinese population, and 80 SNPs were selected for European populations for LDL-C (***Figure 7.2***). Lists of SNPs for SBP in European populations from Keaton et al.[158] are provided in supplementary materials in ***Table S10***.

Figure 7.2. SNP selection for LDL-C and SBP in East Asian and European populations



*In Europeans (UKB), SNP QC excludes SNPs with imputation information less than 0.4 and minor allele frequency less than 0.005.

† The restricted SNPs were chosen based on the following criteria: (1) they were associated with LDL-C at the GWAS significance level ($p < 5 \times 10^{-8}$); (2) they showed no association with triglycerides or HDL-C in either UKB or CKB datasets ($p < 0.05$); and (3) they were not in LD with SNPs significantly associated with triglycerides or HDL-C at GWAS level (using a clumping window of 10,000 kb and an r^2 threshold of 0.001).

Abbreviations. GWAS: Genome-Wide Association Study; **SNP**: Single Nucleotide Polymorphism; **LDL-C**: Low-Density Lipoprotein Cholesterol; **SBP**: Systolic Blood Pressure; **QC**: Quality Control; **LD**: Linkage Disequilibrium; **UKB**: UK Biobank; **CKB**: China Kadoorie Biobank; **HDL-C**: High-Density Lipoprotein Cholesterol.

Table 7.1. Comparison of different variant selection for SBP

GWAS	N SNPs	% Variance Explained	F statistics
BioBank Japan	33	1.15	45.52
CKB (Unadjusted for BMI)	66	2.00	91.17
CKB (Adjusted for BMI)	70	2.20	96.53
Keaton et al. (Europeans)	545	4.24	8766.17

Abbreviations. **SNPs**: Single Nucleotide Polymorphisms; **CKB**: China Kadoorie Biobank; **BMI**: Body Mass Index.

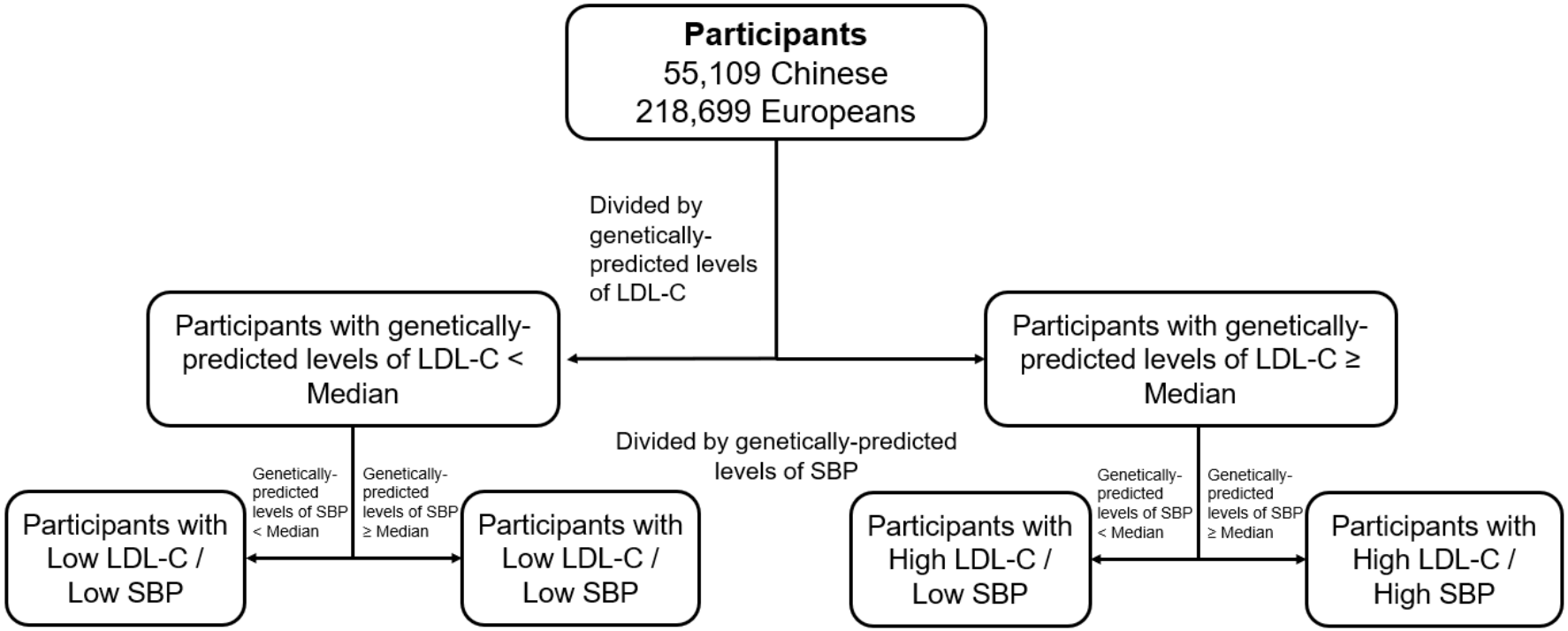
7.2.3. Statistical methods

Participants were classified into 4 (2 × 2) broad categories based on whether their SBP GRS was above the median level and whether LDL-C GRS was above the median level. Dichotomising the genetic scores at the median ensures roughly equal group sizes, which improves statistical efficiency and mirrors the structure of a 2 × 2 factorial RCT.[157]

Participants in the low LDL-C and SBP GRS categories were used as the reference group and the ORs of the exposure to higher LDL-C and/or higher SBP were calculated, with adjustment for the ten national PCs, age, age-squared, genotyping array, sex, and region (CKB only).

By comparing the OR for the combined high LDL-C and high SBP group with the product of the ORs for each exposure separately, this 2 × 2 framework enables estimation of an interaction effect. A significant deviation from this expected multiplicative effect indicates interaction between genetically predicted LDL-C and SBP on the risk of MI. The factorial design is presented in ***Figure 7.3***. All statistical analyses were performed using R version 4.2.2.

Figure 7.3. Factorial design (2 × 2) with participants in four groups based on GRS



Abbreviations. **LDL-C**: Low-Density Lipoprotein Cholesterol; **SBP**: Systolic Blood Pressure; **GRS**: Genetic Risk Score.

Stratified analyses by SBP GRS quartile were conducted for the association of genetically-predicted LDL-C levels (per SD differences) and risk of MI. Similarly, stratified analyses by LDL-C GRS quartile were conducted for the association of genetically-predicted SBP levels (per SD differences) and risk of MI.

In addition, participants were classified into 16 (4 X 4) categories based on the quartiles of SBP or LDL-C GRS to examine the dose-response relationship between exposures to different levels of genetically-predicted SBP and LDL-C. Participants in the first SBP GRS quartile and the first LDL-C GRS quartile were used as the reference group and ORs for other 15 groups were computed. The mean observational levels of SBP and LDL-C for each group were calculated. To calculate the combined exposure to 1 mmol/L higher LDL-C and 10 mmHg higher SBP, a meta-regression model, with the outcome as log(OR) of MI and the exposure as mean levels of LDL-C and SBP, was implemented.

7.3. Results

7.3.1. Genetically-predicted SBP and MI risk

As shown in ***Table 7.2***, in both populations, higher levels of genetically-predicted SBP were associated with higher risk of MI.

Using the GRS unadjusted for BMI in the Chinese population, a 1-SD higher level of genetically-predicted SBP was associated with a 78% higher risk of MI (RR: 1.78, 95%CI: 1.22 – 2.60). When using the GRS adjusted for BMI in the Chinese population, this association was stronger, and a 1-SD higher level of genetically-predicted SBP was associated with a 103% higher risk of MI (RR: 2.03, 95%CI: 1.39 – 2.97). In European populations, a 1-SD higher level of genetically-predicted SBP was associated with a 104% higher risk of MI (RR: 2.04, 95%CI: 1.73 – 2.41).

Table 7.2. Associations of genetically-predicted SBP and risk of MI in Chinese and European populations

GWAS	N SNPs	% Variance Explained	F statistics	OR (95% CI) with MI per 1-SD increase
BioBank Japan	33	1.15	45.52	1.75 (1.12 – 2.73)
CKB (Unadjusted for BMI)	66	2.00	91.17	1.78 (1.22 – 2.60)
CKB (Adjusted for BMI)	70	2.20	96.53	2.03 (1.39 – 2.97)
Keaton et al. (Europeans)	545	4.24	8766.17	2.04 (1.73 – 2.41)

Abbreviations. **SNPs**: Single Nucleotide Polymorphisms; **CKB**: China Kadoorie Biobank; **UKB**: UK Biobank; **BMI**: Body Mass Index; **RR**: Risk Ratio

7.3.2. Combined exposure to genetically-predicted lower LDL-C and SBP

The primary factorial analysis involved categorising participants into four groups based on whether their GRS for LDL-C and SBP were above or below the mean. The results of this 2 × 2 factorial analysis are presented in ***Table 7.3***.

In the Chinese population, compared to the reference group with lower levels of genetically-predicted LDL-C and SBP, individuals with higher levels of genetically-predicted LDL-C had a 15% higher risk of MI (OR: 1.15, 95%CI: 1.01 – 1.31). Those with higher levels of genetically-predicted SBP had a 23% higher risk of MI (OR: 1.23, 95%CI: 1.09 – 1.40). Individuals with higher levels of both risk factors had a 25% higher risk of MI (OR: 1.25, 95%CI: 1.10 – 1.42).

In European populations, a similar pattern was observed. Compared to the reference group, higher levels of genetically-predicted LDL-C were associated with a 5% higher risk of MI (OR: 1.05, 95%CI: 0.95 – 1.15). Higher levels of genetically-predicted SBP were associated with a 25% higher risk of MI (OR: 1.25, 95%CI: 1.14 – 1.36). Individuals with higher genetically-predicted levels of both SBP and LDL-C had a 26% higher risk of MI (OR: 1.26, 95%CI: 1.16 – 1.38).

There was no evidence of an interaction on the multiplicative scale between the effects of genetically-predicted SBP and LDL-C on the risk of MI in either the Chinese population (P for interaction = 0.14) or the European populations (P for interaction = 0.62).

Table 7.3 Factorial MR with combined exposure to lower LDL-C and SBP (adjusted for BMI SBP GRS in CKB)

Group	N events	Mean SBP mmHg	Mean LDL-C mmol/L	RR (95% CI)	*P for interaction*
Chinese					0.14
Low LDL-C / Low SBP	476	129.04	2.22	Ref	
High LDL-C / Low SBP	529	128.71	2.46	1.15 (1.01 – 1.31)	
Low LDL-C / High SBP	572	133.77	2.16	1.23 (1.09 – 1.40)	
High LDL-C / High SBP	571	133.70	2.43	1.25 (1.10 – 1.42)	
1 mmol higher LDL-C and 10mmHg higher SBP				2.18 (1.39 – 3.41)	
Europeans					0.62
Low LDL-C / Low SBP	885	134.53	3.63	Ref	
High LDL-C / Low SBP	931	134.21	3.81	1.05 (0.95 – 1.15)	
Low LDL-C / High SBP	1,089	139.87	3.64	1.25 (1.14 – 1.36)	
High LDL-C / High SBP	1,074	139.66	3.81	1.26 (1.16 – 1.38)	
1 mmol higher LDL-C and 10 mmHg higher SBP				2.15 (1.69 – 2.73)	

*P for interaction derived from the coefficient corresponding to the interaction between dichotomised SBP and LDL (with "high" defined as values above the mean).

Abbreviations. **SBP**: Systolic Blood Pressure; **LDL-C**: Low-Density Lipoprotein Cholesterol; **CKB**: China Kadoorie Biobank; **UKB**: UK Biobank; **BMI**: Body Mass Index; **RR**: Risk Ratio.

These findings were further supported by stratified analyses, which assessed the association of one genetically-predicted risk factor with MI across quartiles of the other. As shown in ***Table 7.4***, the associations of higher levels of genetically-predicted LDL-C with risk of MI were consistent across all four quartiles of the SBP GRS in both populations. There was no evidence of heterogeneity for the associations of genetically predicted levels of LDL-C with risk of MI across SBP strata in the Chinese population (using the SBP GRS adjusted for BMI, P for heterogeneity = 0.29) or in the European population (P for heterogeneity = 0.69).

Similarly, as presented in ***Table 7.5***, the association of higher genetically-predicted levels of SBP with risk of MI was consistent across all four quartiles of the LDL-C GRS. There was no evidence of statistical heterogeneity in either the Chinese population (using the SBP GRS adjusted for BMI, P for heterogeneity = 0.57) or the European population (P for heterogeneity = 0.48). The results were consistent using SBP GRS unadjusted for BMI in the Chinese population (***Tables 7.6*** and ***7.7***).

Table 7.4. Genetically-predicted per SD difference (0.65 mmol/L in Chinese and 0.80 mmol/L in Europeans) in LDL-C and risk of MI, stratified by SBP GRS (adjusted for BMI)

SBP quartile	N events	Mean SBP mmHg*	Mean LDL-C	F statistics	R-squared	RR (95% CI)	P value
Chinese							
1	467	127.57	2.31	56.33	5.20	1.02 (0.65 – 1.58)	
2	533	130.18	2.34	85.29	6.68	1.08 (0.74 – 1.57)	
3	524	132.23	2.34	48.69	4.77	1.79 (1.21 – 2.64)	
4	604	135.23	2.32	98.20	8.53	1.17 (0.80 – 1.70)	
Overall	2,128	131.30	2.33	283.72	6.17	1.23 (1.01 – 1.50)	
						P for trend	0.54
						P for heterogeneity	0.29
Europeans							
1	823	132.74	3.72	956.25	1.76	1.47 (0.83 – 2.59)	
2	993	136.00	3.72	918.84	1.72	1.51 (0.89 – 2.56)	
3	1,021	138.16	3.73	882.80	1.67	1.12 (0.66 – 1.88)	
4	1,142	141.38	3.72	697.97	1.29	1.49 (0.84 – 2.63)	
Overall	3,979	137.07	3.73	3443.31	1.60	1.37 (1.04 – 1.79)	
						P for trend	0.78
						P for heterogeneity	0.69

*Mean observational level in the genetic subset used in the analysis.
Abbreviations. **SBP**: Systolic Blood Pressure; **LDL-C**: Low-Density Lipoprotein Cholesterol; **CKB**: China Kadoorie Biobank; **UKB**: UK Biobank; **BMI**: Body Mass Index; **RR**: Risk Ratio.

Table 7.5. Genetically-predicted per SD difference in SBP (21.54 mmHg in Chinese and 18.49 mmHg in Europeans) and risk of MI, stratified by LDL-C GRS (adjusted for BMI for SBP SNPs in CKB)

LDL-C quartile	N events	Mean LDL-C mmol/L*	Mean SBP	F statistics	R-squared	RR (95% CI)	P value
Chinese							
1	508	2.09	131.35	35.12	2.67	1.52 (1.04 – 2.22)	
2	530	2.33	131.46	24.65	2.32	1.88 (1.28 – 2.78)	
3	529	2.38	131.28	24.91	2.45	2.78 (1.91 – 4.06)	
4	561	2.48	131.12	14.04	1.59	2.22 (1.43 – 3.46)	
Overall	2,128	2.32	131.30	96.53	2.20	2.03 (1.39 – 2.97)	
						P for trend	0.23
						P for heterogeneity	0.57
Europeans							
1	884	3.40	137.42	2162.25	4.16	2.19 (1.54 – 3.12)	
2	1,007	3.65	137.21	2233.93	4.30	1.87 (1.35 – 2.58)	
3	967	3.80	137.00	2299.27	4.40	2.18 (1.57 – 3.03)	
4	1,121	4.04	136.65	2069.08	4.10	2.03 (1.48 – 2.79)	
Overall	3,979	3.72	137.07	8766.17	4.24	2.04 (1.73 – 2.41)	
						P for trend	0.85
						P for heterogeneity	0.48

*In the Chinese population, mean LDL-C levels were obtained from biochemistry subset with genetic data, complete lipid measurements, and control subjects. In European populations, the levels were mean observational level in the genetic subset used in the analysis

Abbreviations. **SBP**: Systolic Blood Pressure; **LDL-C**: Low-Density Lipoprotein Cholesterol; **CKB**: China Kadoorie Biobank; **UKB**: UK Biobank; **BMI**: Body Mass Index; **RR**: Risk Ratio.

Table 7.6. Genetically-predicted per SD difference (0.65 mmol/L in Chinese and 0.80 mmol/L in Europeans) in LDL-C and risk of MI, stratified by SBP GRS (unadjusted for BMI)

SBP quartile	N events	Mean SBP mmHg*		F statistics	R-squared	RR (95% CI)	P value
Chinese							
1	496	127.70	2.32	57.61	5.16	1.56 (1.00 – 2.42)	
2	503	130.25	2.33	73.26	6.41	1.48 (1.01 – 2.17)	
3	542	132.34	2.34	77.79	5.99	0.90 (0.61 – 1.33)	
4	587	134.91	2.32	73.88	6.87	1.17 (0.81 – 1.71)	
Overall	2,128	131.30	2.33	283.72	6.17	1.23 (1.01 – 1.50)	
						P for trend	0.31
						P for heterogeneity	0.21
Europeans							
1	823	132.74	3.72	956.25	1.76	1.47 (0.83 – 2.59)	
2	993	136.00	3.72	918.84	1.72	1.51 (0.89 – 2.56)	
3	1,021	138.16	3.73	882.80	1.67	1.12 (0.66 – 1.88)	
4	1,142	141.38	3.72	697.97	1.29	1.49 (0.84 – 2.63)	
Overall	3,979	137.07	3.73	3443.31	1.60	1.37 (1.04 – 1.79)	
						P for trend	0.78
						P for heterogeneity	0.69

*Mean observational level in the genetic subset used in the analysis.

Abbreviations. **SBP**: Systolic Blood Pressure; **LDL-C**: Low-Density Lipoprotein Cholesterol; **CKB**: China Kadoorie Biobank; **UKB**: UK Biobank; **BMI**: Body Mass Index; **RR**: Risk Ratio.

Table 7.7. Genetically-predicted per SD difference in SBP (21.54 mmHg in Chinese and 18.49 mmHg in Europeans) and risk of MI, stratified by LDL-C GRS (unadjusted for BMI in CKB for SBP SNPs)

LDL-C quartile	N events	Mean LDL-C mmol/L*		F statistics	R-squared	RR (95% CI)	P value
Chinese							
1	508	2.09	131.35	27.34	2.21	2.09 (1.35 – 3.26)	
2	530	2.33	131.46	36.76	2.93	1.69 (1.16 – 2.47)	
3	529	2.38	131.28	17.95	1.60	2.31 (1.57 – 3.41)	
4	561	2.48	131.12	12.71	1.43	1.30 (0.89 – 1.89)	
Overall	2,128	2.32	131.30	91.17	2.00	1.78 (1.22 – 2.60)	
						P for trend	0.43
						P for heterogeneity	0.72
Europeans							
1	884	3.40	137.42	2162.25	4.16	2.19 (1.54 – 3.12)	
2	1,007	3.65	137.21	2233.93	4.30	1.87 (1.35 – 2.58)	
3	967	3.80	137.00	2299.27	4.40	2.18 (1.57 – 3.03)	
4	1,121	4.04	136.65	2069.08	4.10	2.03 (1.48 – 2.79)	
Overall	3,979	3.72	137.07	8766.17	4.24	2.04 (1.73 – 2.41)	
						P for trend	0.85
						P for heterogeneity	0.48

*In the Chinese population, mean LDL-C levels were obtained from biochemistry subset with genetic data, complete lipid measurements, and control subjects. In European populations, the levels were mean observational level in the genetic subset used in the analysis

Abbreviations. **SBP**: Systolic Blood Pressure; **LDL-C**: Low-Density Lipoprotein Cholesterol; **CKB**: China Kadoorie Biobank; **UKB**: UK Biobank; **BMI**: Body Mass Index; **RR**: Risk Ratio.

7.3.3. Dose-response relationships and meta-regression results

To further explore the combined causal associations, a 4 x 4 factorial analysis was conducted by categorising participants into 16 groups based on the quartiles of their GRS for SBP and LDL-C. Overall, this analysis demonstrated a dose-response relationship in both populations (***Table 7.8***, ***Table 7.9***, and ***Table 7.10***).

Relative to the reference group (quartile 1 for both SBP and LDL-C GRS), there was a progressive and graded higher risk of MI with increasing quartiles of either or both GRS. In the Chinese population (using the BMI-adjusted SBP GRS), individuals in the highest quartile for both SBP and LDL-C did not have the highest point estimate for risk; instead, the highest risk was observed among those in the second quartile of the SBP GRS and the fourth quartile of the LDL-C GRS, who had a 72% higher risk of MI (RR: 1.72, 95%CI: 1.34 – 2.20), as shown in ***Table 7.8***. The results were broadly consistent using SBP GRS unadjusted for BMI (***Table 7.9***).

In European populations, the trend was more consistent, with individuals in the highest quartile for both SBP and LDL-C GRS having a 60% higher risk of MI (RR: 1.60, 95%CI: 1.33 – 1.93) compared to the reference group (***Table 7.10***).

A meta-regression model across the factorial groups was used to provide a summary estimate for the combined effect. This analysis estimated that a combined lifelong exposure equivalent to a 1 mmol/L higher level of LDL-C and a 10 mmHg higher level of SBP was associated with a 118% higher risk of MI in the Chinese population (RR: 2.18, 95%CI: 1.39 – 3.41) and a 115% higher risk of MI in European populations (RR: 2.15, 95%CI: 1.69 – 2.73), as presented in ***Table 7.3***.

Table 7.8. Dose-response association of combined exposure to lower LDL-C and SBP in the Chinese population (GRS adjusted for BMI)

SBP quartile	Mean SBP mmHg	LDL quartile	Mean LDL-C mmol/L	N events	RR (95% CI)
1	127.61	1	2.10	112	Ref
2	130.23	1	2.09	108	1.03 (0.78 – 1.36)
3	131.93	1	2.11	127	1.20 (0.92 – 1.56)
4	135.68	1	2.06	167	1.56 (1.21 – 2.00)
1	128.08	2	2.39	123	1.12 (0.86 – 1.47)
2	130.28	2	2.35	138	1.29 (1.00 – 1.68)
3	132.36	2	2.21	135	1.25 (0.97 – 1.63)
4	135.16	2	2.35	137	1.34 (1.03 – 1.74)
1	127.43	3	2.41	112	1.06 (0.81 – 1.39)
2	129.99	3	2.40	114	1.06 (0.81 – 1.40)
3	132.40	3	2.34	127	1.18 (0.90 – 1.53)
4	135.11	3	2.38	144	1.33 (1.03 – 1.72)
1	127.15	4	2.44	122	1.14 (0.87 – 1.49)
2	130.22	4	2.51	176	1.72 (1.34 – 2.20)
3	132.23	4	2.44	150	1.48 (1.15 – 1.91)
4	134.99	4	2.52	156	1.52 (1.18 – 1.96)

Abbreviations. **SBP**: Systolic Blood Pressure; **LDL-C**: Low-Density Lipoprotein Cholesterol; **BMI**: Body Mass Index; **RR**: Risk Ratio.

Table 7.9. Dose-response association of combined exposure to lower LDL-C and SBP in the Chinese population (GRS unadjusted for BMI)

SBP quartile	Mean SBP mmHg	LDL quartile	Mean LDL-C mmol/L	N events	RR (95% CI)	P for trend
1	131.10	1	2.10	110	Ref	0.060
2	134.95	1	2.08	105	0.97 (0.74 – 1.28)	
3	137.78	1	2.11	130	1.23 (0.95 – 1.61)	
4	138.52	1	2.07	169	1.62 (1.26 – 2.08)	
1	132.96	2	2.35	137	1.28 (0.99 – 1.66)	
2	135.00	2	2.35	120	1.09 (0.83 – 1.42)	
3	135.74	2	2.31	158	1.56 (1.21 – 2.01)	
4	141.71	2	2.29	118	1.13 (0.86 – 1.48)	
1	132.49	3	2.38	122	1.15 (0.88 – 1.50)	
2	134.98	3	2.41	106	1.05 (0.80 – 1.39)	
3	134.77	3	2.41	123	1.07 (0.82 – 1.40)	
4	139.37	3	2.31	146	1.39 (1.07 – 1.80)	
1	131.81	4	2.42	150	1.44 (1.11 – 1.86)	
2	134.14	4	2.50	142	1.36 (1.05 – 1.76)	
3	133.89	4	2.51	160	1.58 (1.23 – 2.03)	
4	138.18	4	2.47	152	1.51 (1.17 – 1.95)	

Abbreviations. **SBP**: Systolic Blood Pressure; **LDL-C**: Low-Density Lipoprotein Cholesterol; **BMI**: Body Mass Index; **RR**: Risk Ratio.

Table 7.10. Dose-response association of combined exposure to lower LDL-C and SBP in the European populations

SBP quartile	Mean SBP mmHg	LDL quartile	Mean LDL-C mmol/L	N events	RR (95% CI)	P for trend
1	133.07	1	3.59	195	Ref	0.120
2	136.05	1	3.60	236	1.21 (1.00 – 1.47)	
3	138.35	1	3.60	266	1.35 (1.12 – 1.63)	
4	141.62	1	3.61	272	1.44 (1.19 – 1.73)	
1	132.68	2	3.68	200	1.03 (0.84 – 1.26)	
2	136.08	2	3.69	262	1.36 (1.12 – 1.64)	
3	138.15	2	3.70	252	1.33 (1.10 – 1.60)	
4	141.31	2	3.75	281	1.50 (1.25 – 1.81)	
1	132.73	3	3.77	209	1.05 (0.86 – 1.28)	
2	136.06	3	3.76	241	1.25 (1.03 – 1.51)	
3	138.30	3	3.77	255	1.36 (1.12 – 1.64)	
4	141.46	3	3.77	295	1.57 (1.31 – 1.89)	
1	132.49	4	3.85	219	1.13 (0.93 – 1.37)	
2	135.84	4	3.85	254	1.33 (1.10 – 1.61)	
3	137.83	4	3.85	248	1.32 (1.09 – 1.60)	
4	141.11	4	3.83	294	1.60 (1.33 – 1.93)	

Abbreviations. **SBP**: Systolic Blood Pressure; **LDL-C**: Low-Density Lipoprotein Cholesterol; **BMI**: Body Mass Index; **RR**: Risk Ratio.

7.4. Discussion

The factorial MR analysis in this chapter provides genetic evidence for independent causal relevance of plasma levels of LDL-C and levels of SBP on risk of MI in European and Chinese populations. The consistent absence of a statistically significant interaction between these two major risk factors for risk of MI demonstrated that the effects of lowering plasma levels of LDL-C and levels of SBP simultaneously are likely to be multiplicative. This confirms the observational findings from the previous chapter (Chapter 6), which also demonstrated independent associations of plasma levels of LDL-C and levels of SBP with risk of MI.

The findings of this study are also consistent with factorial MR studies conducted in populations of European ancestry using UKB by Ference et al. in 2019,[74] who reported that lifelong exposure to lower levels of both LDL-C and SBP was associated with independent, log-additive, and dose-dependent lower risks of major coronary events, with no evidence of any interactions. Similarly, a 2024 study by Li et al., also in UKB, used genetic scores to proxy the effects of specific drug classes and similarly reported independent and additive effects on CVD risk.[162] The current analyses used updated instruments from GLGC 2021 [133] and observed comparable results in UKB. The replication of the multiplicative (or log-additive) effect in the Chinese analysis with different genetic instruments and different baseline distributions of plasma lipid traits further confirmed the lack of interactions between lowering SBP and LDL-C on risk of MI. In the current meta-regression analysis, each 1 mmol/L higher LDL-C combined with 10 mmHg higher SBP was associated with a 118% higher risk of MI in the Chinese population and a 115% higher risk in Europeans, corresponding to approximately 54% lower MI risk for equivalent reductions in both factors. These results align with findings from Ference et al.,[74]

who reported that lifelong genetic exposure to 1 mmol/L lower LDL-C and 10 mmHg lower SBP was associated with a 78% lower risk of major coronary events and a 68% lower risk of cardiovascular death. Moreover, an individual participant data meta-analysis of three major polypill trials (TIPS-3, HOPE-3, and PolyIran) demonstrated that after a median follow-up of 5 years, a fixed-dose combination of at least two BP-lowering agents plus a statin reduced LDL-C by 0.58 mmol/L and SBP by 4.7 mmHg compared with control, corresponding to a 38% lower risk of CVD,[23] which is consistent with the predicted effect sizes from MR studies.

The burden of CVD in China is severe and increasing, driven in large part by a high and increasing prevalence of both hypertension and dyslipidaemia.[2,163] The observational and genetic analyses in European populations have consistently shown that these risk factors are independently associated with higher risk of MI.[74,162] Support for the clinical efficacy of polypill strategy with BP- and LDL-lowering components has been provided by RCTs such as HOPE-3[164] and TIPS-3.[155] The present study provides genetic evidence to confirm the findings of these large factorial RCTs that lowering BP and cholesterol simultaneously is an effective strategy for primary prevention of CVD.

This strategy is particularly relevant to the Chinese healthcare context. In China, the primary care infrastructure is fragmented,[165] and a simple, low-cost, fixed-dose combination therapy could be a cost-effective tool for CVD prevention.

These findings must be interpreted in the context of the limitations of this chapter. The validity of the results is contingent upon the three MR IV assumptions. While a restricted instrument for LDL-C was used to reduce pleiotropic effects on other lipid traits, the possibility of unmeasured pleiotropy on non-lipid traits cannot be fully

excluded. Furthermore, statistical power could be limited, particularly in the Chinese population. However, the analyses utilised all available data from the CKB to maximise power. In addition to the stratified analyses by quartiles of LDL-C and SBP, 2 × 2 factorial group analyses were performed as the main analysis, which retained greater power by reducing the number of groups within the constraints of the study. A further limitation is that MR reflects the effects of lifelong, modest differences in risk factor levels, which may not directly correspond to the effects of pharmacological interventions initiated later in life. The interpretation of the effect estimates from the current chapter should be limited to assessing the direction and interaction of effects rather than as quantitative predictions of trial outcomes. In addition, the analyses in this chapter did not account for drug-specific pathways. Therefore, the effect sizes reported here should be interpreted as evidence of the biological nature and independence of LDL-C and SBP effects on risk of MI, rather than as direct predictions of risk reduction achievable in a clinical trial setting, where the median follow-up duration in RCTs of combination therapies has typically been around 5 years.[23]

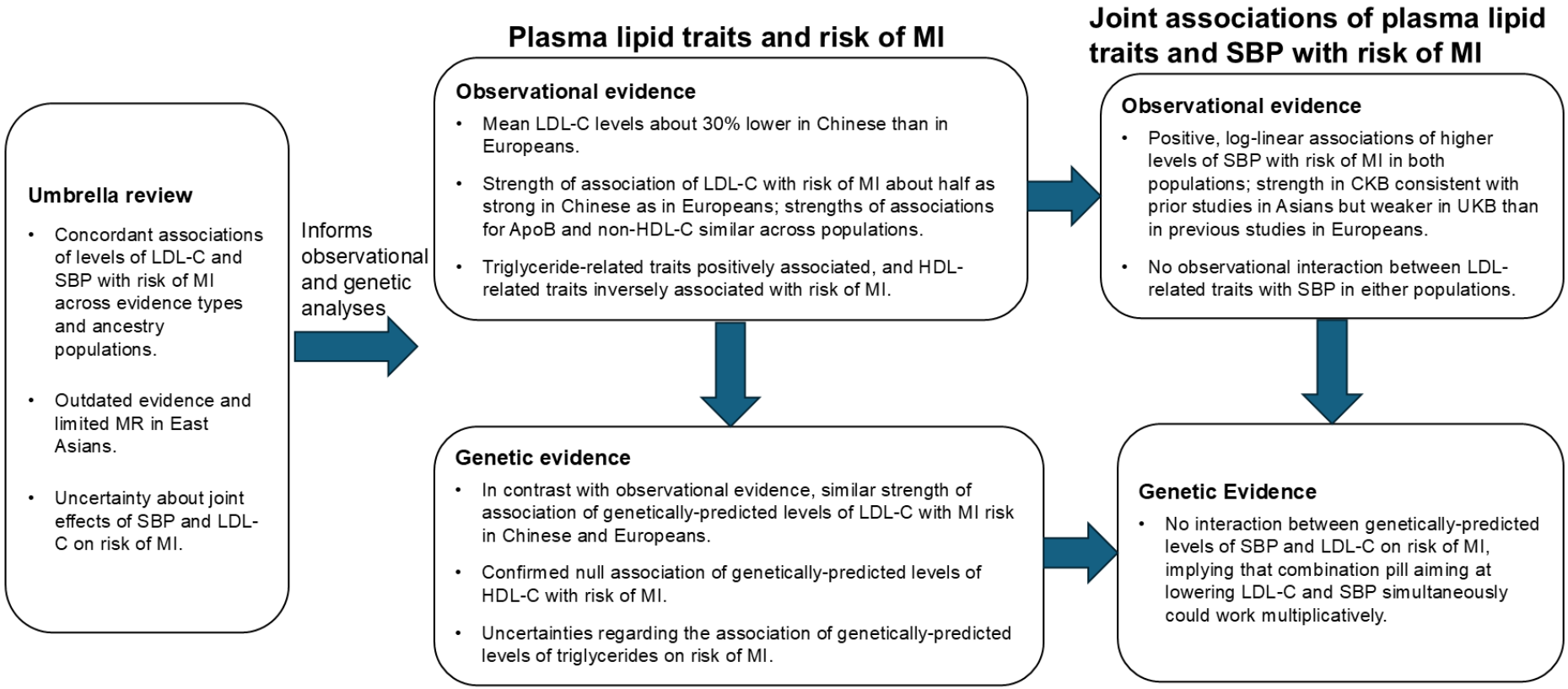
In conclusion, this chapter provides MR evidence in both Chinese and European populations that lifelong genetically predicted levels of LDL-C and SBP lower the risk of MI independently, with no evidence of departure from multiplicative (log-additive) effects when combined. Together with the observational findings from Chapter 6 and the individual MR analyses from Chapter 5, these results strengthen the notion that the effects of lowering SBP and LDL-C on risk of MI are independent. The findings support the potential for the joint lowering of LDL-C and SBP. Building on this evidence, the next section discusses the overall implications of these findings and their relevance for prevention strategies and clinical guidelines.

Chapter 8 Discussion

8.1. Summary of main findings

This thesis provided a comprehensive trans-ethnic investigation of the observational and genetic associations of major plasma lipid traits and SBP with risk of MI, using large-scale and contemporary biobanks in both the Chinese and European populations. The available randomised evidence of LDL-C and BP-lowering was also reviewed. Overall, this thesis provides evidence from observational and genetic studies, and randomised trials for the observational and causal relevance of plasma lipid traits and SBP for primary prevention of MI in Chinese and European populations. ***Figure 8.1*** shows the main findings of each chapter and how they relate to one another.

Figure 8.1. Summary of main findings



Abbreviations. **LDL-C**: Low-Density Lipoprotein Cholesterol; **SBP**: Systolic Blood Pressure; **MI**: Myocardial Infarction; **MR**: Mendelian Randomisation; **ApoB**: Apolipoprotein B; **HDL-C**: High-Density Lipoprotein Cholesterol.

The umbrella review outlined in Chapter 2 combined the available evidence from observational studies, RCTs, and MR on the separate and combined associations of SBP and LDL-C with IHD in East Asian and European populations. Higher plasma levels of LDL-C and higher levels of SBP each demonstrated strong, log-linear, and positive associations with higher risks of IHD across populations, albeit the effect sizes differed by study design. Moreover, studies have consistently demonstrated causal relevance of plasma levels of LDL-C for risk of MI in European populations. The available evidence for interactions between SBP and LDL-C is limited. Most factorial MR and trial evidence supported independence of these two risk factors, while some observational analyses suggested an interaction between SBP and LDL-C on risk of MI. The review identified key gaps in knowledge, including uncertainty about whether findings from earlier studies remain applicable in the contemporary populations with different risk profiles, disease prevalence, and medication use, limited understanding of the causal relevance of plasma lipids in East Asian populations, and insufficient evidence on how SBP and plasma lipids jointly influence the risk of MI. These knowledge gaps motivated trans-ethnic observational, MR, and factorial MR analyses in subsequent chapters in this thesis.

The observational analyses in Chapter 4 highlighted differences in mean levels of LDL-related lipid traits between the Chinese and European populations. While Chinese adults in CKB had approximately 30% lower mean plasma levels of LDL-C than Europeans in UKB, the strength of associations of plasma levels of LDL-C with risk of MI was about 50% weaker in the Chinese than European ancestry populations and appeared to be non-linear, with little evidence of an association below plasma levels of LDL-C at 3.0 mmol/L. In contrast, other atherogenic markers, particularly plasma levels of ApoB and non-HDL-C, were log-linearly and positively

associated with higher risk of MI with comparable strength of associations across both ancestry populations. This suggests that plasma levels of ApoB and non-HDL-C may be more reliable markers of MI risk compared to LDL-C, especially when baseline plasma levels of LDL-C are low. Furthermore, plasma levels of remnant-C were also more strongly associated with risk of MI than triglycerides in both ancestry populations, while the associations of plasma levels of HDL-C and ApoA1 with risk of MI also showed some trans-ethnic differences.

To investigate the causal relevance of plasma lipid traits with risk of MI, Chapter 5 employed an MR approach. These analyses provided evidence that higher levels of genetically-predicted LDL-C were associated with risk of MI, and importantly, that the strength of these associations was broadly similar in both Chinese and European populations. This suggests that the weaker observational association of plasma levels of LDL-C with risk of MI in Chinese compared to Europeans could be attributable to confounding or other biases. The MR analyses also clarified the roles of other lipids and found no evidence for causal relevance of HDL-C on risk of MI. The causal relevance of triglycerides on risk of MI also appeared to be mediated through other atherogenic lipoproteins, as the associations of genetically-predicted levels of triglycerides with risk of MI were attenuated to the null when using a restricted genetic score designed to minimise pleiotropic association with levels of LDL-C and HDL-C.

Chapter 6 extended the observational analyses to the associations of levels of SBP and risk of MI. The results confirmed the associations of levels of SBP with risk of MI were strong and log-linear in both Chinese and European populations. This is consistent with the accumulating body of evidence summarised in the umbrella review. More importantly, this chapter also examined possible interaction of plasma

levels of LDL-C and levels of SBP with risk of MI and found no evidence of interaction between SBP and LDL-C or ApoB with risk of MI, suggesting that these effects were independent and provide observational evidence for combined interventions as proposed in the combination pill strategy aiming at lowering levels of SBP and LDL-C for primary prevention of MI.

Finally, Chapter 7 utilised a factorial MR design to test the combined causal relevance of plasma levels of LDL-C and levels of SBP on risk of MI. These analyses confirmed the observational findings outlined in Chapter 6, demonstrating no interaction of genetically-predicted levels of LDL-C and SBP for risk of MI. The results also suggested a dose-response relationship of lower levels of LDL-C and SBP, where combined lifelong exposure to both lower levels of LDL-C and SBP was associated with a substantially lower risk of MI in both ancestry populations. These results provide genetic evidence to support the concept of combined BP-lowering and lipid-modifying medications for primary prevention of MI in trans-ancestry populations.

8.2. Population health and clinical implications

8.2.1. Optimal plasma lipid traits as risk markers of MI.

The findings from Chapter 4 contribute to the evidence regarding the optimal plasma lipid traits for reliable cardiovascular risk assessment. Plasma levels of LDL-C have been the primary target of lipid-modifying therapy and indicator for risk of MI since the measurement became widely available.[166] Recent evidence suggests that measures reflecting the total number of atherogenic lipoprotein particles, such as ApoB or non-HDL-C, may be better predictors of risk compared to LDL-C, which measures the content of cholesterol carried by LDL particles.[15,17,41,127] In settings where pre-treatment plasma levels of LDL-C are low on average, such as in the Chinese population, this thesis demonstrated that plasma levels of ApoB and non-HDL-C offered more reliable and consistent log-linear associations with risk of MI than plasma levels of LDL-C. This shows their reliability as MI risk markers across diverse populations and at lower ranges of LDL-C. These findings are particularly relevant for the Chinese population, where not only are plasma levels of LDL-related markers lower, but the particle size of LDL, as indicated by LDL-C/ApoB ratio in this study, can also differ. For risk assessment and potentially for monitoring therapeutic effects of RCTs, plasma levels of ApoB and non-HDL-C may be better markers of atherogenic particle burden. This is consistent with the ApoB particle model of atherosclerosis, which proposes that the numbers of ApoB-containing lipoproteins trapped in the arterial wall are the fundamental drivers of atherosclerosis, whereas the cholesterol mass they carry is more variable and less directly related to risk.[107] Furthermore, previous large-scale analyses show that even at the same LDL-C or non-HDL-C concentration, ApoB levels can vary widely, and individuals with higher ApoB have consistently higher rates of ASCVD.[107,127] Together, these results

strengthen the findings from this thesis that ApoB provides a more accurate and reliable measure of atherogenic risk than LDL-C. [105] This contrasts with the conclusions of a previous meta-analysis from the ERFC, which found that additional LDL-related measures, including ApoB, were associated with MI risk with similar strengths as non-HDL-C and LDL-C.[9] However, most ERFC studies were conducted in the last century, using older assays and were largely limited to European ancestry cohorts. The current analyses incorporate more contemporary data with harmonised laboratory methods, a more extended range of LDL-related lipid traits in diverse populations, thereby providing more reliable evidence that updates the previous conclusions about the relative importance of levels of ApoB and non-HDL-C versus LDL-C for risk of MI.

Current clinical guidelines in China,[67] as in many Western countries,[4,5,50] still prioritise LDL-C as the primary target for risk stratification and risk management. The findings of this thesis suggest that a greater emphasis on ApoB and non-HDL-C in guidelines may be warranted, though further genetic and randomised evidence is required.

Furthermore, plasma levels of remnant-C showed stronger associations with risk of MI than plasma levels of triglycerides in Chinese and European populations, reinforcing the importance of atherogenic remnant particles for risk stratification. This is consistent with previous studies showing that elevated remnant-C is associated with risk of MI and increased cardiovascular and all-cause mortality.[39,125] Unlike triglycerides, which fluctuate with dietary intake, remnant-C directly reflects the cholesterol content of triglyceride-rich lipoproteins that can penetrate the arterial wall, deposit cholesterol, and promote atherosclerosis.[39] Genetic studies have further supported its causal role, demonstrating that higher levels of remnant-C were

associated with an increased risk of IHD, independent of HDL-C levels.[44] These findings highlight remnant-C as a potentially important marker of triglyceride-rich lipoproteins. Current lipid management and CVD risk assessment frameworks[4,67] that rely on triglycerides would benefit from incorporating remnant-C as a more direct measure of triglyceride-rich lipoproteins. The findings also suggest that remnant-C may be a useful target for residual risk of MI beyond LDL-C lowering, in Chinese as well as European populations.

8.2.2. Generalisability of lipid-modifying and BP-lowering therapies.

The trans-ethnic MR evidence suggests the generalisability of LDL-C-lowering benefits for risk of MI to the Chinese population despite lower pre-treatment plasma levels of LDL-C. These results contribute to the discussion about whether we can extrapolate results of statin trials conducted in Western populations to the Chinese population. The findings support prioritising LDL-C–lowering therapy for primary prevention of MI according to absolute cardiovascular risk, rather than baseline LDL-C levels. This is consistent with the conclusions of the CTT Collaboration, which advocate more intensive LDL-C lowering in individuals at highest risk of cardiovascular disease.[120] As highlighted by Jackson et al., reliance on categorical labels such as hypertension or hypercholesterolaemia is of limited clinical value.[167] This concept of shifting the focus to initiation of treatment from levels of SBP and LDL-C to overall risk of cardiovascular disease is reflected in the changes of RCT inclusion criteria. Previous clinical trials for primary prevention for MI using BP-lowering and LDL-C lowering medication selected participants with higher levels of each exposure. For example, West of Scotland Coronary Prevention Study (WOSCOPS), conducted in the 1990s, was the first primary prevention trial of lipid-

lowering drug therapy to show a reduction in all-cause mortality. The trial selected participants with very high plasma levels of cholesterol, with the initial screening criteria requiring total cholesterol of at least 6.5 mmol/L.[168] Similarly, early BP-lowering trials focused on treating levels,[36] but more recent trials prioritised selection of participants with high or medium absolute risk of CVD, rather than high levels of SBP. For example, the Systolic Blood Pressure Intervention Trial (SPRINT) provided support for the benefit of treatment of SBP to a target of less than 120 mmHg compared to the traditional treatment to a target of less than 140 mmHg.[169] More recent factorial trials targeting BP- and LDL-C-lowering simultaneously, such as International Polycap Study-3 (TIPS-3), enrolled participants based on intermediate global cardiovascular risk, rather than high individual levels.[155] This shift is further reflected in the updated 2025 American guidelines for BP management, where the treatment thresholds for primary prevention for CVD were altered from the higher cut-offs used in the 2017 guidelines (SBP ≥ 140 mmHg or DBP ≥ 90 mmHg) to the broader criteria that recommend initiation of BP-lowering treatment at SBP ≥ 130 mmHg or DBP ≥ 80 mmHg after lifestyle intervention, thereby expanding eligibility to a wider population at risk.[170]

Overall, this thesis provides support for a public health strategy that prioritises the initiation of lipid-modifying therapy in individuals at high absolute risk of CVD rather than solely those with high LDL-C levels. Despite lower mean pre-treatment levels of LDL-C in the Chinese population compared with European populations, the observational associations of plasma levels of ApoB and non-HDL-C with risk of MI and the associations of genetically-predicted levels of LDL-C with risk of MI were consistent across both ancestry populations. These findings are also consistent with current European Society of Cardiology (ESC) guidelines, which recommend

initiation of statins for primary prevention of IHD based on overall risk based on Systematic Coronary Risk Evaluation 2 model, rather than levels of plasma lipids.[171] The Chinese guidelines also support initiating statins for individuals at high absolute ASCVD risk, rather than only high LDL-C levels.[67]

The genetic evidence from Chapter 5 provides support for the causal relevance of LDL-C for risk of MI in both Chinese and European ancestry populations. The comparable effect sizes observed across populations indicate broadly consistent associations of genetically-predicted levels of LDL-C with risk of MI, suggesting similar proportional effects of LDL-C lowering on the risk of MI. These findings support the importance of LDL-C reduction for primary prevention in the Chinese population and are consistent with the current Chinese guidelines for lipid management, which advocate wider use of statin therapy to control plasma lipid levels.[67] In addition, current Chinese guidelines recommend PCSK9 inhibitors as alternatives to first-line statin therapy combined with lifestyle modification for lipid management, and similarly, the ESC guidelines recommend the use of PCSK9 inhibitors when statins cannot be used.[5,67] Recently, there has been growing interest in PCSK9 injection therapy, which has been shown to safely reduce plasma levels of LDL-C even with existing use of statins.[172,173] An oral antisense oligonucleotide targeting PCSK9 to lower lipid levels has also been proposed.[174] Future RCTs in China may consider incorporating such novel lipid-lowering approaches that further modify plasma lipid levels, given the proven benefits of LDL-C reduction even at low levels. However, it is worth noting that a large-scale RCT would be unlikely to focus on MI alone. Instead, trials typically use composite outcomes such as major vascular events (MVE), for example, as examined by the CTT Collaboration and defined as the first occurrence of coronary death, non-fatal MI, coronary revascularisation, or

stroke.[120] Previous analyses from CKB by Sun et al. showed that LDL-C was positively associated with ischaemic stroke and inversely associated with intracerebral haemorrhage, with findings confirmed by genetic analyses and LDL-C-lowering trials.[123] Importantly, they demonstrated that the overall benefits of lowering LDL-C for ischaemic stroke and major vascular events outweigh the increase in haemorrhagic stroke risk.[123] Taken together, these results indicate that LDL-C lowering would be expected to confer a net benefit for the prevention of overall stroke and MVE in China.

8.2.3. Multi-factorial intervention for the primary prevention of MI.

The consistent finding of independent effects for SBP and LDL-C across observational and genetic analyses has direct clinical relevance and provides strong support for combined interventions aiming to lower levels of SBP and LDL-C simultaneously. The absence of interaction implies that the proportional benefits of lowering BP apply regardless of baseline LDL-C, and vice versa.

Given the log-linear associations for MI extending across the usual ranges of both risk factors in both populations, treatment policies should target both risk factors concurrently to achieve multiplicative risk reductions. This is consistent with previous recommendations that integrating multiple risk factors into absolute risk prediction for CVD provides a more effective and clinically meaningful basis for prevention than treating individual factors in isolation.[167] However, as mentioned, a large-scale RCT would be unlikely to focus on MI alone. Previous analyses from CKB demonstrated the overall benefits for CVD with lipid-lowering.[73] Similarly, recent analyses in CKB demonstrated that both observational and genetic analyses showed log-linear positive associations of SBP with major vascular events and other major CVD types

in the range of 120 to 170 mmHg, with particularly strong associations observed in younger populations.[57]

This evidence supports combination therapy approaches, a fixed-dose combination of multiple preventive medications originally proposed by Wald and Law as a public health strategy to reduce the burden of CVD.[21] The original idea proposed by Wald and Law included modifying levels of LDL-C, SBP, serum homocysteine (by folic acid), and platelet function (by aspirin). Evidence from trials in the 1990s evaluating aspirin for primary prevention suggested modest reductions in MI incidence but an increased risk of bleeding. To balance these risks and benefits, earlier international guidelines recommended aspirin only for individuals at very high 10-year cardiovascular risk. However, more recent large-scale trials have demonstrated little or no net benefit of aspirin for primary prevention and, in some cases, potential harm. Consequently, most recent guidelines now recommend against using aspirin for primary CVD prevention.[175] Similarly, although the initial observational studies suggested that lowering plasma levels of homocysteine might reduce cardiovascular risk,[176] subsequent meta-analyses of RCTs have shown no clear evidence that routine folic acid supplementation reduces cardiovascular events.[177] Thus, this thesis concentrated on LDL-C and SBP as the most important modifiable causal risk factors for CVD prevention.

The observational results from Chapter 6 and the subsequent factorial MR analyses from Chapter 7 provide observational and genetic epidemiological support for fixed-dose combination approaches that jointly lower levels of SBP and plasma levels of LDL-C. The premise of this approach is that its components act on independent causal pathways, leading to a combined benefit that is multiplicative.[21] By demonstrating that genetically-predicted levels of LDL-C and SBP are associated

with risk of MI independently and without interaction, this research provides causal evidence for this assumption in diverse populations.

This genetic validation complements evidence from recent landmark RCTs on primary prevention of CVD, such as HOPE-3,[164] TIPS-3,[155] and PolyIran,[178] which have demonstrated the efficacy of this approach (***Table 8.1***). While these RCTs test specific drug combinations over a limited duration of follow-up (typically median follow-up is 5 years), the MR analyses in this thesis model the lifelong effects of modifying LDL-C and SBP, thereby offering a broader perspective on the potential benefits of risk factor reduction.[21] In addition, whereas the trial populations had mean ages of about 60 years and mean baseline pre-treatment levels of SBP above 130 mmHg and LDL-C levels above 3 mmol/L, the observational and genetic analyses here cover a wider age range (30–80 years) and broader pretreatment distributions of risk factors. For example, in CKB, women aged 30–39 years had mean SBP levels below 120 mmHg, while participants in the oldest age groups had mean SBP levels above 140 mmHg. Similarly, mean LDL-C levels were as low as ~2 mmol/L in the youngest groups in CKB. The current analyses also involve substantially larger numbers of participants than the ~6,000 included in each of the individual trials.

By incorporating expanded age ranges, lifelong risk factor effects, and a wider range of LDL-C and SBP levels, this thesis provides complementary evidence to RCTs and helps inform the design of future trials on combination therapy. The suggestion of independent effects of lowering LDL-C and SBP on risk of MI indicates that such interventions could be extended to a broader proportion of the population, particularly those above age 50 years where CVD risk rises substantially.[179] Future trials aiming to inform population-level implementation of combination therapy should therefore consider expanding the eligibility criteria, including individuals aged above

50 years and a broader range of baseline risk factor levels, and inclusion of more diverse ancestry populations to maximise the generalisability of subsequent trials to the relevant populations in China and Europe.

Table 8.1. Characteristics of three major polypill randomised clinical trials

	TIPS-3[155]	**HOPE-3**[164]	**PolyIran**[178]
Study location	Bangladesh, Canada, Colombia, India, Indonesia, Malaysia, Philippines, Tanzania, Tunisia	Argentina, Australia, Brazil, Canada, China, Colombia, Czech Republic, Ecuador, Hungary, India, Israel, South Korea, Malaysia, Netherlands, Philippines, Russia, Slovakia, South Africa, Sweden, UK, Ukraine	Iran
Trial population	5713 participants without known vascular disease but at intermediate CVD risk*	6348 participants without known CVD but at intermediate CVD risk,* and were randomly assigned to double active group or double placebo group	Participants without a history of vascular disease (n=6101)
Intervention	2 × 2 × 2 factorial design: daily oral polypill consisting of simvastatin 40 mg, ramipril 10 mg, atenolol 100 mg, and hydrochlorothiazide 25 mg; daily aspirin 75 mg; monthly oral vitamin D 60 000 IU	2 × 2 factorial design: daily oral rosuvastatin 10 mg; daily oral candesartan 16 mg and hydrochlorothiazide 12·5 mg	Daily oral polypill consisting of atorvastatin 20 mg, hydrochlorothiazide 12·5 mg, enalapril 5 mg (or valsartan 40 mg), and aspirin 81 mg
Controls	Matching placebos	Matching placebos	Minimal care (blood pressure measurement and risk factor counselling)
Median follow-up (IQR), years	4·4 (3·2–5·9)	5·5 (5·1–6·2)	5·0 (4·9–5·0)
Mean age (SD), years	63·9 (6·6)	65·7 (6·3)	59·3 (6·7)
Mean SBP (SD), mmHg	144·5 (16·8)	138·1 (14·7)	130·8 (22·4)
Mean LDL-C (SD), mmol/L	3.1 (1.1)	3.3 (0.9)	3.0 (0.9)

*Based on estimated cardiovascular risk by the Framingham Risk Score. Table was reproduced from Joseph et al.[23]

Abbreviations. **TIPS-3**: International Polycap Study-3; **HOPE-3**: Heart Outcomes Prevention Evaluation-3; **CVD**: Cardiovascular Disease; **IQR**: Interquartile Range; **SD**: Standard Deviation; **SBP**: Systolic Blood Pressure; **LDL-C**: Low-Density Lipoprotein Cholesterol.

Recent modelling further underscores the potential impact of scaling up fixed-dose combination (single-pill combination, SPC) therapies worldwide. Watkins et al. projected that widespread adoption of SPCs for primary and secondary prevention could prevent up to 72 million deaths and 130 million nonfatal CVD events by 2050, with the largest absolute benefits observed in South and East Asia because of the large population size. This reinforces the cost-effectiveness and population-level impact of such interventions, particularly in low- and middle-income settings such as China, where uptake of preventive drugs remains suboptimal.[180]

In both UKB and CKB, mean levels of SBP exceeded 130 mmHg in men and women despite some use of antihypertensive medication. According to the 2025 AHA guidelines, such individuals already meet the criteria for treatment initiation for BP.[170] Modelling studies suggest that in the UK, adoption of a cardiovascular polypill could further prevent about 47 nonfatal cardiovascular events and 10 cardiovascular deaths per 1,000 treated over 10 years compared with multiple monotherapy.[22] While no direct analysis on primary prevention of CVD was identified for the cost-effectiveness of a polypill strategy in China, an analysis on secondary prevention of CVD projected that a polypill can further prevent 52 major adverse cardiovascular events per 1,000 treated over 5 years compared with current monotherapy approach.[181]

In addition, SPCs consistently improve adherence compared with usual care. Meta-analyses of randomised trials from the Single Pill to Avert Cardiovascular Events (SPACE) collaboration, which focused mainly on primary prevention, reported a 58% improvement in adherence for polypills compared with monotherapies.[62] Improved adherence is expected to amplify the population-level benefits of combination therapy in both China and the UK.

Taking the combination therapy concept further, pharmacological interventions can be implemented alongside lifestyle changes. Approximately half of CVD is attributable to modifiable behavioural and lifestyle factors,[3] and the AHA's Life's Essential 8 framework highlights the importance of targeting diet, physical activity, smoking, sleep, and other behaviours alongside medical therapy.[182] Combining lifestyle modification with SPCs could therefore provide complementary risk reductions, enhancing both population-level prevention and long-term adherence in drug treatment. In light of the strong evidence for the effectiveness of lowering LDL-C and SBP, including the findings presented in this thesis, and given the large attributable burden of modifiable lifestyle factors, future multi-factorial prevention strategies could consider integrated approaches that combine LDL- and BP-lowering medications with effective lifestyle interventions.

8.3. Strengths and limitations

8.3.1. Strengths

Key strengths of this thesis include the use of two very large, contemporary biobanks with harmonised assays, reliable outcome ascertainment, and standardised covariates in Chinese and European populations, enabling direct cross-population comparisons.

Correction for regression dilution in observational analyses using repeat measures allowed for the estimation of associations with usual long-term exposure levels, which provides a more accurate reflection of real risk estimates than single baseline measurements.

The thesis also benefited from the application of standardised study protocols and extensive quality assurance procedures across study centres, ensuring the reliability of exposure and outcome measurements. Linkage to national health insurance systems and disease registries enabled comprehensive and virtually complete follow-up, with accurate and validated classification of cardiovascular outcomes, such as MI used in this thesis.

The analytical framework combining observational analyses with MR and factorial MR, alongside comparison with RCT evidence from the umbrella review, enhanced the robustness of findings.

In addition, the diverse population and range of exposures captured in CKB and UKB allowed exploration of associations across different environmental, genetic, and metabolic contexts. This diversity provided unique insights into the consistency of risk factor effects across ranges and ancestry groups. The harmonisation of

laboratory assays, particularly for lipid biomarkers, ensured comparability of results between the two cohorts.

8.3.2. Limitations

First, sample and study design differed across analyses, with a nested case-control subset in CKB for plasma lipid traits and a prospective cohort analysis for SBP, which may introduce heterogeneity in study design and statistical estimates, for example, OR versus RR. However, outcomes in CKB accrued during follow-up from baseline despite the nested case-control design, so the results should be broadly comparable to those derived from a prospective cohort study.

Second, this thesis focused primarily on MI as the outcome. MI is only one component of IHD, and IHD itself is only one part of the wider spectrum of CVD. This choice was partly constrained by data availability, as only MI outcomes, not IHD, were available in the CKB biochemistry nested case-control subset. Previous large-scale studies have generally demonstrated concordance of associations of CVD risk factors with both MI and IHD.[183-185] Stroke outcomes, including both ischaemic and haemorrhagic stroke, have been examined in detail in previous CKB analyses.[57,73]

Third, selection bias is probably present in both CKB and UKB, which are not representative of their underlying populations.[119,186] Internal validity of exposure-outcome associations is less sensitive to such bias but the extrapolation of the effect sizes of the associations to the general populations might be inaccurate. In CKB, participants were recruited through non-probabilistic sampling with a relatively low participation rate (~30%),[99] and both cohorts are subject to volunteer effect, whereby those who participate in research studies tend to be healthier and of higher

socioeconomic status than the general population.[187] Related to this, by restricting the analyses to White British individuals, the study assumes homogeneity across European populations. However, considerable genetic, environmental, and demographic heterogeneity exists even within European ancestry populations, and effect sizes observed in a White British sample may not extrapolate well to, for example, southern or eastern European populations.[188]

Fourth, statistical power was limited for some MR analyses, particularly in the Chinese population, for example, when studying the associations of levels of genetically-predicted triglycerides with risk of MI using restricted genetic instruments, resulting in wider CIs and greater uncertainty. Similarly, factorial MR analyses stratified by genetic scores were underpowered in some subgroups, limiting the precision of interaction tests. Replication in larger East Asian datasets is therefore needed.

Fifth, inherent limitations of MR remain, including potential violation of IV assumptions and differences between lifelong genetically mediated exposure and effects of pharmacological interventions initiated later in life.[129,148] Some instruments for East Asian analyses included CKB participants, which can bias results towards observational estimates, although extensive sensitivity analyses, including jack-knife approaches to avoid sample overlap, were conducted.[146] Although restricted instruments for LDL-C were used to minimise pleiotropic associations with other lipid traits, unmeasured pleiotropy, including pleiotropic effects of SBP instruments, cannot be ruled out, and this may lead to violation of IV assumptions.

Sixth, only limited forms of MR analyses were conducted. A key question for future work concerns the specific pathways through which lipid modification and BP

reduction influence cardiovascular outcomes.[135,189] Drug-target MR would help distinguish the effects of pharmacological mechanisms, but this thesis considered instruments proxying overall lipid and SBP levels. Furthermore, while observational analyses indicated possible non-linear associations of LDL-C with MI risk in Chinese populations, this thesis did not conduct non-linear MR analyses. Although extensions to linear MR have been developed to explore non-linear exposure–outcome relationships, they remain uncertain in terms of the validity of the methodologies.[190-192] For example, residual-based stratification methods rely heavily on assumptions such as homogeneity and monotonicity, and recent evaluations have shown that violations of these assumptions can produce biased or implausible results, even in negative control settings. Alternative approaches, such as doubly ranked methods, also face challenges, and the origins of these biases are not yet fully understood.[190-192] Given these limitations and the ongoing debate over methodological validity, non-linear MR was not conducted in this thesis.

8.4. Future research

8.4.1. Replication of results in larger and more diverse databases

Expanding these analyses to additional East Asian and multi-ethnic cohorts with harmonised plasma lipid traits, repeated measures, and adjudicated outcomes will increase statistical power to refine causal estimates. This is particularly important for triglycerides and remnant-related traits, where the present analyses had limited precision. In this thesis, power was limited for restricted genetic analyses of triglycerides, and the restricted analyses produced findings that were not fully consistent with multivariable MR, highlighting the need for replication and further investigation. Moreover, while the observational analyses examined novel lipid traits such as ApoB, ApoA1, and remnant-C, these were not included in the MR analyses due to the absence of sufficiently strong instruments. As GWAS continue to expand, updated instruments may allow robust genetic analyses of these traits in future research.[193] Larger datasets will also enable more detailed exploration of potential heterogeneity by age, sex, and regional practice patterns, which could further inform targeted prevention strategies.

8.4.2. Drug target MR

Future research could employ drug target MR, which focuses on variants in or near genes encoding therapeutic targets.[148,189] For example, variants in HMGCR (the statin target) and PCSK9 will proxy LDL-C lowering [194] and variants in antihypertensive target genes (e.g., ADRB1) will proxy BP-lowering.[195] This approach can complement biomarker-based MR approaches in the current thesis by more closely approximating the effects of specific classes of lipid-lowering and BP-lowering agents. Such analyses would help to prioritise targets for combination

therapy and anticipate the treatment effects of specific drug combinations in diverse ancestry populations.

Some analyses are already feasible within the current databases. For example, missense variants in lipid and BP drug target regions can be studied in the Chinese population. Previous analyses focused on the CETP region for vascular and non-vascular diseases[196] and in the PCSK9 region for LDL-C lowering.[197] Further analysis can study novel lipid-modifying regions and apply factorial MR analysis to study the effects of additional lipid-modifying pathways on statins (HMGCR). For drug target regions lacking strong predictive variants, MR using summary statistics from large consortia should be used in place of GRS approaches to increase power. Implementation of factorial drug target MR to mimic combination therapy poses additional challenges, even in European populations with larger numbers of events, since this would require stratification by two or more independent genetic scores. Although weights for constructing these scores can be extracted externally, factorial MR analyses need individual-level data to properly evaluate joint effects. Achieving sufficient power will therefore depend on collaborative efforts to combine multiple cohorts with harmonised observational and genetic data.

8.4.3. Mediation and analysis of underlying mechanisms

Further work is needed to clarify the causal pathways by which triglyceride-rich lipoproteins, remnant-C, APOC3, and ApoB-containing particles influence MI risk. Methodologies such as multivariable MR and formal mediation analysis frameworks could be used to explore the direct and indirect effects of these lipid fractions while minimising the influence of pleiotropy.[198,199] This would provide a more in-depth understanding of the specific atherogenic components of dyslipidaemia beyond LDL-

C. In particular, MVMR analyses with sufficient power can estimate the direct effects of each lipid trait on risk of MI after accounting for their correlations, while standard MR provides total effects. Comparing these estimates enables quantification of the proportion of MI risk mediated through specific lipid pathways. Two-step or network MR approaches can also be applied to evaluate sequential pathways (e.g., MR analysis on triglycerides and remnant-C, then on risk of MI), allowing formal estimation of mediated effects.[136,200] However, current GWAS for relatively novel lipid traits, such as ApoB and remnant-C, remain less well powered than those for conventional lipids.[193] As large-scale meta-analyses and updated GWAS with improved coverage and sample sizes become available, these approaches are likely to become increasingly feasible.

8.4.4. Novel lipids and other markers to address residual risk of MI

Beyond conventional lipids, remnant-related measures and other particle metrics, for example, lipoprotein(a), may capture a significant component of residual cardiovascular risk.[201-203] Integrating these further markers with established ones, such as LDL-C and non-HDL-C, could improve risk assessment, particularly in populations where LDL-C is low on average, such as in China. However, the benefits of lowering lipoprotein(a) would be limited to a subset with high plasma levels of lipoprotein(a).[203] Future studies should assess whether these markers provide incremental predictive value beyond ApoB or non-HDL-C, which this thesis has shown to be reliable predictors for risk of MI in observational analyses. Recently, the availability of large-scale proteomics datasets has enabled more detailed pathway analyses. For example, recent studies have shown that proteomic panels can significantly improve risk prediction of ischaemic heart disease in both Chinese and

European populations and can also help to identify novel drug targets through integrated proteomic–genetic analyses.[204,205] Such approaches can be implemented to explore mechanistic pathways, for example by using broader proteomic profiling to identify proteins that mediate the effects of lipids on MI risk, and to prioritise therapeutic targets beyond traditional lipid biomarkers through the discovery of novel proteomics-based markers.

8.4.5. Modifiable factors for primary prevention of MI

Given the established independence of the SBP and LDL-related causal pathways, evaluating pragmatic, risk-based, multi-component strategies should be a research priority. This includes the design and implementation of large-scale RCTs of fixed-dose combinations (polypills) in Chinese primary care settings to assess real-world effectiveness, adherence, and cost-effectiveness.

Furthermore, future research could also explore associations beyond SBP and LDL-related causal pathways and examine effects of other major modifiable cardiovascular risk factors, including lowering levels of BMI, smoking, and alcohol intake and increasing physical activity.[3,206] These results can inform a more comprehensive public health primary prevention strategy for CVD in diverse populations. For example, a recent factorial MR study using UKB demonstrated that genetically predicted lowering of Lp(a) had a similar magnitude of effect in reducing cardiovascular events whether considered alone or in combination with LDL-C reduction or lifestyle improvements, including lower smoking and alcohol consumption, higher physical activity, and lower BMI.[207]

In line with the American Heart Association's Life's Essential 8 framework, integrating lifestyle modification, including healthy diet, adequate sleep, smoking cessation, regular physical activity, and weight management, with pharmaceutical interventions such as statins and antihypertensives could yield greater and more sustainable risk reductions than either approach alone.[182] Future research should therefore evaluate pragmatic strategies that combine combination therapies with lifestyle modifications, similar to the analyses on Lp(a), LDL-C and lifestyle factors mentioned earlier,[207] to promote improvements in these lifestyle factors, thereby addressing both biological and behavioural determinants of CVD.

8.5. Conclusions

This thesis provides a trans-ethnic investigation into the associations of plasma lipid traits and systolic blood pressure with risk of MI, integrating observational and genetic evidence from two large contemporary biobanks in China and the UK. The analyses demonstrated that higher levels of LDL-C and SBP are each strongly and causally associated with higher risk of MI, with consistent evidence across populations despite about 30% lower levels of LDL-C in the Chinese ancestry population. LDL-related lipid traits, including ApoB and non-HDL-C, emerged as more reliable markers of MI risk than LDL-C, particularly in settings where LDL-C levels are low. These findings update the conclusions of previous meta-analyses, in which ApoB and non-HDL-C showed associations with risk of MI that were similar to those of LDL-C. The MR analyses confirmed the causal relevance of LDL-C in the Chinese population and clarified the limited roles of HDL-C and triglycerides, while factorial MR established the independent effects of levels of LDL-C and SBP on risk

of MI across ancestry populations, showing that combined lifelong exposure to lower levels of both traits substantially reduces MI risk.

Together, these findings contribute to the growing body of evidence supporting risk-based, multi-factorial strategies for primary prevention of MI and CVD in general. By addressing key evidence gaps in East Asian populations and triangulating evidence across observational, genetic, and trial results, the thesis strengthens the case for combination therapy approaches that target both lipids and BP. The results have direct implications for primary prevention of MI and the design of future large-scale RCTs, particularly in China. This thesis highlights the importance of integrated prevention policies that extend beyond single-risk-factor thresholds to a comprehensive, risk-based approach, with the potential to inform clinical guidelines and improve population health across ethnicities.

References

1. Thygesen K, Alpert JS, Jaffe AS, et al. Fourth universal definition of myocardial infarction (2018). *Circulation.* 2018; **138**(20): e618-e51.

2. Liu S, Li Y, Zeng X, et al. Burden of cardiovascular diseases in China, 1990-2016: findings from the 2016 global burden of disease study. *JAMA Cardiol.* 2019; **4**(4): 342-52.

3. Global Cardiovascular Risk Consortium. Global effect of modifiable risk factors on cardiovascular disease and mortality. *N Engl J Med.* 2023; **389**(14): 1273-85.

4. Fegers-Wustrow I, Gianos E, Halle M, Yang E. Comparison of American and European Guidelines for Primary Prevention of Cardiovascular Disease. *J Am Coll Cardiol.* 2022; **79**(13): 1304-13.

5. Mach F, Baigent C, Catapano AL, et al. 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk: The Task Force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and European Atherosclerosis Society (EAS). *Eur Heart J.* 2020; **41**(1): 111-88.

6. Doi T, Langsted A, Nordestgaard BG. Lipoproteins, cholesterol, and atherosclerotic cardiovascular disease in East Asians and Europeans. *J Atheroscler Thromb.* 2023; **30**(11): 1525-46.

7. Jia Y, Wen J, Qureshi R, et al. Effect of redundant clinical trials from mainland China evaluating statins in patients with coronary artery disease: cross sectional study. *BMJ.* 2021; **372**: n48.

8. Chinese Society of Cardiology of Chinese Medical Association, et al. Chinese guideline on the primary prevention of cardiovascular diseases. *Cardiol Discov.* 2021; **1**(02): 70-104.

9. Emerging Risk Factors Collaboration. Major Lipids, Apolipoproteins, and Risk of Vascular Disease. *JAMA.* 2009; **302**(18): 1993-2000.

10. Cao J, Nomura SO, Steffen BT, et al. Apolipoprotein B discordance with low-density lipoprotein cholesterol and non-high-density lipoprotein cholesterol in relation to coronary artery calcification in the Multi-Ethnic Study of Atherosclerosis (MESA). *J Clin Lipidol.* 2020; **14**(1): 109-21.e5.

11. Hagström E, Steg PG, Szarek M, et al. Apolipoprotein B, residual cardiovascular risk after acute coronary syndrome, and effects of alirocumab. *Circulation.* 2022; **146**(9): 657-72.

12. Johannesen CDL, Mortensen MB, Langsted A, Nordestgaard BG. Apolipoprotein B and non-HDL cholesterol better reflect residual risk than LDL cholesterol in statin-treated patients. *J Am Coll Cardiol.* 2021; **77**(11): 1439-50.

13. Kim CW, Hong S, Chang Y, Lee JA, Shin H, Ryu S. Discordance Between Apolipoprotein B and Low-Density Lipoprotein Cholesterol and Progression of Coronary Artery Calcification in Middle Age. *Circ J.* 2021; **85**(6): 900-7.

14. Lawler PR, Akinkuolie AO, Ridker PM, et al. Discordance between Circulating Atherogenic Cholesterol Mass and Lipoprotein Particle Concentration in Relation to Future Coronary Events in Women. *Clin Chem.* 2017; **63**(4): 870-9.

15. Marston NA, Giugliano RP, Melloni GE, et al. Association of apolipoprotein B–containing lipoproteins and risk of myocardial infarction in individuals with and without atherosclerosis: distinguishing between particle concentration, type, and content. *JAMA Cardiol.* 2022; **7**(3): 250-6.

16. Murphy A, Faria-Neto JR, Al-Rasadi K, et al. World heart federation cholesterol roadmap. *Glob Heart.* 2017; **12**(3): 179-97.e5.

17. Pencina MJ, D'Agostino RB, Zdrojewski T, et al. Apolipoprotein B improves risk assessment of future coronary heart disease in the Framingham Heart Study beyond LDL-C and non-HDL-C. *Eur J Prev Cardiol.* 2015; **22**(10): 1321-7.

18. Su X, Cai X, Pan Y, et al. Discordance of apolipoprotein B with low-density lipoprotein cholesterol or non-high-density lipoprotein cholesterol and coronary atherosclerosis. *Eur J Prev Cardiol.* 2022; **29**(18): 2349-58.

19. Welsh C, Celis-Morales CA, Brown R, et al. Comparison of conventional lipoprotein tests and apolipoproteins in the prediction of cardiovascular disease: data from UK Biobank. *Circulation.* 2019; **140**(7): 542-52.

20. Wilkins JT, Li RC, Sniderman A, Chan C, Lloyd-Jones DM. Discordance Between Apolipoprotein B and LDL-Cholesterol in Young Adults Predicts Coronary Artery Calcification. *J Am Coll Cardiol.* 2016; **67**(2): 193-201.

21. Wald NJ, Law MR. A strategy to reduce cardiovascular disease by more than 80%. *BMJ.* 2003; **326**(7404): 1419.

22. Fuster V, Sanz G, Castellano JM. The journey of the cardiovascular polypill from its conception to the WHO List of Essential Medicines. *Nat Cardiovasc Res.* 2025; **4**(3): 259-65.

23. Joseph P, Roshandel G, Gao P, et al. Fixed-dose combination therapies with and without aspirin for primary prevention of cardiovascular disease: an individual participant data meta-analysis. *Lancet.* 2021; **398**(10306): 1133-46.

24. Gu A, Yue Y, Desai RP, Argulian E. Racial and ethnic differences in antihypertensive medication use and blood pressure control among US adults with hypertension: the National Health and Nutrition Examination Survey, 2003 to 2012. *Circ Cardiovasc Qual Outcomes.* 2017; **10**(1): e003166.

25. Yin R, Yin L, Li L, et al. Hypertension in China: burdens, guidelines and policy responses: a state-of-the-art review. *J Hum Hypertens.* 2022; **36**(2): 126-34.

26. Sanderson E, Glymour MM, Holmes MV, et al. Mendelian randomization. *Nat Rev Methods Primers.* 2022; **2**(1): 6.

27. Thygesen K, Alpert JS, Jaffe AS, Simoons ML, Chaitman BR, White HD. Third Universal Definition of Myocardial Infarction. *Circulation.* 2012; **126**(16): 2020-35.

28. Bhatt DL, Lopes RD, Harrington RA. Diagnosis and treatment of acute coronary syndromes: a review. *JAMA.* 2022; **327**(7): 662-75.

29. van Oosterhout RE, de Boer AR, Maas AH, Rutten FH, Bots ML, Peters SA. Sex differences in symptom presentation in acute coronary syndromes: a systematic review and meta‐analysis. *J Am Coll Cardiol.* 2020; **9**(9): e014733.

30. Higuma T, Soeda T, Abe N, et al. A Combined Optical Coherence Tomography and Intravascular Ultrasound Study on Plaque Rupture, Plaque Erosion, and Calcified Nodule in Patients With ST-Segment Elevation Myocardial Infarction: Incidence, Morphologic Characteristics, and Outcomes After Percutaneous Coronary Intervention. *JACC Cardiovasc Interv.* 2015; **8**(9): 1166-76.

31. Roth GA, Mensah GA, Johnson CO, et al. Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the GBD 2019 study. *J Am Coll Cardiol.* 2020; **76**(25): 2982-3021.

32. Tan SCW, Zheng B-B, Tang M-L, Chu H, Zhao Y-T, Weng C. Global Burden of Cardiovascular Diseases and its Risk Factors, 1990–2021: A Systematic Analysis for the Global Burden of Disease Study 2021. *QJM.* 2025: hcaf022.

33. Li J, Li X, Wang Q, et al. ST-segment elevation myocardial infarction in China from 2001 to 2011 (the China PEACE-Retrospective Acute Myocardial Infarction Study): a retrospective analysis of hospital data. *Lancet.* 2015; **385**(9966): 441-51.

34. Murray CJ, Aravkin AY, Zheng P, et al. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet.* 2020; **396**(10258): 1223-49.

35. Reiner Ž. Statins in the primary prevention of cardiovascular disease. *Nat Rev Cardiol.* 2013; **10**(8): 453-64.

36. Law MR, Morris JK, Wald NJ. Use of blood pressure lowering drugs in the prevention of cardiovascular disease: Meta-analysis of 147 randomised trials in the context of expectations from prospective epidemiological studies. *BMJ.* 2009; **338(7705)**: 1245.

37. Hegele RA. Plasma lipoproteins: genetic influences and clinical implications. *Nat Rev Genet.* 2009; **10**(2): 109-21.

38. Libby P, Loscalzo J, Ridker PM, et al. Inflammation, immunity, and infection in atherothrombosis: JACC review topic of the week. *J Am Coll Cardiol.* 2018; **72**(17): 2071-81.

39. Sandesara PB, Virani SS, Fazio S, Shapiro MD. The Forgotten Lipids: Triglycerides, Remnant Cholesterol, and Atherosclerotic Cardiovascular Disease Risk. *Endocr Rev.* 2019; **40**(2): 537-57.

40. Holmes MV, Asselbergs FW, Palmer TM, et al. Mendelian randomization of blood lipids for coronary heart disease. *Eur Heart J.* 2015; **36**(9): 539-50.

41. Richardson TG, Sanderson E, Palmer TM, et al. Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: A multivariable Mendelian randomisation analysis. *PLoS Med.* 2020; **17**(3): e1003062.

42. Mihaylova B, Emberson J, Blackwell L, et al. The effects of lowering LDL cholesterol with statin therapy in people at low risk of vascular disease: meta-analysis of individual data from 27 randomised trials. *Lancet.* 2012; **380**(9841): 581-90.

43. Laufs U, Parhofer KG, Ginsberg HN, Hegele RA. Clinical review on triglycerides. *Eur Heart J.* 2019; **41**(1): 99-109c.

44. Nordestgaard BG, Varbo A. Triglycerides and cardiovascular disease. *Lancet.* 2014; **384**(9943): 626-35.

45. Do R, Willer CJ, Schmidt EM, et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nat Genet.* 2013; **45**(11): 1345-52.

46. Rader DJ, Hovingh GK. HDL and cardiovascular disease. *Lancet.* 2014; **384**(9943): 618-25.

47. Kannel WB, Dawber TR, Friedman GD, Glennon WE, McNamara PM. Risk factors in coronary heart disease: an evaluation of several serum lipids as predictors of coronary heart disease, the Framingham Study. *Ann Intern Med.* 1964; **61**: 888-99.

48. Barter PJ, Caulfield M, Eriksson M, et al. Effects of torcetrapib in patients at high risk for coronary events. *N Engl J Med.* 2007; **357**(21): 2109-22.

49. Schwartz GG, Olsson AG, Abt M, et al. Effects of dalcetrapib in patients with a recent acute coronary syndrome. *N Engl J Med.* 2012; **367**(22): 2089-99.

50. Whelton PK, Carey RM, Aronow WS, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Hypertension.* 2018; **71**(6): e13-e115.

51. Fuchs FD, Whelton PK. High Blood Pressure and Cardiovascular Disease. *Hypertension.* 2020; **75**(2): 285-92.

52. Prospective Studies Collaboration. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet.* 2002; **360**(9349): 1903-13.

53. Asia Pacific Cohort Studies Collaboration. Blood pressure and cardiovascular disease in the Asia Pacific region. *J Hypertens.* 2003; **21**(4): 707-16.

54. Satoh M, Ohkubo T, Asayama K, et al. Combined effect of blood pressure and total cholesterol levels on long-term risks of subtypes of cardiovascular death: Evidence for Cardiovascular Prevention from Observational Cohorts in Japan. *Hypertension* 2015; **65**(3): 517-24.

55. Lacey B, Lewington S, Clarke R, et al. Age-specific relation of blood pressure to vascular and non-vascular chronic disease in China: a prospective study of 0.5 million adults. *Lancet Glob Health.* 2018; **6**(6).

56. Cho SMJ, Koyama S, Ruan Y, et al. Measured Blood Pressure, Genetically Predicted Blood Pressure, and Cardiovascular Disease Risk in the UK Biobank. *JAMA Cardiol.* 2022; **7**(11): 1129-37.

57. Clarke R, Wright N, Walters R, Gan W, Guo Y, Millwood IY. Genetically Predicted Differences in Systolic Blood Pressure and Risk of Cardiovascular and Noncardiovascular Diseases: A Mendelian Randomization Study in Chinese Adults. *Hypertension.* 2023.

58. Wang K, Shi X, Zhu Z, et al. Mendelian randomization analysis of 37 clinical factors and coronary artery disease in East Asian and European populations. *Genome Med.* 2022; **14**(1): 63.

59. Rahimi K, Bidel Z, Nazarzadeh M, et al. Pharmacological blood pressure lowering for primary and secondary prevention of cardiovascular disease across different levels of blood pressure: an individual participant-level data meta-analysis. *Lancet.* 2021; **397**(10285): 1625-36.

60. Prospective Studies Collaboration. Blood cholesterol and vascular mortality by age, sex, and blood pressure: a meta-analysis of individual data from 61 prospective studies with 55 000 vascular deaths. *Lancet.* 2007; **370**(9602): 1829-39.

61. Thom S, Poulter N, Field J, et al. Effects of a Fixed-Dose Combination Strategy on Adherence and Risk Factors in Patients With or at High Risk of CVD: The UMPIRE Randomized Clinical Trial. *JAMA.* 2013; **310**(9): 918-29.

62. Baumgartner A, Drame K, Geutjens S, Airaksinen M. Does the Polypill Improve Patient Adherence Compared to Its Individual Formulations? A Systematic Review. *Pharmaceutics.* 2020; **12**(2): 190.

63. Attaei MW, Khatib R, McKee M, et al. Availability and affordability of blood pressure-lowering medicines and the effect on blood pressure control in high-income, middle-income, and low-income countries: an analysis of the PURE study data. *Lancet Public Health.* 2017; **2**(9): e411-e9.

64. Marcus ME, Manne-Goehler J, Theilmann M, et al. Use of statins for the prevention of cardiovascular disease in 41 low-income and middle-income countries: a cross-sectional study of nationally representative, individual-level data. *Lancet Glob Health.* 2022; **10**(3): e369-e79.

65. Tsui L, Ye P, Xu S, et al. Adverse drug reactions of statin therapy in China from 1989 to 2019: a national database analysis. *Eur J Hosp Pharm.* 2023; **30**(e1): e82-e9.

66. Xi L, Yan G, Jing L, et al. Underuse of statins in patients with atherosclerotic ischemic stroke in China. *Chin Med J.* 2012; **125**(10): 1703-7.

67. Li JJ, Zhao SP, Zhao D, et al. 2023 Chinese guideline for lipid management. *Front Pharmacol.* 2023; **14**: 1190934.

68. Aromataris E, Fernandez R, Godfrey CM, Holly C, Khalil H, Tungpunkom P. Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. *JBI Evid Implement.* 2015; **13**(3): 132-40.

69. Ioannidis JP. Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. *CMAJ.* 2009; **181**(8): 488-93.

70. DeFilippis AP, Nasir K, Blaha MJ. Myocardial infarction as a clinical end point in research: what are we really talking about? *Circ Res.* 2019; **124**(12): 1701-3.

71. Nabel EG, Braunwald E. A tale of coronary artery disease and myocardial infarction. *N Engl J Med.* 2012; **366**(1): 54-63.

72. Asia Pacific Cohort Studies Collaboration. Cholesterol, diabetes and major cardiovascular diseases in the Asia-Pacific region. *Diabetologia.* 2007; **50**(11): 2289-97.

73. Sun L, Clarke R, Bennett D, et al. Causal associations of blood lipids with risk of ischemic stroke and intracerebral hemorrhage in Chinese adults. *Nat Med.* 2019; **25**(4): 569-74.

74. Ference BA, Bhatt DL, Catapano AL, et al. Association of Genetic Variants Related to Combined Exposure to Lower Low-Density Lipoproteins and Lower Systolic Blood Pressure With Lifetime Risk of Cardiovascular Disease. *JAMA.* 2019; **322**(14): 1381-91.

75. Veritas Health Innovation. Covidence systematic review software. Available at www.covidence.org.

76. Clarke R, Shipley M, Lewington S, et al. Underestimation of risk associations due to regression dilution in long-term follow-up of prospective studies. *Am J Epidemiol.* 1999; **150**(4): 341-53.

77. Burgess S, Thompson SG. Mendelian randomization: methods for using genetic variants in causal estimation: CRC Press; 2015.

78. Holmes MV, Millwood IY, Kartsonaki C, et al. Lipids, Lipoproteins, and Metabolites and Risk of Myocardial Infarction and Stroke. *J Am Coll Cardiol.* 2018; **71**(6): 620-32.

79. Asia Pacific Cohort Studies Collaboration. Joint Effects of Systolic Blood Pressure and Serum Cholesterol on Cardiovascular Disease in the Asia Pacific Region. *Circulation.* 2005; **112**(22): 3384-90.

80. Ference BA, Julius S, Mahajan N, Levy PD, Williams KA, Sr., Flack JM. Clinical effect of naturally random allocation to lower systolic blood pressure beginning before the development of hypertension. *Hypertension.* 2014; **63**(6): 1182-8.

81. Higgins H, Mason AM, Larsson SC, Gill D, Langenberg C, Burgess S. Estimating the Population Benefits of Blood Pressure Lowering: A Wide-Angled Mendelian Randomization Study in UK Biobank. *J Am Heart Assoc.* 2021; **10**(17): e021098.

82. Malik R, Georgakis MK, Vujkovic M, et al. Relationship Between Blood Pressure and Incident Cardiovascular Disease: Linear and Nonlinear Mendelian Randomization Analyses. *Hypertension.* 2021; **77**(6): 2004-13.

83. Wan EYF, Fung WT, Schooling CM, et al. Blood Pressure and Risk of Cardiovascular Disease in UK Biobank: A Mendelian Randomization Study. *Hypertension.* 2021; **77**(2): 367-75.

84. Gill D, Georgakis MK, Zuber V, et al. Genetically Predicted Midlife Blood Pressure and Coronary Artery Disease Risk: Mendelian Randomization Analysis. *J Am Heart Assoc.* 2020; **9**(14): e016773.

85. Allara E, Morani G, Carter P, et al. Genetic determinants of lipids and cardiovascular disease outcomes: a wide-angled Mendelian randomization investigation. *Circ Genom Precis Med.* 2019; **12**(12): e002711.

86. Burgess S, Freitag DF, Khan H, Gorman DN, Thompson SG. Using Multivariable Mendelian Randomization to Disentangle the Causal Effects of Lipid Fractions. *PLoS One.* 2014; **9**(10): e108891.

87. Ference BA, Yoo W, Alesh I, et al. Effect of long-term exposure to lower low-density lipoprotein cholesterol beginning early in life on the risk of coronary heart disease: a Mendelian randomization analysis. *J Am Coll Cardiol.* 2012; **60**(25): 2631-9.

88. Lee SH, Lee J-Y, Kim Gh, et al. Two-Sample Mendelian Randomization Study of Lipid levels and Ischemic Heart Disease. *Korean Circ J.* 2020; **50**(10): 940-8.

89. Valdes-Marquez E, Parish S, Clarke R, Stari T, Worrall BB, Hopewell JC. Relative effects of LDL-C on ischemic stroke and coronary disease: a Mendelian randomization study. *Neurology.* 2019; **92**(11): e1176-e87.

90. White J, Swerdlow DI, Preiss D, et al. Association of Lipid Fractions With Risks for Coronary Artery Disease and Diabetes. *JAMA Cardiol.* 2016; **1**(6): 692-9.

91. Briel M, Ferreira-Gonzalez I, You JJ, et al. Association between change in high density lipoprotein cholesterol and cardiovascular disease morbidity and

mortality: systematic review and meta-regression analysis. *BMJ.* 2009; **338**: b92.

92. Genser B, März W. Low density lipoprotein cholesterol, statins and cardiovascular events: a meta–analysis. *Clin Res Cardiol.* 2007; **96**(7): 526-.

93. Silverman MG, Ference BA, Im K, et al. Association Between Lowering LDL-C and Cardiovascular Risk Reduction Among Different Therapeutic Interventions: A Systematic Review and Meta-analysis. *JAMA.* 2016; **316**(12): 1289-97.

94. Wang N, Fulcher J, Abeysuriya N, et al. Intensive LDL cholesterol-lowering treatment beyond current recommendations for the prevention of major vascular events: a systematic review and meta-analysis of randomised trials including 327 037 participants. *Lancet Diabetes Endocrinol.* 2020; **8**(1): 36-49.

95. Gong X, Chen L, Song B, et al. Associations of lipid profiles with the risk of ischemic and hemorrhagic stroke: A systematic review and meta-analysis of prospective cohort studies. *Front Cardiovasc Med.* 2022; **9**.

96. Beheshti S, Madsen CM, Varbo A, Benn M, Nordestgaard BG. Relationship of Familial Hypercholesterolemia and High Low-Density Lipoprotein Cholesterol to Ischemic Stroke. *Circulation.* 2018; **138**(6): 578-89.

97. He J, Zhu Z, Bundy JD, Dorans KS, Chen J, Hamm LL. Trends in cardiovascular risk factors in US adults by race and ethnicity and socioeconomic status, 1999-2018. *JAMA.* 2021; **326**(13): 1286-98.

98. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature.* 2018; **562**(7726): 203-9.

99. Chen Z, Chen J, Collins R, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol.* 2011; **40**(6): 1652-66.

100. Walters RG, Millwood IY, Lin K, et al. Genotyping and population characteristics of the China Kadoorie Biobank. *Cell Genom.* 2023; **3**(8).

101. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* 2015; **12**(3): e1001779.

102. China Kadoorie Biobank. Survey data: China Kadoorie Biobank. 2025. https://www.ckbiobank.org/study-resources/survey-data (accessed Janurary 20 2025).

103. UK Biobank. Genotyping and quality control of UK Biobank, a large-scale, extensively phenotyped prospective resource. *Available at https://biobank.ctsu.ox.ac.uk/crystal/docs/genotyping_qc.pdf* (accessed April 2025)

104. Geneletti S, Richardson S, Best N. Adjusting for selection bias in retrospective, case–control studies. *Biostatistics.* 2009; **10**(1): 17-31.

105. Glavinovic T, Thanassoulis G, de Graaf J, Couture P, Hegele RA, Sniderman AD. Physiological bases for the superiority of apolipoprotein b over low-density lipoprotein cholesterol and non–high-density lipoprotein cholesterol as a marker of cardiovascular risk. *J Am Heart Assoc.* 2022; **11**(20): e025858.

106. Parish S, Peto R, Palmer A, et al. The joint effects of apolipoprotein B, apolipoprotein A1, LDL cholesterol, and HDL cholesterol on risk: 3510 cases of acute myocardial infarction and 9805 controls. *Eur Heart J.* 2009; **30**(17): 2137-46.

107. Sniderman AD, Thanassoulis G, Glavinovic T, et al. Apolipoprotein B Particles and Cardiovascular Disease: A Narrative Review. *JAMA Cardiol.* 2019; **4**(12): 1287-95.

108. Sarwar N, Danesh J, Eiriksdottir G, et al. Triglycerides and the risk of coronary heart disease: 10 158 incident cases among 262 525 participants in 29 Western prospective studies. *Circulation.* 2007; **115**(4): 450-8.

109. VanderWeele TJ. Optimal approximate conversions of odds ratios and hazard ratios to risk ratios. *Biometrics.* 2020; **76**(3): 746-52.

110. Haffner SM. Relationship of metabolic risk factors and development of cardiovascular disease and diabetes. *Obesity.* 2006; **14**(S6): 121S-7S.

111. Lewington S, Li L, Sherliker P, et al. Seasonal variation in blood pressure and its relationship with outdoor temperature in 10 diverse regions of China: the China Kadoorie Biobank. *J Hypertens.* 2012; **30**(7): 1383-91.

112. Guo YT, Mazidi M, Wright N, et al. Acute Impact of Nonoptimal Ambient Temperatures on Plasma Levels of 3000 Proteins in Chinese Adults. *Environ Sci Technol.* 2025; **59**(10): 4868-82.

113. Xia X, Chan KH, Niu Y, et al. Modelling personal temperature exposure using household and outdoor temperature and questionnaire data: Implications for epidemiological studies. *Environ Int.* 2024; **192**: 109060.

114. San H, Lewington S. Lexis expansion-age-at-risk adjustment for survival analysis. Brussels: Statistics and Pharmacokinetics Pharmaceutical Users Software Exchange conference; 2013.

115. Easton DF, Peto J, Babiker AG. Floating absolute risk: an alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. *Stat Med.* 1991; **10**(7): 1025-35.

116. Lewis F, Butler A, Gilbert L. A unified approach to model selection using the likelihood ratio test. *Methods Ecol Evol.* 2011; **2**(2): 155-62.

117. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw.* 2010; **33**(1): 1.

118. Frost C, Thompson SG. Correcting for regression dilution bias: comparison of methods for a single predictor variable. *J R Stat Soc Ser A Stat Soc.* 2000; **163**(2): 173-89.

119. Chen Z, Peto R, Collins R, MacMahon S, Lu J, Li W. Serum cholesterol concentration and coronary heart disease in population with low cholesterol concentrations. *BMJ.* 1991; **303**(6797): 276-82.

120. Baigent C, Blackwell L, Emberson J, et al. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. *Lancet.* 2010; **376**(9753): 1670-81.

121. Penson PE, Pirro M, Banach M. LDL-C: lower is better for longer—even at low risk. *BMC Med.* 2020; **18**(1): 320.

122. Husain MJ, Spencer G, Nugent R, Kostova D, Richter P. The cost-effectiveness of hyperlipidemia medication in low-and middle-income countries: a review. *Global Heart.* 2022; **17**(1): 18.

123. Sun L. Major blood lipids and risk of ischaemic stroke and intracerebral haemorrhage in Chinese adults. [PhD thesis]. Oxford: University of Oxford; 2017.

124. Castañer O, Pintó X, Subirana I, et al. Remnant Cholesterol, Not LDL Cholesterol, Is Associated With Incident Cardiovascular Disease. *J Am Coll Cardiol.* 2020; **76**(23): 2712-24.

125. Wadström BN, Pedersen KM, Wulff AB, Nordestgaard BG. Elevated remnant cholesterol, plasma triglycerides, and cardiovascular and non-cardiovascular mortality. *Eur Heart J.* 2023; **44**(16): 1432-45.

126. Rhee EJ, Byrne CD, Sung KC. The HDL cholesterol/apolipoprotein A-I ratio: an indicator of cardiovascular disease. *Curr Opin Endocrinol Diabetes Obes.* 2017; **24**(2): 148-53.

127. Sniderman AD, Dufresne L, Pencina KM, Bilgic S, Thanassoulis G, Pencina MJ. Discordance among apoB, non–high-density lipoprotein cholesterol, and triglycerides: implications for cardiovascular prevention. *Eur Heart J.* 2024: ehae258.

128. Voight BF, Peloso GM, Orho-Melander M, et al. Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. *Lancet.* 2012; **380**(9841): 572-80.

129. Davies NM, Holmes MV, Smith GD. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ.* 2018; **362**.

130. Pierce BL, Ahsan H, VanderWeele TJ. Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants. *Int J Epidemiol.* 2011; **40**(3): 740-52.

131. Hemani G, Bowden J, Davey Smith G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum Mol Genet.* 2018; **27**(R2): R195-R208.

132. Graham SE, Clarke SL, Wu KH, et al. The power of genetic diversity in genome-wide association studies of lipids. *Nature.* 2021; **600**(7890): 675-9.

133. Global Lipids Genetics Consortium Results. Ancestry-specific GWAS summary statistics for HDL-C, LDL-C, nonHDL-C, TC and TG. 2025. https://csg.sph.umich.edu/willer/public/glgc-lipids2021/results/ancestry_specific/ (accessed April 5th 2025).

134. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife.* 2018; **7**: e34408.

135. Su B, Fan Z, Wu J, Zhan H. Genetic association of lipid-lowering drug target genes with pancreatic cancer: a Mendelian randomization study. *Sci Rep.* 2025; **15**(1): 3282.

136. Sanderson E. Multivariable Mendelian randomization and mediation. *Cold Spring Harb Perspect Med.* 2021; **11**(2): a038984.

137. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol.* 2017; **32**: 377-89.

138. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol.* 2016; **40**(4): 304-14.

139. Burgess S, Zuber V, Gkatzionis A, Foley CN. Modal-based estimation via heterogeneity-penalized weighting: model averaging for consistent and efficient estimation in Mendelian randomization when a plurality of candidate instruments are valid. *Int J Epidemiol.* 2018; **47**(4): 1242-54.

140. Zhang W, Ghosh D. A general approach to sensitivity analysis for Mendelian randomization. Stat Biosci. 2021; **13**(1): 34-55.

141. Brion M-JA, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. *Int J Epidemiol.* 2013; **42**(5): 1497-501.

142. Sabatine MS, Wiviott SD, Im K, Murphy SA, Giugliano RP. Efficacy and safety of further lowering of low-density lipoprotein cholesterol in patients starting with very low levels: a meta-analysis. *JAMA Cardiol.* 2018; **3**(9): 823-8.

143. O'Keeffe AG, Nazareth I, Petersen I. Time trends in the prescription of statins for the primary prevention of cardiovascular disease in the United Kingdom: a cohort study using The Health Improvement Network primary care data. *Clin Epidemiol.* 2016: 123-32.

144. Jing R, Yao H, Yan Q, et al. Trends and gaps in statin use for cardiovascular disease prevention in type 2 diabetes: A real-world study in Shanghai, China. *Endocr Pract.* 2023; **29**(10): 747-53.

145. Björnson E, Adiels M, Gummesson A, et al. Quantifying Triglyceride-Rich Lipoprotein Atherogenicity, Associations With Inflammation, and Implications for Risk Assessment Using Non-HDL Cholesterol. *J Am Coll Cardiol.* 2024; **84**(14): 1328-38.

146. Pozarickij A, Gan W, Lin K, et al. Causal relevance of different blood pressure traits on risk of cardiovascular diseases: GWAS and Mendelian randomisation in 100,000 Chinese adults. *Nat Commun.* 2024; **15**(1): 6265.

147. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. *Genet Epidemiol.* 2016; **40**(7): 597-608.

148. Ference BA, Holmes MV, Smith GD. Using Mendelian randomization to improve the design of randomized trials. *Cold Spring Harb Perspect Med.* 2021; **11**(7): a040980.

149. Brummett BH, Babyak MA, Siegler IC, et al. Systolic blood pressure, socioeconomic status, and biobehavioral risk factors in a nationally representative US young adult sample. *Hypertension.* 2011; **58**(2): 161-6.

150. Knol MJ, VanderWeele TJ. Recommendations for presenting analyses of effect modification and interaction. *Int J Epidemiol.* 2012; **41**(2): 514-20.

151. Hoaglin DC. Misunderstandings about Q and 'Cochran's Q test'in meta-analysis. *Stat Med.* 2016; **35**(4): 485-95.

152. Ortega FB, Lavie CJ, Blair SN. Obesity and cardiovascular disease. *Circ Res.* 2016; **118**(11): 1752-70.

153. Joseph K, Mehrabadi A, Lisonkova S. Confounding by indication and related concepts. *Curr Epidemiol Rep.* 2014; **1**(1): 1-8.

154. Zhao D, Qi Y, Zheng Z, et al. Dietary factors associated with hypertension. *Nat Rev Cardiol.* 2011; **8**(8): 456-65.

155. Yusuf S, Joseph P, Dans A, et al. Polypill with or without aspirin in persons without cardiovascular disease. *N Engl J Med.* 2021; **384**(3): 216-28.

156. Kahan BC, Hall SS, Beller EM, et al. Reporting of factorial randomized trials: extension of the CONSORT 2010 statement. *JAMA* 2023; **330**(21): 2106-14.

157. Rees JM, Foley CN, Burgess S. Factorial Mendelian randomization: using genetic variants to assess interactions. *Int J Epidemiol.* 2020; **49**(4): 1147-58.

158. Keaton JM, Kamali Z, Xie T, et al. Genome-wide analysis in over 1 million individuals of European ancestry yields improved polygenic risk scores for blood pressure traits. *Nat Genet.* 2024; **56**(5): 778-91.

159. Wu C-FJ. Jackknife, bootstrap and other resampling methods in regression analysis. *Ann Stat.* 1986; **14**(4): 1261-95.

160. Fang S, Hemani G, Richardson TG, Gaunt TR, Davey Smith G. Evaluating and implementing block jackknife resampling Mendelian randomization to mitigate bias induced by overlapping samples. *Hum Mol Genet.* 2023; **32**(2): 192-203.

161. Sakaue S, Kanai M, Tanigawa Y, et al. A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet.* 2021; **53**(10): 1415-24.

162. Li Y, Liu H, Shen C, et al. Association of genetic variants related to combined lipid-lowering and antihypertensive therapies with risk of cardiovascular disease: 2× 2 factorial Mendelian randomization analyses. *BMC Med.* 2024; **22**(1): 201.

163. Cao X, Zhao Z, Kang Y, et al. The burden of cardiovascular disease attributable to high systolic blood pressure across China, 2005–18: a population-based study. *Lancet Public Health.* 2022; **7**(12): e1027-e40.

164. Lonn EM, Bosch J, López-Jaramillo P, et al. Blood-Pressure Lowering in Intermediate-Risk Persons without Cardiovascular Disease. *N Engl J Med.* 2016; **374**(21): 2009-20.

165. Li X, Lu J, Hu S, et al. The primary health-care system in China. *Lancet.* 2017; **390**(10112): 2584-94.

166. Ridker PM. LDL cholesterol: controversies and future therapeutic directions. *Lancet.* 2014; **384**(9943): 607-17.

167. Jackson R, Lawes CMM, Bennett DA, Milne RJ, Rodgers A. Treatment with drugs to lower blood pressure and blood cholesterol based on an individual's absolute cardiovascular risk. *Lancet.* 2005; **365**(9457): 434-41.

168. Shepherd J. The west of Scotland Coronary Prevention study: A trial of cholesterol reduction in scottish men. *Am J Cardiol.* 1995; **76**(9, Supplement 1): 113C-7C.

169. SPRINT Research Group. A randomized trial of intensive versus standard blood-pressure control. *N Engl J Med.* 2015; **373**(22): 2103-16.

170. Writing Committee Members, Jones DW, Ferdinand KC, 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 on Clinical Practice Guidelines. *Circulation.* 2025; **152**(11): e114-e218.

171. Mortensen MB, Tybjærg-Hansen A, Nordestgaard BG. Statin Eligibility for Primary Prevention of Cardiovascular Disease According to 2021 European Prevention Guidelines Compared With Other International Guidelines. *JAMA Cardiol.* 2022; **7**(8): 836-43.

172. Coppinger C, Movahed MR, Azemawah V, Peyton L, Gregory J, Hashemzadeh M. A comprehensive review of PCSK9 inhibitors. *J Cardiovasc Pharmacol Ther.* 2022; **27**: 10742484221100107.

173. Sabatine MS, Giugliano RP, Keech AC, et al. Evolocumab and clinical outcomes in patients with cardiovascular disease. *N Engl J Med.* 2017; **376**(18): 1713-22.

174. Gennemark P, Walter K, Clemmensen N, et al. An oral antisense oligonucleotide for PCSK9 inhibition. *Sci Transl Med.* 2021; **13**(593): eabe9117.

175. Raber I, McCarthy CP, Vaduganathan M, et al. The rise and fall of aspirin in the primary prevention of cardiovascular disease. *Lancet.* 2019; **393**(10186): 2155-67.

176. Wald DS, Law M, Morris JK. Homocysteine and cardiovascular disease: evidence on causality from a meta-analysis. *BMJ.* 2002; **325**(7374): 1202.

177. Clarke R, Halsey J, Bennett D, Lewington S. Homocysteine and vascular disease: review of published results of the homocysteine-lowering trials. *J Inherit Metab Dis.* 2011; **34**(1): 83-91.

178. Roshandel G, Khoshnia M, Poustchi H, et al. Effectiveness of polypill for primary and secondary prevention of cardiovascular diseases (PolyIran): a pragmatic, cluster-randomised trial. *Lancet.* 2019; **394**(10199): 672-83.

179. Rodgers JL, Jones J, Bolleddu SI, et al. Cardiovascular risks associated with gender and aging. *J Cardiovasc Dev Dis.* 2019; **6**(2): 19.

180. Watkins DA, Pickersgill SJ, Flood D, et al. Global impact of fixed-dose combination therapies on cardiovascular mortality and events, 2023-2050: a modeling study. *J Am Coll Cardiol.* 2025; **86**(3): 149-61.

181. Lin JK, Moran AE, Bibbins-Domingo K, et al. Cost-effectiveness of a fixed-dose combination pill for secondary prevention of cardiovascular disease in China, India, Mexico, Nigeria, and South Africa: a modelling study. *Lancet Glob Health.* 2019; **7**(10): e1346-e58.

182. Lloyd-Jones DM, Allen NB, Anderson CA, et al. Life's essential 8: updating and enhancing the American Heart Association's construct of cardiovascular health: a presidential advisory from the American Heart Association. *Circulation.* 2022; **146**(5): e18-e43.

183. Nordestgaard BG, Benn M, Schnohr P, Tybjærg-Hansen A. Nonfasting triglycerides and risk of myocardial infarction, ischemic heart disease, and death in men and women. *JAMA.* 2007; **298**(3): 299-308.

184. Peng K, Li X, Wang Z, Li M, Yang Y. Association of low-density lipoprotein cholesterol levels with the risk of mortality and cardiovascular events: A meta-analysis of cohort studies with 1,232,694 participants. *Medicine (Baltimore).* 2022; **101**(48): e32003.

185. Ndumele CE, Matsushita K, Lazo M, et al. Obesity and subtypes of incident cardiovascular disease. *J Am Heart Assoc.* 2016; **5**(8): e003921.

186. Swanson JM. The UK Biobank and selection bias. *Lancet.* 2012; **380**(9837): 110.

187. Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am J Epidemiol.* 2017; **186**(9): 1026-34.

188. Tishkoff SA, Verrelli BC. Patterns of human genetic diversity: implications for human evolutionary history and disease. *Annu Rev Genomics Hum Genet.* 2003; **4**(1): 293-340.

189. Walker VM, Davies NM, Hemani G, et al. Using the MR-Base platform to investigate risk factors and drug targets for thousands of phenotypes. *Wellcome Open Res.* 2019; **4**: 113.

190. Wade KH, Hamilton FW, Carslake D, Sattar N, Davey Smith G, Timpson NJ. Challenges in undertaking nonlinear Mendelian randomization. *Obesity.* 2023; **31**(12): 2887-90.

191. Burgess S. Violation of the constant genetic effect assumption can result in biased estimates for non-linear Mendelian randomization. *Hum Hered.* 2023; **88**(1): 79-90.

192. Hamilton FW, Hughes DA, Spiller W, Tilling K, Davey Smith G. Non-linear Mendelian randomization: detection of biases using negative controls with a focus on BMI, Vitamin D and LDL cholesterol. *Eur J Epidemiol.* 2024; **39**(5): 451-65.

193. Ottensmann L, Tabassum R, Ruotsalainen SE, et al. Genome-wide association analysis of plasma lipidome identifies 495 genetic associations. *Nat Commun.* 2023; **14**(1): 6934.

194. Rosoff DB, Bell AS, Wagner J, et al. Assessing the impact of PCSK9 and HMGCR inhibition on liver function: Drug-target Mendelian randomization analyses in four ancestries. *Cell Mol Gastroenterol Hepatol.* 2024; **17**(1): 29-40.

195. Fan B, Zhang J, Zhao JV. Systematic review of Mendelian randomization studies on antihypertensive drugs. *BMC Med.* 2024; **22**(1): 547.

196. Millwood IY, Bennett DA, Holmes MV, et al. Association of CETP Gene Variants With Risk for Vascular and Nonvascular Diseases Among Chinese Adults. *JAMA Cardiol.* 2018; **3**(1): 34-43.

197. Holmes MV, Kartsonaki C, Boxall R, et al. PCSK9 genetic variants and risk of vascular and non-vascular diseases in Chinese and UK populations. *Eur J Prev Cardiol.* 2024; **31**(8): 1015-25.

198. Ding Y, Yang S, He M, Fan S, Tao X, Lu W. Lipid Metabolism Traits Mediate the Effect of Psoriasis on Myocardial Infarction Risk: A Two-Step Mendelian Randomization Study. *Metabolites.* 2023; **13**(9): 976.

199. Rees JM, Wood AM, Burgess S. Extending the MR-Egger method for multivariable Mendelian randomization to correct for both measured and unmeasured pleiotropy. *Stat Med.* 2017; **36**(29): 4705-18.

200. Sanderson E, Spiller W, Bowden J. Testing and correcting for weak and pleiotropic instruments in two-sample multivariable Mendelian randomization. *Stat Med.* 2021; **40**(25): 5434-52.

201. Clarke R, Peden JF, Hopewell JC, et al. Genetic variants associated with Lp (a) lipoprotein level and coronary disease. *N Engl J Med.* 2009; **361**(26): 2518-28.

202. Clarke R, Von Ende A, Schmidt LE, et al. Apolipoprotein proteomics for residual lipid-related risk in coronary heart disease. *Circ Res.* 2023; **132**(4): 452-64.

203. Clarke R, Wright N, Lin K, et al. Causal relevance of Lp (a) for coronary heart disease and stroke types in East Asian and European ancestry populations: a Mendelian randomization study. *Circulation.* 2025; **151**(24): 1699-711.

204. Mazidi M, Wright N, Yao P, et al. Plasma proteomics to identify drug targets for ischemic heart disease. *J Am Coll Cardiol.* 2023; **82**(20): 1906-20.

205. Mazidi M, Wright N, Yao P, et al. Risk prediction of ischemic heart disease using plasma proteomics, conventional risk factors and polygenic scores in Chinese and European adults. *Eur J Epidemiol.* 2024; **39**(11): 1229-40.

206. Mozaffarian D, Wilson PW, Kannel WB. Beyond established and novel risk factors: lifestyle risk factors for cardiovascular disease. *Circulation.* 2008; **117**(23): 3031-8.

207. Wang L, Jiang F, Sun J, et al. Factorial Mendelian randomization of lipoprotein (a) lowering, low-density lipoprotein cholesterol lowering, and lifestyle improvements: joint associations with cardiovascular risk. *Int J Epidemiol.* 2025; **54**(2): dyaf020.

Supplementary materials

Table S1. LDL-C SNPs for the East Asian Population used in MR analysis

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs406315	19	45384116	PVRL2	A	G	0.073	-0.36005	0.010381	1.00E-200
rs602633	1	109821511	PSRC1	G	T	0.0662	-0.19194	0.010447	2.15E-75
rs2738464	19	11242307	SPC24	C	G	0.283	-0.10797	0.005945	1.02E-73
rs6453131	5	74644706	HMGCR	G	T	0.498	-0.08151	0.004995	7.22E-60
rs13306194	2	21252534	APOB	G	A	0.114	-0.12348	0.007892	3.52E-55
rs151193009	1	55509585	PCSK9	C	T	0.00889	-0.48598	0.033324	3.59E-48
rs2519093	9	136141870	ABO	C	T	0.216	0.086043	0.006019	2.38E-46
rs77303550	16	72079657	TXNL4B	C	T	0.244	-0.08952	0.006396	1.62E-44
rs157592	19	45424514	APOC1	A	C	0.0931	0.116394	0.010017	3.27E-31
rs737337	19	11347493	DOCK6	T	C	0.265	-0.06188	0.005865	4.96E-26
rs6874202	5	156391628	TIMD4	C	T	0.253	-0.05573	0.005774	4.77E-22
rs6663252	1	55630151	USP24	T	C	0.12	-0.07509	0.007818	7.64E-22
rs7898735	10	74712546	PLA2G12B	T	C	0.238	-0.05779	0.006134	4.46E-21
rs2954027	8	126485294	RP11-136O12.2	A	T	0.451	0.043207	0.005076	1.72E-17
rs73596816	6	161017363	LPA	G	A	0.0321	0.115993	0.014067	1.64E-16
rs11571836	13	32973439	LRG_293	A	G	0.376	-0.04291	0.005243	2.72E-16
rs7789194	7	25992510	MIR148A	C	T	0.421	0.043026	0.005533	7.44E-15
rs5758600	22	42538670	CYP2D7P	G	T	0.433	0.043929	0.005678	1.02E-14
rs187976859	19	45533628	RELB	A	C	0.0053	-0.30524	0.039976	2.25E-14
rs2642438	1	220970028	MARC1	G	A	0.177	-0.05093	0.006827	8.63E-14
rs62133263	2	21374231	TDRD15	G	C	0.0759	0.069649	0.009487	2.11E-13
rs181711349	6	161108293	RP1-81D8.4	C	A	0.00312	0.368891	0.050591	3.06E-13
rs2419607	10	113978499	GPAM	T	G	0.403	-0.03661	0.005157	1.26E-12

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs11320208	1	234850422	RP4-781K5.7	C	A	0.256	-0.04019	0.005696	1.71E-12
rs72911441	1	55552273	USP24	A	G	0.0536	0.077341	0.011106	3.31E-12
rs9376090	6	135411228	HBS1L	T	C	0.28	-0.03983	0.005724	3.42E-12
rs62074055	17	45771933	TBKBP1	G	C	0.332	0.038156	0.005484	3.45E-12
rs2539981	2	63104904	EHBP1	C	T	0.328	-0.03731	0.005492	1.09E-11
rs1883025	9	107664301	ABCA1	C	T	0.234	-0.03971	0.00588	1.44E-11
rs12162136	17	67082266	ABCA6	A	G	0.4	0.034092	0.005091	2.12E-11
rs6741916	2	158419968	ACVR1C	A	T	0.271	-0.04136	0.006327	6.27E-11
rs11601507	11	5701074	TRIM5	C	A	0.0843	0.059695	0.009267	1.18E-10
rs174559	11	61581656	FADS1	G	A	0.41	-0.03583	0.005602	1.59E-10
rs7140110	13	114544024	GAS6	T	C	0.217	0.042974	0.006737	1.79E-10
rs112784971	8	59406490	CYP7A1	C	T	0.244	-0.03633	0.005854	5.44E-10
rs3752442	4	3446883	HGFAC	A	G	0.453	-0.03297	0.005393	9.74E-10
rs59379014	11	126228000	ST3GAL4	C	T	0.0938	0.053765	0.008809	1.04E-09
rs12229026	12	100834253	SLC17A8	T	C	0.282	0.034689	0.005712	1.26E-09
rs4646776	12	112230019	ALDH2	G	C	0.195	0.039033	0.006442	1.37E-09
rs7212349	17	26733698	CTD-2350C19.2	C	T	0.274	-0.03403	0.005752	3.28E-09
rs6493583	15	53096084	RP11-209K10.2	C	G	0.174	-0.04098	0.006943	3.59E-09
rs59857465	9	130719709	FAM102A	G	A	0.0763	0.056316	0.009819	9.74E-09
rs2054345	8	9199305	RP11-115J16.1	T	C	0.0119	-0.13521	0.023638	1.06E-08
rs9274662	6	32636494	HLA-DQB1	C	T	0.167	0.047445	0.008387	1.54E-08
rs10903130	1	25769221	TMEM57	A	T	0.299	0.030631	0.005443	1.82E-08
rs2737265	8	116667634	TRPS1	A	G	0.299	-0.03001	0.005416	2.99E-08

Abbreviations. **LDL-C**: Low-Density Lipoprotein Cholesterol; **SNP**: Single Nucleotide Polymorphism; **RSID**: Reference SNP Cluster ID; **GRCh37**: Genome Reference Consortium Human Build 37

Table S2. HDL-C SNPs for the East Asian Population used in MR analysis

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs74368849	11	116622299	BUD13	G	A	0.0615	-0.30029	0.008877	1.00E-200
rs72786786	16	56985514	AC012181.1	G	A	0.171	0.242173	0.00603	1.00E-200
rs2070895	15	58723939	LIPC	G	A	0.442	0.127025	0.004357	7.21E-187
rs77069344	8	19821782	LPL	T	G	0.107	0.166403	0.006721	2.44E-135
rs7350481	11	116586283	BUD13	C	T	0.239	-0.13091	0.005943	1.58E-107
rs2740488	9	107661742	ABCA1	A	C	0.245	-0.09904	0.004792	6.93E-95
rs11082764	18	47119579	LIPG	A	G	0.422	0.081604	0.004229	5.87E-83
rs4775042	15	58674730	ALDH1A2	G	C	0.459	-0.07607	0.004296	3.67E-70
rs429358	19	45411941	APOE	T	C	0.0933	-0.11431	0.007454	4.46E-53
rs3782886	12	112110489	BRAP	T	C	0.206	-0.08389	0.005631	3.41E-50
rs35950463	12	125369124	SCARB1	G	A	0.256	0.067019	0.005132	5.68E-39
rs255056	16	68016185	DPEP3	G	C	0.127	0.078691	0.006432	2.03E-34
rs3745683	19	11348521	C19orf80	G	A	0.256	-0.06503	0.005778	2.21E-29
rs1883023	9	107564846	ABCA1	C	T	0.382	-0.05741	0.005154	8.06E-29
rs9304381	18	47158234	SMUG1P1	T	C	0.234	-0.05118	0.004904	1.67E-25
rs247610	16	56972345	HERPUD1	T	G	0.442	-0.04925	0.005009	8.32E-23
rs2792703	10	113931690	GPAM	T	C	0.274	0.043195	0.004674	2.41E-20
rs2737212	8	116621214	TRPS1	T	C	0.198	-0.04898	0.005328	3.86E-20
rs35499210	6	161023491	LPA	G	T	0.433	-0.03802	0.004189	1.12E-19
rs6458349	6	43759789	VEGFA	A	G	0.157	0.053855	0.006241	6.20E-18
rs4821114	22	21970810	UBE2L3	G	C	0.367	-0.03719	0.004317	7.00E-18
rs12633551	3	12360494	PPARG	C	T	0.262	-0.04066	0.004727	7.88E-18
rs111859808	20	44558849	PCIF1	G	A	0.0603	-0.07442	0.008667	8.96E-18
rs143722364	7	80328352	CD36	T	C	0.0245	0.113859	0.013532	3.96E-17
rs579953	15	58641514	ALDH1A2	C	G	0.305	0.038034	0.004557	7.07E-17
rs3741521	12	125264293	SCARB1	G	A	0.39	-0.04277	0.005183	1.58E-16
rs10444449	12	123764271	RNA5SP375	A	G	0.288	0.03837	0.004655	1.68E-16

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs10886864	10	122929537	LINC01153	T	C	0.346	0.036106	0.004464	6.02E-16
rs448092	19	54784936	LILRB2	T	C	0.272	-0.05022	0.00623	7.59E-16
rs1917368	7	17911752	SNX13	G	T	0.471	-0.03947	0.004944	1.42E-15
rs59842799	11	117092340	PCSK7	C	T	0.0318	0.093139	0.01213	1.61E-14
rs12992267	2	21215645	RP11-116D2.1	C	T	0.0833	-0.05519	0.007461	1.40E-13
rs2925979	16	81534790	CMIP	C	T	0.392	-0.03842	0.005235	2.15E-13
rs2269434	11	47360412	LRG_386	C	T	0.355	-0.03231	0.004422	2.75E-13
rs2422493	9	107690995	RP11-217B7.2	G	A	0.412	0.030736	0.004338	1.39E-12
rs333947	1	110470764	CSF1	G	A	0.282	-0.03556	0.005055	1.99E-12
rs11236534	11	75514528	RP11-535A19.2	T	C	0.236	-0.0344	0.004904	2.32E-12
rs10850379	12	110002777	MMAB	C	T	0.214	0.035564	0.005071	2.33E-12
rs111997200	19	45466325	CLPTM1	C	T	0.461	-0.03005	0.004314	3.27E-12
rs7137797	12	125327352	SCARB1	A	G	0.422	-0.02938	0.004303	8.61E-12
rs4418728	10	94839724	CYP26A1	T	G	0.262	-0.03319	0.004937	1.78E-11
rs113879410	1	40073011	HEYL	G	C	0.123	-0.04286	0.006397	2.08E-11
rs77648222	15	70182851	LINC00593	A	G	0.21	0.033702	0.00506	2.72E-11
rs11993276	8	71095359	NCOA2	G	A	0.415	-0.02851	0.004303	3.48E-11
rs4846913	1	230294715	GALNT2	A	C	0.186	-0.03571	0.005402	3.81E-11
rs2642438	1	220970028	MARC1	G	A	0.177	-0.03671	0.005629	6.94E-11
rs931992	17	37821435	TCAP	G	T	0.437	0.033194	0.005202	1.77E-10
rs12565182	1	93861589	Y_RNA	G	T	0.363	0.028579	0.004482	1.82E-10
rs117592139	6	32357059	HCG23	G	A	0.129	0.041788	0.006719	5.00E-10
rs13248499	8	121860589	RP11-713M15.2	A	G	0.283	0.030316	0.004896	5.94E-10
rs73544986	16	57044322	NLRC5	T	C	0.132	-0.04803	0.007886	1.13E-09
rs1115721	4	100636286	RP11-766F14.2	T	G	0.216	0.030424	0.005055	1.75E-09
rs79831695	11	117022328	PAFAH1B2	A	G	0.0148	0.102684	0.017078	1.83E-09
rs10866704	5	176676461	LRG_512	T	A	0.474	-0.02654	0.004424	1.98E-09
rs75766425	14	52511911	NID2	G	C	0.123	0.038348	0.006416	2.27E-09
rs174565	11	61591636	FADS1	C	G	0.407	-0.0266	0.004529	4.29E-09

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs2126260	8	9185081	RP11-115J16.1	C	T	0.0154	-0.10036	0.01736	7.42E-09
rs12753251	1	178529898	RNA5SP69	A	G	0.496	0.023873	0.004149	8.73E-09
rs146544283	11	117147061	RNF214	C	T	0.00849	-0.13286	0.023143	9.43E-09
rs11098353	4	79704102	BMP2K	A	C	0.0742	-0.05651	0.009851	9.70E-09
rs12450500	17	76395750	PGS1	C	T	0.483	0.023543	0.004132	1.21E-08
rs2980870	8	126483840	RP11-136O12.2	G	A	0.33	-0.02517	0.004448	1.53E-08
rs16868954	5	39451982	DAB2	C	T	0.13	0.042262	0.007506	1.79E-08
rs11773763	7	50271499	AC020743.3	C	T	0.32	0.025033	0.004565	4.18E-08

Abbreviations. **HDL-C**: High-Density Lipoprotein Cholesterol; **SNP**: Single Nucleotide Polymorphism; **RSID**: Reference SNP Cluster ID; **GRCh37**: Genome Reference Consortium Human Build 37

Table S3. Triglycerides SNPs for the East Asian Population used in MR analysis

SNP	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs7350481	11	116586283	BUD13	C	T	0.239	0.226663	0.005957	1.00E-200
rs74368849	11	116622299	BUD13	G	A	0.059	0.372043	0.010737	1.00E-200
rs74855321	8	19836984	LPL	C	T	0.0986	-0.17722	0.008276	9.92E-102
rs1260326	2	27730940	GCKR	T	C	0.48	-0.09829	0.005089	4.14E-83
rs13246993	7	73022746	MLXIPL	G	A	0.105	-0.1348	0.008279	1.33E-59
rs483082	19	45416178	APOC1	G	T	0.166	0.112978	0.007302	5.40E-54
rs2954021	8	126482077	RP11-136O12.2	G	A	0.435	0.070738	0.004969	5.53E-46
rs1077835	15	58723426	LIPC	A	G	0.413	0.070267	0.005337	1.38E-39
rs4495740	1	63124465	DOCK7	T	G	0.186	-0.08178	0.00635	6.01E-38
rs7165077	15	58686809	ALDH1A2	C	T	0.359	-0.04968	0.005426	5.36E-20
rs4704834	5	156443066	HAVCR1	G	A	0.248	-0.05324	0.005829	6.63E-20
rs58542926	19	19379549	TM6SF2	C	T	0.067	-0.08998	0.010139	7.06E-19
rs35570672	8	18272635	NAT2	T	C	0.434	-0.04073	0.004999	3.68E-16
rs117157737	11	116833013	SIK3	G	T	0.0116	-0.18975	0.023951	2.33E-15
rs6706783	2	21191409	RP11-116D2.1	T	A	0.105	0.062194	0.008008	8.07E-15
rs16990971	20	44601293	ZNF335	A	G	0.0609	0.079202	0.010342	1.89E-14
rs7394579	11	61581450	FADS1	A	G	0.4	0.043543	0.005723	2.78E-14
rs3752442	4	3446883	HGFAC	A	G	0.453	-0.03816	0.005346	9.48E-13
rs116111528	6	32313512	C6orf10	C	T	0.0457	0.086836	0.012388	2.39E-12
rs821840	16	56993886	CETP	A	G	0.163	-0.04441	0.00678	5.73E-11
rs1475537	9	92212750	GADD45G	C	T	0.263	0.036954	0.00574	1.21E-10
rs72670104	4	88151605	KLHL8	C	T	0.339	0.036689	0.005718	1.40E-10
rs79702711	11	117128389	RNF214	C	G	0.019	-0.1192	0.018695	1.82E-10
rs7078456	10	65269479	REEP3	T	C	0.33	0.034303	0.005405	2.20E-10
rs998584	6	43757896	VEGFA	A	C	0.44	-0.03488	0.005523	2.71E-10
rs4418728	10	94839724	CYP26A1	T	G	0.214	0.039434	0.006344	5.09E-10

SNP	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs2059248	16	72322376	RP11-328J14.2	T	G	0.462	0.031554	0.005255	1.92E-09
rs34828061	8	59382169	UBXN2B	A	G	0.262	-0.03336	0.00564	3.32E-09
rs1985109	19	19469907	MAU2	T	G	0.0572	-0.06751	0.011455	3.77E-09
rs80001814	11	116854372	SIK3	G	A	0.00925	0.155823	0.02687	6.66E-09
rs9687832	5	55861595	AC022431.2	G	A	0.111	0.045111	0.007846	8.95E-09
rs2291119	11	47298201	MADD	C	T	0.299	0.030925	0.00549	1.77E-08
rs10886864	10	122929537	LINC01153	T	C	0.366	-0.02964	0.005275	1.91E-08
rs6889596	5	123030802	KRT18P16	T	C	0.277	-0.03159	0.005643	2.15E-08
rs9378236	6	31293035	XXbac-BPG248L24.10	A	T	0.266	0.039442	0.007059	2.30E-08
rs12229430	12	94861277	CCDC41	G	C	0.367	-0.03113	0.005705	4.85E-08

Abbreviations. **SNP**: Single Nucleotide Polymorphism; **RSID**: Reference SNP Cluster ID; **GRCh37**: Genome Reference Consortium Human Build 37

Table S4. LDL-C SNPs for European Populations used in MR analysis

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs12732125	1	55470153	BSND	C	T	0.0182	-0.20637	0.006354	1.00E-200
rs553741	1	55520408	PCSK9	G	C	0.361	0.055371	0.001753	1.00E-200
rs4970829	1	109757295	SARS	G	A	0.0716	-0.11002	0.003178	1.00E-200
rs4075673	2	21150787	RP11-116D2.1	T	C	0.442	-0.05893	0.001658	1.00E-200
rs1712249	2	21385974	TDRD15	T	C	0.249	0.071659	0.001895	1.00E-200
rs56266464	2	44054991	ABCG5	G	A	0.0617	-0.10941	0.003426	1.00E-200
rs2112653	5	74562564	ANKRD31	T	C	0.441	0.063054	0.001721	1.00E-200
rs118039278	6	160985526	LPA	G	A	0.0637	0.111285	0.003389	1.00E-200
rs2954017	8	126476873	RP11-136O12.2	T	C	0.472	-0.05055	0.001656	1.00E-200
rs2519093	9	136141870	ABO	C	T	0.189	0.073474	0.002058	1.00E-200
rs1109375	19	10914719	DNM2	C	G	0.17	-0.07082	0.002176	1.00E-200
rs2738445	19	11227070	LRG_274	C	T	0.43	0.052976	0.001666	1.00E-200
rs58542926	19	19379549	TM6SF2	C	T	0.0752	-0.09876	0.003094	1.00E-200
rs62119267	19	45134682	IGSF23	A	C	0.0198	-0.23267	0.005924	1.00E-200
rs531660643	19	45251806	BCL3	G	T	0.0162	-0.41297	0.007347	1.00E-200
rs113330691	19	45295886	CBLC	G	A	0.0327	-0.23012	0.004551	1.00E-200
rs146275714	19	45365447	PVRL2	G	A	0.0283	0.173572	0.005381	1.00E-200
rs6670347	1	109815368	CELSR2	T	C	0.0397	-0.11496	0.004184	3.18E-166
rs77542162	17	67081278	ABCA6	A	G	0.018	0.169014	0.00626	1.57E-160
rs12924886	16	72075593	TXNL4B	A	T	0.192	0.054227	0.002078	3.81E-150
rs1883711	20	39179822	SNORD112	G	C	0.0322	0.127743	0.004983	5.99E-145
rs174546	11	61569830	FADS1	C	T	0.347	-0.04367	0.001716	9.15E-143
rs41289514	19	45359667	PVRL2	A	G	0.00967	0.215772	0.008559	3.04E-140
rs553427	1	234852760	RP4-781K5.7	C	T	0.469	0.038281	0.001669	1.83E-116
rs117733303	6	160922870	LPAL2	A	G	0.0171	0.141976	0.006346	7.13E-111
rs2618566	20	17844684	AL035045.1	G	T	0.339	-0.0404	0.001806	8.46E-111

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs7534572	1	62999675	DOCK7	C	G	0.316	0.039596	0.001867	7.60E-100
rs144900553	19	10798974	ILF3	C	T	0.0136	-0.15784	0.007625	3.52E-95
rs9987289	8	9183358	RP11-115J16.1	A	G	0.0916	0.057486	0.002816	1.34E-92
rs10260606	7	44584551	NPC1L1	G	C	0.201	0.042917	0.002115	1.50E-91
rs1169288	12	121416650	LRG_522	A	C	0.326	0.035422	0.001754	9.94E-91
rs247617	16	56990716	AC012181.1	C	A	0.319	-0.0346	0.001747	2.97E-87
rs12657266	5	156396003	TIMD4	C	T	0.365	0.032558	0.001695	3.05E-82
rs55696093	7	21605973	DNAH11	A	G	0.206	0.038404	0.002027	4.70E-80
rs11206517	1	55526428	PCSK9	T	G	0.0393	0.081149	0.004296	1.40E-79
rs76970536	11	126250680	ST3GAL4	G	A	0.0713	0.058869	0.003137	1.42E-78
rs12208357	6	160543148	SLC22A1	C	T	0.0694	0.059346	0.003168	2.62E-78
rs9297994	8	59392324	CYP7A1	G	A	0.339	-0.03158	0.001727	9.24E-75
rs1801689	17	64210580	APOH	A	C	0.0243	0.094514	0.005169	1.06E-74
rs6689	6	32627700	XXbac-BPG254F23.6	A	G	0.21	0.041386	0.002299	1.98E-72
rs7746081	6	16126934	MYLIP	G	A	0.33	-0.03108	0.001737	1.42E-71
rs72631343	17	67191270	ABCA10	C	G	0.131	-0.04123	0.002429	1.34E-64
rs148601586	19	45346666	PVRL2	C	G	0.0106	0.138212	0.008217	1.75E-63
rs80215559	6	25918225	SLC17A2	T	C	0.0615	-0.05765	0.003433	2.84E-63
rs140488605	2	44063346	ABCG5	C	T	0.0424	0.074537	0.004445	4.10E-63
rs55714927	17	7080316	ASGR1	C	T	0.187	-0.03524	0.002197	6.56E-58
rs17036094	1	109834940	MYBPHL	A	C	0.0114	-0.12078	0.007598	6.83E-57
rs3184504	12	111884608	SH2B3	T	C	0.468	0.025259	0.001647	4.56E-53
rs11870935	17	45732605	KPNB1	G	A	0.491	-0.02488	0.001646	1.38E-51
rs10903129	1	25768937	TMEM57	A	G	0.458	0.024687	0.001636	1.94E-51
rs7250652	19	11302606	KANK2	A	G	0.439	0.025048	0.001676	1.65E-50
rs150474434	2	118845121	INSIG2	G	A	0.0919	-0.04226	0.002847	7.99E-50
rs79598313	1	27284913	C1orf172	C	T	0.0237	0.074494	0.005277	2.94E-45

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs2642438	1	220970028	MARC1	A	G	0.293	0.025448	0.001805	3.82E-45
rs138483078	1	55489240	BSND	G	A	0.142	0.032972	0.002356	1.70E-44
rs11787335	8	145044104	PLEC	C	T	0.353	0.02383	0.001739	9.48E-43
rs1177809	3	12466490	PPARG	A	G	0.344	-0.02356	0.001732	3.82E-42
rs7758845	6	135428537	HBS1L	A	C	0.279	-0.02461	0.001819	1.06E-41
rs2072560	11	116661826	APOA5	T	C	0.0647	-0.04419	0.003311	1.21E-40
rs9423289	10	124704695	C10orf88	C	T	0.423	0.021993	0.00166	4.71E-40
rs2792751	10	113940329	GPAM	T	C	0.285	-0.02324	0.001794	2.29E-38
rs6093446	20	39780932	PLCG1	G	A	0.269	0.024017	0.001856	2.78E-38
rs34707604	4	69491456	RP11-1267H10.3	T	C	0.242	0.030061	0.002327	3.42E-38
rs34503352	19	58651296	ZNF329	G	A	0.161	-0.02867	0.002253	4.27E-37
rs11601507	11	5701074	TRIM5	C	A	0.07	0.040413	0.003176	4.36E-37
rs1260326	2	27730940	GCKR	T	C	0.39	-0.02116	0.001675	1.41E-36
rs704	17	26694861	SARM1	G	A	0.478	0.020362	0.00164	2.10E-35
rs360804	2	62939397	EHBP1	G	A	0.316	-0.02205	0.001779	2.70E-35
rs7515577	1	93009438	EVI5	C	A	0.207	0.025174	0.00204	5.67E-35
rs11621792	14	24871926	NYNRIN	C	T	0.448	0.02077	0.001689	9.39E-35
rs72654437	19	45416099	APOC1	G	A	0.0295	0.066203	0.005395	1.31E-34
rs71311871	3	58420613	PDHB	A	G	0.081	-0.0363	0.002972	2.57E-34
rs12086676	1	55738663	RP11-101C11.1	C	T	0.174	-0.02652	0.002175	3.29E-34
rs2737252	8	116663898	TRPS1	G	A	0.276	-0.02204	0.001826	1.48E-33
rs17580	14	94847262	SERPINA1	T	A	0.034	0.053693	0.004501	8.39E-33
rs61016611	20	34212151	CPNE1	G	A	0.101	-0.03234	0.00272	1.36E-32
rs28615248	8	55451193	RP11-53M11.2	T	C	0.195	0.024553	0.002077	3.03E-32
rs7327867	13	32968591	LRG_293	G	A	0.474	-0.01926	0.001652	2.01E-31
rs10184673	2	169827796	ABCB11	G	A	0.409	-0.01953	0.001685	4.55E-31
rs10832956	11	18645668	SPTY2D1	T	C	0.279	0.021153	0.001826	4.94E-31

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs17050272	2	121306440	AC073257.2	G	A	0.426	-0.01998	0.001726	5.50E-31
rs2740488	9	107661742	ABCA1	A	C	0.26	-0.02124	0.00187	6.65E-30
rs10932008	2	204048840	NBEAL1	G	A	0.184	0.024506	0.002208	1.26E-28
rs77704739	5	52080909	CTD-2288O8.1	T	C	0.0361	-0.04885	0.004401	1.27E-28
rs1800961	20	43042364	HNF4A	C	T	0.0348	-0.04906	0.004489	8.47E-28
rs9837622	3	32514647	CMTM7	T	A	0.0725	-0.03502	0.003215	1.23E-27
rs67890964	16	83979317	RP11-505K9.4	T	C	0.385	-0.01889	0.001761	7.87E-27
rs3780181	9	2640759	VLDLR	A	G	0.0708	-0.03479	0.003245	8.09E-27
rs79391862	15	53739426	WDR72	A	C	0.0231	-0.06463	0.006039	1.00E-26
rs11620731	14	70817141	SYNJ2BP-COX16	T	C	0.147	0.024901	0.002328	1.07E-26
rs78946096	3	132188163	DNAJC13	A	G	0.0471	-0.0413	0.003876	1.65E-26
rs7202323	16	72217113	PMFBP1	T	G	0.227	-0.02087	0.001968	2.93E-26
rs12551960	9	19267440	DENND4C	C	T	0.0794	0.033333	0.003145	3.04E-26
rs10263252	7	1049949	C7orf50	G	A	0.208	-0.02146	0.002034	4.98E-26
rs564449	7	100321138	EPO	T	G	0.114	-0.02708	0.002584	1.07E-25
rs28929474	14	94844947	SERPINA1	C	T	0.0169	0.065287	0.006253	1.62E-25
rs896311	7	25934357	CTD-2227E11.1	G	A	0.303	0.018231	0.001805	5.37E-24
rs548088667	2	44087934	ABCG8	C	T	0.0183	0.076381	0.007642	1.61E-23
rs2125345	17	73782191	UNK	T	C	0.291	-0.018	0.001802	1.71E-23
rs13189347	5	122828560	Y_RNA	A	C	0.448	-0.01654	0.001659	2.15E-23
rs11754773	6	34577257	C6orf106	A	G	0.0977	-0.02736	0.002759	3.61E-23
rs17665178	6	52435243	TRAM2	C	G	0.304	-0.01828	0.00185	4.93E-23
rs1375131	2	135954797	ZRANB3	T	C	0.341	0.020758	0.002115	9.67E-23
rs4751995	10	118397884	PNLIPRP2	A	G	0.47	0.015939	0.001638	2.20E-22
rs2251024	12	51103633	DIP2B	C	T	0.339	-0.017	0.00175	2.63E-22
rs204473	19	45487519	CLPTM1	A	G	0.0223	0.054915	0.005685	4.51E-22

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs9496567	6	100602753	PRDX2P4	G	A	0.233	-0.01878	0.00195	6.07E-22
rs28497720	4	100487370	MTTP	C	T	0.25	-0.01813	0.001885	6.75E-22
rs34019521	12	618790	B4GALNT3	G	C	0.252	0.01851	0.001941	1.52E-21
rs79429216	19	45445517	APOC4	G	A	0.00891	0.086959	0.00915	2.02E-21
rs13108218	4	3443931	HGFAC	A	G	0.387	-0.01643	0.001741	3.87E-21
rs6894249	5	131797547	C5orf56	A	G	0.389	0.01572	0.001685	1.05E-20
rs2068888	10	94839642	CYP26A1	G	A	0.452	-0.01537	0.001648	1.11E-20
rs1556857	6	116325559	FRK	T	C	0.399	-0.01544	0.001665	1.80E-20
rs1250259	2	216300482	FN1	T	A	0.264	0.017467	0.001889	2.35E-20
rs2465956	2	109065004	GCC2	T	C	0.195	0.019302	0.00209	2.57E-20
rs80098465	1	56323794	RP11-90C4.1	G	A	0.0334	-0.04276	0.004637	2.93E-20
rs12445804	16	11706100	LITAF	G	A	0.077	0.029058	0.003168	4.59E-20
rs146534110	6	160578069	SLC22A1	G	T	0.0108	0.074576	0.008151	5.73E-20
rs887829	2	234668570	UGT1A3	C	T	0.328	-0.01577	0.001736	1.09E-19
rs9524538	13	95218192	TGDS	G	A	0.235	0.017433	0.001926	1.38E-19
rs10761750	10	65128619	JMJD1C	G	A	0.488	0.014896	0.001647	1.52E-19
rs6602909	13	114551993	GAS6	T	C	0.332	0.01619	0.0018	2.37E-19
rs12451056	17	76396188	PGS1	C	T	0.143	-0.02037	0.002315	1.38E-18
rs13379043	14	74250126	ELMSAN1	T	C	0.268	-0.01617	0.001847	2.03E-18
rs10184376	2	158442005	ACVR1C	C	T	0.0737	-0.02713	0.0031	2.12E-18
rs900776	8	21918089	DMTN	A	C	0.167	-0.01949	0.002237	3.01E-18
rs111654618	19	45300824	CBLC	G	A	0.00925	0.078842	0.009068	3.48E-18
rs6792725	3	24520283	THRB	A	G	0.325	-0.016	0.00185	5.38E-18
rs11057841	12	125316743	SCARB1	C	T	0.142	0.020512	0.002373	5.48E-18
rs77960347	18	47109955	LIPG	A	G	0.0121	0.063971	0.007405	5.71E-18
rs3010277	5	72014400	CTC-347C20.1	A	G	0.21	-0.0173	0.002008	6.88E-18
rs12320328	12	25408464	KRAS	A	G	0.0824	-0.02573	0.002988	7.23E-18
rs969075	20	17792323	AL035045.1	T	C	0.334	0.015262	0.00178	9.98E-18

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs267733	1	150958836	ANXA9	A	G	0.153	-0.01931	0.002254	1.07E-17
rs116276872	1	110068291	GPR61	C	T	0.0188	-0.05479	0.006462	2.26E-17
rs16988435	21	33058571	SCAF4	C	T	0.0514	0.031968	0.003776	2.53E-17
rs1689801	1	182165484	GS1-122H1.2	G	A	0.322	0.014794	0.001748	2.62E-17
rs13161656	5	141916867	AC005592.2	T	C	0.108	0.02246	0.002655	2.71E-17
rs4465	22	35708790	TOM1	C	T	0.358	-0.0144	0.001724	6.62E-17
rs7864568	9	78212428	RP11-470F18.1	A	G	0.321	0.015536	0.001863	7.53E-17
rs11638576	15	75260387	SCAMP5	G	A	0.396	0.014563	0.001757	1.14E-16
rs6940814	6	42924932	RPL24P4	A	G	0.416	-0.01366	0.001657	1.70E-16
rs10858093	1	109949756	PSMA5	C	T	0.0234	0.050055	0.006116	2.74E-16
rs62509311	8	74907295	LY96	A	T	0.277	-0.01511	0.001853	3.60E-16
rs12055389	6	161091952	RP1-81D8.4	C	T	0.0592	-0.02842	0.0035	4.64E-16
rs12054451	3	122064369	CSTA	T	G	0.26	0.015059	0.001866	7.04E-16
rs11999532	9	139371405	SEC16A	G	C	0.267	-0.01483	0.001843	8.71E-16
rs7046887	9	78736048	PCSK5	T	C	0.464	-0.01319	0.001647	1.14E-15
rs6601302	8	9239458	RP11-115J16.2	T	G	0.248	-0.01546	0.001935	1.38E-15
rs17647249	17	65195226	HELZ	T	C	0.499	-0.01338	0.001686	2.14E-15
rs438568	20	12958687	RP5-1069C8.2	A	G	0.392	0.01325	0.001678	2.84E-15
rs8121509	20	62712053	OPRL1	T	C	0.452	-0.01387	0.001764	3.75E-15
rs3738622	1	235110859	RP11-443B7.3	G	T	0.203	-0.01643	0.00209	3.81E-15
rs12016920	13	114639025	LINC00565	T	C	0.192	-0.01681	0.002146	4.72E-15
rs3820897	2	3642361	LRG_350	T	C	0.198	-0.01606	0.002065	7.26E-15
rs117264457	19	45404432	TOMM40	G	A	0.0196	-0.04922	0.006332	7.70E-15
rs12533280	7	36171953	RP11-377H23.1	C	T	0.2	0.015997	0.002063	8.96E-15
rs73066228	19	59013196	SLC27A5	A	G	0.164	0.01733	0.002236	9.29E-15

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs60847460	10	113983758	GPAM	C	T	0.138	-0.01815	0.002349	1.12E-14
rs150688657	17	7505801	LRG_285	G	A	0.103	0.021183	0.002743	1.14E-14
rs745763	2	43455654	THADA	T	C	0.168	0.01819	0.002356	1.16E-14
rs72836561	17	41926126	CD300LG	C	T	0.0296	-0.03719	0.00482	1.21E-14
rs4390169	1	155106054	EFNA1	A	G	0.487	-0.01256	0.001633	1.48E-14
rs240762	6	100985574	ASCC3	C	T	0.406	-0.01308	0.001701	1.50E-14
rs6478851	9	131561110	TBC1D13	G	A	0.234	-0.01491	0.00194	1.54E-14
rs11175540	12	40586295	AC079630.4	T	A	0.0647	0.026107	0.003415	2.09E-14
rs75907879	1	16170903	RP11-169K16.9	C	T	0.121	0.020073	0.002627	2.14E-14
rs1584688	3	160267858	KPNA4	C	T	0.449	-0.01297	0.001707	2.98E-14
rs55726838	2	44071788	ABCG8	G	A	0.00907	0.070093	0.009235	3.19E-14
rs45505501	11	118405305	TMEM25	G	A	0.137	0.018413	0.002427	3.27E-14
rs4939883	18	47167214	SMUG1P1	T	C	0.175	0.01618	0.002146	4.65E-14
rs28498684	9	16900695	BNC2	A	G	0.41	-0.01252	0.001671	6.64E-14
rs17617028	4	106064683	TET2	G	A	0.223	0.014583	0.001954	8.56E-14
rs642803	11	65560620	OVOL1	C	T	0.452	0.012257	0.001662	1.63E-13
rs6967728	7	97915637	BRI3	A	G	0.182	0.016	0.002193	2.97E-13
rs3732359	3	119536429	NR1I2	G	A	0.236	-0.01393	0.001922	4.35E-13
rs551473284	13	111038325	COL4A2	T	C	0.367	0.015159	0.002096	4.78E-13
rs112170089	6	160502019	IGF2R	G	A	0.0164	0.055013	0.00761	4.86E-13
rs10193634	2	64933985	SERTAD2	C	A	0.391	0.012354	0.001717	6.22E-13
rs138335	22	41227086	ST13	C	G	0.347	-0.01275	0.001785	9.17E-13
rs3747207	22	44324855	PNPLA3	G	A	0.226	-0.01395	0.001957	1.03E-12
rs6458949	6	53508265	KLHL31	G	T	0.259	0.01324	0.001866	1.30E-12
rs3745686	19	11564993	ELAVL3	G	A	0.0421	-0.03019	0.004256	1.30E-12
rs9795910	12	51795623	SLC4A8	A	G	0.376	-0.01231	0.001738	1.42E-12
rs72768400	5	75046359	SLC25A5P9	G	T	0.0372	-0.03449	0.004876	1.51E-12
rs11635502	15	63771730	AC007950.1	G	A	0.306	0.012657	0.001792	1.62E-12

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs71536537	6	32523877	HLA-DRB6	C	T	0.473	0.026135	0.00371	1.86E-12
rs11105294	12	89858839	POC1B	G	A	0.339	0.012324	0.001757	2.33E-12
rs28677840	19	44200667	IRGC	C	A	0.0694	-0.02233	0.003202	3.08E-12
rs56118251	3	171534525	PLD1	A	G	0.155	0.015415	0.002233	5.06E-12
rs7206039	16	71645914	RP11-432I5.2	C	A	0.371	0.012093	0.001757	5.90E-12
rs2970902	2	88429183	FABP1	G	C	0.344	0.011941	0.001737	6.30E-12
rs5746498	22	18435794	MICAL3	C	T	0.237	-0.01392	0.002027	6.60E-12
rs414850	21	40543337	PCBP2P1	A	C	0.411	-0.01165	0.001703	7.89E-12
rs7955221	12	100850750	NR1H4	A	C	0.453	-0.0115	0.001691	1.02E-11
rs79778025	1	26449337	PDIK1L	A	G	0.176	-0.01647	0.002434	1.33E-11
rs2296288	1	183072590	LAMC1	T	C	0.442	0.011191	0.00166	1.58E-11
rs5752963	22	30203833	ASCC2	G	A	0.036	0.028915	0.004312	2.01E-11
rs2911971	8	6607634	AGPAT5	C	G	0.374	0.011506	0.001717	2.07E-11
rs8681	1	45468606	HECTD3	A	G	0.255	0.012666	0.001891	2.09E-11
rs113776580	5	74950420	ANKDD1B	G	A	0.0143	0.046195	0.006899	2.14E-11
rs11846704	14	35186694	LRG_213	C	T	0.267	-0.01232	0.001845	2.42E-11
rs12271225	11	5690544	TRIM5	A	T	0.133	0.015969	0.002398	2.76E-11
rs9653945	3	142660706	RP11-372E1.4	G	A	0.344	-0.01135	0.001715	3.61E-11
rs67777803	17	27323322	SEZ6	G	T	0.168	-0.01464	0.002211	3.63E-11
rs9805560	13	50076079	PHF11	G	A	0.465	-0.01091	0.00165	3.84E-11
rs10910522	1	235011725	RN7SL668P	G	A	0.433	-0.01108	0.001683	4.59E-11
rs2295027	20	47582884	ARFGEF2	G	A	0.321	0.011947	0.001817	4.86E-11
rs1123571	1	2326009	RER1	G	A	0.473	-0.01111	0.001693	5.26E-11
rs35313547	19	58352806	ZNF587B	T	C	0.158	-0.01461	0.002231	5.83E-11
rs7184567	16	31021078	STX1B	C	T	0.375	0.011223	0.001717	6.27E-11
rs2665404	17	57875396	VMP1	C	T	0.459	0.01086	0.001662	6.43E-11
rs42122	7	72841823	FZD9	G	A	0.161	0.014852	0.002281	7.41E-11
rs615552	9	22026077	CDKN2B-AS1	T	C	0.438	0.010866	0.001669	7.52E-11

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs6682862	1	177938437	RP4-798P15.3	G	A	0.162	-0.01449	0.002231	8.30E-11
rs74907759	6	160857193	SLC22A3	A	G	0.0093	-0.05917	0.009116	8.54E-11
rs140365836	19	45284978	CBLC	G	A	0.00436	0.113169	0.017455	8.95E-11
rs793900	4	185253019	RP11-290F5.2	C	T	0.408	0.01075	0.001664	1.03E-10
rs4946713	6	106374015	RN7SKP211	C	A	0.443	-0.01068	0.001653	1.03E-10
rs77502095	12	133119022	FBRSL1	G	A	0.141	0.016054	0.002487	1.08E-10
rs1618633	19	2821030	ZNF554	C	G	0.353	0.013548	0.002103	1.17E-10
rs75393372	6	139187883	ECT2L	T	C	0.0776	0.019691	0.003071	1.43E-10
rs72971192	6	127088318	RPS4XP9	T	C	0.237	-0.01278	0.002002	1.70E-10
rs6107650	20	5526113	GPCPD1	A	G	0.39	-0.01084	0.001703	1.94E-10
rs603424	10	102075479	PKD2L1	G	A	0.169	0.014236	0.002238	2.02E-10
rs13057311	22	50878196	PPP6R2	G	A	0.267	-0.01211	0.001909	2.22E-10
rs55921103	3	69810294	MITF	G	T	0.364	0.011092	0.001751	2.37E-10
rs17776811	14	90066451	RP11-33N16.3	A	C	0.403	-0.01057	0.00167	2.42E-10
rs61886346	10	96101364	NOC3L	C	T	0.059	-0.02203	0.003496	2.96E-10
rs35046559	20	62344480	RP4-583P15.15	A	G	0.246	0.012059	0.001915	3.05E-10
rs207637	13	33085469	N4BP2L2	T	G	0.385	0.010671	0.001695	3.08E-10
rs9410207	9	91404799	MIR4289	T	C	0.066	-0.02099	0.003339	3.25E-10
rs115458560	1	110061974	AMIGO1	T	C	0.0141	-0.04659	0.007424	3.49E-10
rs17526980	8	28950309	KIF13B	C	T	0.0298	0.029566	0.004737	4.34E-10
rs10936349	3	100453859	TFG	T	C	0.341	-0.01119	0.0018	5.07E-10
rs8041391	15	89102448	RP11-97O12.3	C	T	0.154	0.014712	0.002374	5.72E-10
rs60417583	20	62890877	PCMTD2	G	T	0.165	0.014687	0.002372	5.92E-10
rs2297400	9	107599481	ABCA1	T	C	0.125	0.015622	0.002525	6.10E-10

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs11650379	17	29460252	RP11-142O6.1	A	G	0.0128	-0.04748	0.007692	6.70E-10
rs112575086	4	54429890	LNX1	C	T	0.121	-0.01546	0.002506	6.91E-10
rs56223611	7	87083157	ABCB4	C	A	0.142	-0.01448	0.002349	7.07E-10
rs28811342	17	18125845	LLGL1	T	C	0.2	0.012613	0.002055	8.32E-10
rs13112099	4	69537733	UGT2B15	T	G	0.471	0.011357	0.001851	8.50E-10
rs756559	15	101907189	PCSK6	A	G	0.31	-0.01078	0.001762	9.42E-10
rs2085723	4	151206057	LRBA	G	T	0.342	0.010544	0.001725	9.85E-10
rs17586023	4	100272950	ADH1C	A	G	0.113	0.015989	0.002618	1.02E-09
rs4804815	19	7842669	CLEC4M	C	G	0.264	-0.01167	0.001913	1.05E-09
rs11761517	7	158532706	ESYT2	C	T	0.458	-0.01041	0.001712	1.18E-09
rs4239702	20	44749251	CD40	T	C	0.282	0.010985	0.001809	1.27E-09
rs7300593	12	53790450	SP1	T	C	0.172	0.01314	0.002167	1.34E-09
rs147654565	12	21029965	SLCO1B7	G	A	0.0323	-0.0283	0.004669	1.35E-09
rs72647039	4	74321365	AFP	C	T	0.0201	0.036537	0.006033	1.40E-09
rs6794351	3	136292279	STAG1	C	T	0.24	0.012329	0.002039	1.47E-09
rs4760388	12	92802078	CLLU1OS	T	C	0.478	0.009971	0.00165	1.50E-09
rs13005230	2	21914611	AC018742.1	C	A	0.0962	-0.02001	0.003313	1.54E-09
rs13403394	2	21942492	AC067959.1	C	T	0.0316	-0.02816	0.004675	1.72E-09
rs76643124	9	136356448	SLC2A6	G	A	0.0286	-0.02979	0.004962	1.94E-09
rs189527671	2	202945407	AC079354.1	G	A	0.0228	0.038922	0.006501	2.14E-09
rs10462024	17	8051639	PER1	A	G	0.415	0.010412	0.001739	2.16E-09
rs880973	2	21078879	AC012361.1	A	G	0.223	-0.01182	0.001976	2.22E-09
rs61882680	11	46370636	DGKZ	C	T	0.0295	-0.02937	0.004928	2.51E-09
rs9824581	3	125066925	ZNF148	G	A	0.267	0.011002	0.001847	2.58E-09
rs77631946	11	63688374	RCOR2	C	A	0.0895	-0.01748	0.002935	2.60E-09
rs4234798	4	7219933	SORCS2	T	G	0.393	0.009946	0.001672	2.72E-09
rs145275181	5	74201340	CTD-2503O16.4	T	G	0.0107	0.055208	0.009287	2.77E-09

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs113548930	7	44847350	PPIA	T	A	0.0374	0.026746	0.004511	3.05E-09
rs34927723	7	21422925	RNU1-15P	C	T	0.157	-0.01351	0.00228	3.09E-09
rs2247139	12	122380274	WDR66	A	G	0.136	-0.01415	0.002389	3.15E-09
rs576285576	16	88527501	ZFPM1	A	G	0.111	-0.02378	0.00402	3.33E-09
rs10229853	7	155022821	AC099552.3	A	G	0.0873	0.017633	0.002981	3.34E-09
rs1044573	20	25206654	ENTPD6	A	G	0.495	0.009854	0.001678	4.28E-09
rs16861497	3	186731327	ST6GAL1	G	T	0.395	-0.00986	0.001679	4.37E-09
rs368280	8	81379656	Y_RNA	C	A	0.253	0.011092	0.001891	4.46E-09
rs59593290	17	65813326	AC006534.1	G	A	0.0843	0.017419	0.00297	4.48E-09
rs241056	5	176528106	FGFR4	G	A	0.256	-0.01126	0.001931	5.50E-09
rs10818580	9	124415022	DAB2IP	G	A	0.263	0.011009	0.001891	5.78E-09
rs72729610	4	154190965	TRIM2	A	G	0.167	-0.01274	0.002191	5.99E-09
rs11075253	16	15148646	NTAN1	C	A	0.302	0.010562	0.001818	6.27E-09
rs61754230	12	72179446	RAB21	C	T	0.0136	0.042023	0.00727	7.46E-09
rs11000443	10	74665640	OIT3	C	A	0.0433	-0.02304	0.003998	8.28E-09
rs117590445	20	61321110	SLCO4A1	C	T	0.014	-0.04396	0.007634	8.48E-09
rs62075819	17	46982544	SUMO2P17	C	T	0.358	0.010105	0.001755	8.50E-09
rs7404072	16	73096750	RP11-346C20.3	C	T	0.289	0.01056	0.001834	8.52E-09
rs342467	4	88052219	AFF1	T	C	0.401	0.009634	0.001674	8.75E-09
rs11159099	14	75235397	YLPM1	A	G	0.461	-0.00951	0.001667	1.17E-08
rs138204164	4	120123417	RP11-455G16.1	C	G	0.135	-0.01379	0.00242	1.20E-08
rs1136348	1	16661427	FBXO42	C	T	0.154	-0.01333	0.002345	1.31E-08
rs62075782	17	40536804	LRG_112	G	C	0.226	-0.01212	0.002132	1.33E-08
rs41279736	1	110043057	CYB561D1	A	G	0.0287	0.031727	0.00559	1.38E-08
rs113371168	6	11837368	RP3-420J14.1	G	A	0.0678	-0.01849	0.003259	1.39E-08
rs62399909	6	34339793	RPS10-NUDT3	C	A	0.0868	0.016464	0.002918	1.67E-08
rs2302434	7	75630183	STYXL1	C	T	0.177	0.01216	0.002166	1.99E-08

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs9909417	17	18679442	FBXW10	G	A	0.288	0.01031	0.001837	2.01E-08
rs12454507	18	77825981	RBFADN	A	G	0.288	0.01065	0.001898	2.01E-08
rs1870140	10	82246749	TSPAN14	A	G	0.159	0.012703	0.002268	2.13E-08
rs1031101	10	54533360	MBL2	A	G	0.142	0.013044	0.00233	2.16E-08
rs189974330	1	109729937	KIAA1324	G	A	0.00737	0.055546	0.009924	2.18E-08
rs28375625	15	49742494	FGF7	A	C	0.459	0.009294	0.001664	2.33E-08
rs77860716	11	95456770	RP11-644L4.1	A	G	0.0484	0.022583	0.004052	2.50E-08
rs78531123	10	104082953	GBF1	G	A	0.0576	0.019672	0.003531	2.52E-08
rs543486395	5	156375335	TIMD4	C	A	0.0236	-0.04216	0.007573	2.59E-08
rs17532301	13	41609047	ELF1	G	A	0.0679	-0.01818	0.003274	2.83E-08
rs235343	21	46249752	SUMO3	A	C	0.436	0.009349	0.00169	3.19E-08
rs76738473	19	33870906	CEBPG	T	C	0.0279	-0.02698	0.004912	3.98E-08
rs13357800	5	156422653	TIMD4	C	T	0.426	0.009137	0.001665	4.06E-08
rs1571536	9	92215638	GADD45G	T	C	0.485	-0.009	0.001643	4.26E-08
rs2469204	15	77256130	RN7SL278P	T	G	0.379	-0.0095	0.001734	4.27E-08
rs825474	12	124567132	FAM101A	G	A	0.416	0.009049	0.001654	4.46E-08
rs16967478	16	15864284	MYH11	G	A	0.157	-0.01223	0.00224	4.74E-08
rs11846741	14	64238580	RP11-804M7.2	A	G	0.0742	0.017111	0.003135	4.83E-08
rs10814052	9	33853571	UBE2R2	C	A	0.162	0.012171	0.002232	4.93E-08
rs6426060	1	224384711	DEGS1	A	T	0.119	0.01386	0.002542	4.98E-08

Abbreviations. **LDL-C**: Low-Density Lipoprotein Cholesterol; **SNP**: Single Nucleotide Polymorphism; **RSID**: Reference SNP Cluster ID; **GRCh37**: Genome Reference Consortium Human Build 37

Table S5. HDL-C SNPs for European Populations used in MR analysis

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs6728178	2	21193946	RP11-116D2.1	G	A	0.238	0.057192	0.001866	1.00E-200
rs6984305	8	9178268	RP11-115J16.1	A	T	0.111	0.078118	0.00253	1.00E-200
rs11986182	8	19727447	INTS10	A	T	0.163	0.066349	0.002182	1.00E-200
rs62568179	9	107643786	ABCA1	G	A	0.122	-0.07546	0.002435	1.00E-200
rs11216117	11	116596129	BUD13	C	A	0.0707	-0.0985	0.003093	1.00E-200
rs145347194	15	58670135	ALDH1A2	T	C	0.0867	0.09298	0.003007	1.00E-200
rs17821316	15	58703069	LIPC	A	C	0.485	-0.06771	0.001666	1.00E-200
rs77050717	16	56767381	NUP93	G	A	0.0153	-0.20667	0.006548	1.00E-200
rs72786781	16	56970210	HERPUD1	T	A	0.0255	0.170665	0.005302	1.00E-200
rs12444708	16	56984482	HERPUD1	C	T	0.125	0.077391	0.002505	1.00E-200
rs16942887	16	67928042	PSKH1	G	A	0.124	0.072873	0.002407	1.00E-200
rs76868109	17	41776943	CTC-501O10.1	A	C	0.0305	-0.15745	0.005113	1.00E-200
rs80204526	18	47066144	LIPG	C	A	0.00999	0.277406	0.008356	1.00E-200
rs4121823	18	47144223	LIPG	T	A	0.16	0.06944	0.002202	1.00E-200
rs116843064	19	8429323	ANGPTL4	G	A	0.0201	0.217611	0.00571	1.00E-200
rs6857	19	45392254	CTB-129P6.4	C	T	0.169	-0.06823	0.002153	1.00E-200
rs1800961	20	43042364	HNF4A	C	T	0.0347	-0.13115	0.004379	4.47E-197
rs103294	19	54797848	AC010518.3	C	T	0.213	0.057212	0.001952	8.98E-189
rs150799387	8	19766707	LPL	G	A	0.0135	-0.20293	0.006951	2.47E-187
rs11071371	15	58576226	ALDH1A2	C	T	0.129	-0.06988	0.002394	2.83E-187
rs72771479	16	57002121	CETP	A	T	0.0169	0.182532	0.006439	9.45E-177
rs4846914	1	230295691	GALNT2	G	A	0.402	0.045854	0.001619	1.72E-176
rs10838684	11	47283862	NR1H3	A	G	0.314	0.044598	0.001746	7.38E-144
rs2954038	8	126507389	RP11-136O12.2	C	A	0.304	0.043892	0.001744	8.16E-140
rs13107325	4	103188709	SLC39A8	C	T	0.0596	-0.08104	0.003273	2.40E-135
rs61941677	12	125325010	SCARB1	C	G	0.13	-0.05965	0.002443	1.28E-131

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs174545	11	61569306	FADS1	C	G	0.347	-0.04077	0.001675	7.63E-131
rs7200805	16	56989015	AC012181.1	T	C	0.0271	0.121368	0.0051	3.46E-125
rs3824477	9	107588328	ABCA1	G	A	0.0331	0.105132	0.004558	1.08E-117
rs838876	12	125259888	SCARB1	A	G	0.332	-0.03921	0.001759	4.74E-110
rs2943652	2	227108446	AC068138.1	C	T	0.369	-0.03623	0.001658	6.83E-106
rs3768321	1	40035928	RP11-69E11.8	G	T	0.191	-0.0443	0.002072	2.16E-101
rs686030	9	15304782	TTC39B	C	A	0.142	0.048229	0.002277	1.55E-99
rs11751347	6	161092438	RP1-81D8.4	C	T	0.0939	-0.05661	0.002739	6.66E-95
rs147464145	16	56988503	AC012181.1	C	T	0.0104	0.160384	0.007891	7.54E-92
rs55707100	15	43820717	MAP1A	C	T	0.027	-0.09975	0.004943	1.46E-90
rs113286621	16	56969016	HERPUD1	G	A	0.0105	0.166657	0.008278	3.84E-90
rs141469619	11	116714293	SIK3	A	G	0.00874	-0.18894	0.009415	1.41E-89
rs629301	1	109818306	CELSR2	G	T	0.219	-0.03789	0.001909	1.18E-87
rs4969182	17	76393030	PGS1	T	C	0.492	0.030908	0.001592	6.17E-84
rs2925979	16	81534790	CMIP	T	C	0.301	0.033452	0.001733	4.62E-83
rs55747707	7	73037366	MLXIPL	G	A	0.194	0.038846	0.002035	3.31E-81
rs2293889	8	116599199	TRPS1	T	G	0.422	0.030622	0.001619	9.04E-80
rs737337	19	11347493	DOCK6	T	C	0.0871	-0.05232	0.002806	1.28E-77
rs2744944	6	34658080	C6orf106	A	C	0.363	0.03121	0.001711	2.38E-74
rs881844	17	37810218	STARD3	C	G	0.337	0.031158	0.001713	6.56E-74
rs77617917	20	44563217	PCIF1	G	A	0.0761	-0.05442	0.003002	1.91E-73
rs11762784	7	130424646	KLF14	G	A	0.495	0.029563	0.001636	5.47E-73
rs78058190	2	219699999	PRKAG3	G	A	0.0529	-0.07489	0.004157	1.39E-72
rs12721090	11	116700775	APOC3	C	T	0.0333	0.080683	0.004486	2.62E-72
rs2070512	22	21949411	UBE2L3	A	C	0.21	-0.03508	0.001961	1.54E-71
rs13389219	2	165528876	COBLL1	C	T	0.403	0.028923	0.001625	7.18E-71
rs7954144	12	110015893	LRG_156	A	G	0.47	0.028426	0.001608	6.35E-70
rs1109560	12	123757905	CDK2AP1	C	A	0.0938	0.04649	0.00276	1.21E-63
rs2792751	10	113940329	GPAM	T	C	0.284	-0.02904	0.00175	7.65E-62

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs2777802	9	107569337	ABCA1	T	C	0.23	0.031378	0.001896	1.68E-61
rs75460349	1	27180088	ZDHHC18	A	C	0.0237	-0.08447	0.005158	2.83E-60
rs3767964	1	93815490	DR1	T	C	0.362	-0.02735	0.001683	2.10E-59
rs1689796	1	182151236	GS1-122H1.2	T	G	0.325	-0.0273	0.001705	1.01E-57
rs6905288	6	43758873	VEGFA	G	A	0.442	-0.02646	0.001664	6.58E-57
rs79949326	7	6461310	DAGLB	C	T	0.249	0.029538	0.001865	1.73E-56
rs144311893	19	45423944	APOC1	C	T	0.0186	0.091044	0.006019	1.09E-51
rs2494747	14	105258437	AKT1	G	T	0.385	-0.0264	0.001751	2.17E-51
rs35511894	15	58723709	LIPC	C	T	0.00926	0.130649	0.008777	4.15E-50
rs28650790	5	55861464	AC022431.2	C	T	0.176	-0.03052	0.002075	5.59E-49
rs2642438	1	220970028	MARC1	A	G	0.293	0.025783	0.001759	1.26E-48
rs837500	12	125077567	NCOR2	C	T	0.369	0.025374	0.001752	1.60E-47
rs11172147	12	57696677	R3HDM2	G	A	0.247	0.027062	0.00187	1.75E-47
rs10242866	7	17920613	SNX13	C	T	0.392	-0.0234	0.001646	7.13E-46
rs7613875	3	49971514	MON1A	C	A	0.478	-0.02283	0.001614	2.00E-45
rs148754782	8	19910576	AC100802.3	G	A	0.0313	0.065897	0.004676	4.19E-45
rs17585887	6	139835498	RP11-12A2.3	T	C	0.421	0.022612	0.001608	6.85E-45
rs34931250	17	66879927	ABCA8	C	T	0.0587	-0.04874	0.003483	1.67E-44
rs61854123	10	46065451	MARCH8	G	A	0.241	0.026326	0.001885	2.41E-44
rs673335	11	75450576	RN7SL786P	T	C	0.18	-0.02887	0.002072	3.98E-44
rs74444445	8	19852491	LPL	T	C	0.021	-0.07853	0.005709	4.63E-43
rs687339	3	135932359	RP11-463H24.1	C	T	0.216	-0.02694	0.001963	7.01E-43
rs10049088	3	156797648	LINC00880	C	T	0.383	0.022548	0.001643	7.38E-43
rs139524394	11	116624922	BUD13	C	T	0.0105	-0.11302	0.008362	1.26E-41
rs2800708	6	127437617	RSPO3	C	T	0.471	-0.02139	0.001605	1.61E-40
rs56731460	8	9194670	RP11-115J16.1	G	A	0.226	0.025418	0.001942	3.70E-39
rs72647336	8	126445055	TRIB1	G	A	0.0456	-0.05168	0.004058	3.78E-37
rs1516725	3	185824004	DGKG	T	C	0.133	0.03002	0.002378	1.53E-36
rs138962675	11	116789027	SIK3	C	T	0.00997	0.107989	0.008563	1.83E-36

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs57316347	12	124389053	DNAH10	A	G	0.312	0.021963	0.001754	5.90E-36
rs4418728	10	94839724	CYP26A1	G	T	0.452	0.020024	0.001609	1.45E-35
rs17875416	10	115796337	ADRB1	A	G	0.115	0.030296	0.00248	2.52E-34
rs28418580	4	89742244	FAM13A	C	T	0.463	-0.01943	0.001599	5.33E-34
rs34409228	7	1102674	C7orf50	C	T	0.126	0.029463	0.002446	2.01E-33
rs11231694	11	63862882	RP11-21A7A.2	G	C	0.0569	-0.04154	0.003461	3.46E-33
rs56925758	8	144299352	GPIHBP1	G	A	0.197	0.024403	0.002034	3.78E-33
rs113271699	11	116650571	ZNF259	C	A	0.0142	0.081201	0.006794	6.29E-33
rs2899297	22	38594668	PLA2G6	G	A	0.401	-0.01962	0.001642	6.36E-33
rs267738	1	150940625	CERS2	T	G	0.209	0.023227	0.001964	2.91E-32
rs4080140	11	116676590	AP006216.5	G	A	0.0233	0.063713	0.005405	4.50E-32
rs137967290	9	107666614	ABCA1	G	A	0.0163	0.075881	0.006452	6.16E-32
rs4077194	1	178533832	RNA5SP69	G	T	0.487	-0.01881	0.001601	7.09E-32
rs4805881	19	33896432	PEPD	A	C	0.334	0.019646	0.001693	3.80E-31
rs235374	21	46294986	PTTG1IP	G	C	0.431	-0.01898	0.001638	4.75E-31
rs79968526	16	56911069	SLC12A3	T	C	0.023	-0.06548	0.005656	5.45E-31
rs7134375	12	20473758	RP11-284H19.1	C	A	0.433	0.018382	0.001606	2.47E-30
rs35137994	19	8429066	ANGPTL4	C	T	0.0591	0.038914	0.003424	6.23E-30
rs6088461	20	32905161	AHCY	G	T	0.425	-0.01895	0.001669	7.39E-30
rs142638686	8	19892948	AC100802.3	C	T	0.012	0.083914	0.007399	8.14E-30
rs76716928	6	31295941	XXbac-BPG248L24.10	C	G	0.364	-0.02077	0.001843	1.80E-29
rs33999310	3	52883332	TMEM110	G	A	0.104	0.029234	0.00261	4.00E-29
rs12549603	8	126632639	RP11-136O12.2	G	A	0.277	-0.0198	0.001776	6.91E-29
rs11065979	12	112059557	BRAP	C	T	0.427	-0.01814	0.001656	6.28E-28
rs6542680	2	3640142	LRG_350	C	T	0.197	-0.02229	0.002039	8.13E-28
rs74833298	3	12343549	PPARG	C	A	0.0275	-0.05291	0.004846	9.31E-28
rs1650527	5	158022724	RP11-32D16.1	C	T	0.226	-0.02071	0.001903	1.35E-27

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs112792569	16	56658745	MT1E	G	C	0.0188	0.063901	0.005896	2.27E-27
rs1425485	4	157681890	PDGFC	C	T	0.23	0.020519	0.001895	2.59E-27
rs151235402	20	569164	TCF15	C	T	0.0168	-0.06927	0.006402	2.78E-27
rs2305637	3	47045846	NBEAL2	C	T	0.17	-0.02326	0.002155	3.71E-27
rs2306363	11	65405600	SIPA1	G	T	0.199	0.021846	0.00204	9.22E-27
rs13379042	14	74250100	ELMSAN1	T	C	0.266	0.019332	0.001806	9.53E-27
rs553427	1	234852760	RP4-781K5.7	C	T	0.469	0.017059	0.001624	8.52E-26
rs1415293	1	219730006	RP11-95P13.2	A	T	0.312	0.018046	0.00173	1.74E-25
rs73165503	7	150519279	AOC1	G	C	0.0955	-0.02766	0.002679	5.49E-25
rs4375701	17	45712218	RP11-580I16.2	C	T	0.488	0.016692	0.001623	8.05E-25
rs2540949	2	65284231	CEP68	A	T	0.387	0.016536	0.001644	8.55E-24
rs188501176	18	47325634	RP11-886H22.1	A	G	0.0142	-0.07141	0.007106	9.21E-24
rs4946955	6	109639710	CCDC162P	C	G	0.307	-0.01719	0.001726	2.31E-23
rs11644601	16	15172118	RRN3	T	C	0.299	0.017711	0.001779	2.43E-23
rs72801474	5	132444128	HSPA4	G	A	0.0964	0.027674	0.002799	4.66E-23
rs12533197	7	36192606	EEPD1	T	G	0.445	-0.01599	0.001626	8.05E-23
rs7581601	2	203497786	FAM117B	A	C	0.317	-0.01692	0.001724	9.63E-23
rs149074578	20	44625065	MMP9	C	A	0.0236	-0.0536	0.005467	1.08E-22
rs2862954	10	101912064	ERLIN1	T	C	0.474	0.015558	0.001595	1.73E-22
rs6710091	2	239597	SH3YL1	C	G	0.351	0.016353	0.001694	4.76E-22
rs523288	18	57848369	RP11-795H16.2	A	T	0.239	-0.01837	0.001908	6.18E-22
rs35374654	3	185926247	DGKG	A	T	0.223	0.019523	0.002031	6.98E-22
rs11694828	2	228565656	SLC19A3	G	A	0.444	0.016018	0.001675	1.13E-21
rs2307111	5	75003678	POC5	T	C	0.394	0.015632	0.001634	1.13E-21
rs3803800	17	7462969	TNFSF13	A	G	0.22	0.018318	0.001924	1.72E-21
rs11766379	7	150208971	STRADBP1	T	C	0.246	-0.01779	0.00188	2.95E-21
rs12358424	10	33641149	RP11-90B22.1	T	G	0.0888	-0.02702	0.00286	3.46E-21
rs187544997	8	19890641	AC100802.3	C	G	0.0125	-0.07304	0.007735	3.63E-21
rs9286789	6	32439696	HLA-DRB9	T	G	0.0504	-0.03787	0.004044	7.60E-21

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs112540547	16	68844713	LRG_301	A	G	0.0565	-0.03328	0.003556	8.20E-21
rs111933537	12	123248255	DENR	A	C	0.0878	0.027963	0.00303	2.77E-20
rs34246353	8	71138299	NCOA2	T	C	0.0939	-0.02496	0.002709	3.19E-20
rs9976784	21	46907479	COL18A1	G	A	0.202	-0.01829	0.001989	3.77E-20
rs3775217	4	88019467	AFF1	A	G	0.378	0.015079	0.001644	4.59E-20
rs117687565	18	47147524	LIPG	C	T	0.0123	0.068638	0.007506	5.98E-20
rs3747973	1	205677148	NUCKS1	A	G	0.414	0.01505	0.001647	6.45E-20
rs1664781	5	53276301	ARL15	G	A	0.314	-0.0156	0.001712	8.02E-20
rs3815026	1	110230073	GSTM2	C	G	0.471	0.023442	0.002579	9.86E-20
rs3900595	11	14313329	RRAS2	T	G	0.128	0.021781	0.002418	2.13E-19
rs9608955	22	30900829	SEC14L4	C	T	0.226	-0.01709	0.001898	2.26E-19
rs1047891	2	211540507	LRG_336	C	A	0.304	-0.0157	0.001756	3.83E-19
rs3739020	2	136602491	MCM6	T	G	0.229	0.017795	0.002004	6.76E-19
rs876039	7	50308811	AC020743.4	G	C	0.312	0.015172	0.001713	8.41E-19
rs4714556	6	41993229	CCND3	A	G	0.382	0.014691	0.001659	8.45E-19
rs12502605	4	26031528	RP11-324H7.1	T	C	0.18	-0.01818	0.002072	1.70E-18
rs1767141	1	23734350	TCEA3	A	C	0.0938	0.024298	0.002778	2.20E-18
rs34311866	4	951947	TMEM175	T	C	0.185	-0.01791	0.002062	3.72E-18
rs7920002	10	17259766	VIM-AS1	T	A	0.416	0.01427	0.001647	4.46E-18
rs1661052	11	2942593	SLC22A18	G	A	0.0917	0.024859	0.002874	5.17E-18
rs11170538	12	53796744	SP1	T	C	0.478	-0.01673	0.001935	5.23E-18
rs7404696	16	1012393	LMF1	T	C	0.275	0.015844	0.001835	6.01E-18
rs3732356	3	119529113	NR1I2	G	T	0.0688	-0.02763	0.003208	7.17E-18
rs7622114	3	36960660	TRANK1	C	A	0.431	0.013971	0.001627	8.92E-18
rs10767659	11	27686196	BDNF	G	T	0.326	0.01457	0.001702	1.14E-17
rs3806415	1	156698265	RRNAD1	C	T	0.323	0.0144	0.001704	2.90E-17
rs4675812	2	242395674	FARP2	G	A	0.409	0.013704	0.001629	4.04E-17
rs118175831	18	47066204	LIPG	C	T	0.0236	0.046108	0.005486	4.28E-17
rs672341	6	153455994	RGS17	G	A	0.404	0.013646	0.001636	7.45E-17

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs79634051	11	14561945	PSMA1	G	C	0.0243	0.046524	0.005579	7.48E-17
rs12234017	5	118711834	TNFAIP8	A	G	0.259	0.015209	0.001827	8.52E-17
rs11688682	2	121347612	AC073257.1	G	C	0.249	0.01821	0.002192	9.76E-17
rs9374588	6	116311701	FRK	G	A	0.326	0.014143	0.001706	1.14E-16
rs1652343	18	21131929	NPC1	T	C	0.491	0.013217	0.001599	1.39E-16
rs1889930	10	95320505	FFAR4	G	A	0.201	-0.01641	0.001992	1.75E-16
rs4804101	19	7220745	INSR	G	T	0.436	-0.01336	0.001628	2.24E-16
rs140201358	11	823586	PNPLA2	C	G	0.0127	-0.05905	0.007221	2.89E-16
rs549058	13	51201045	DLEU1	G	T	0.126	0.019587	0.002403	3.65E-16
rs9694004	8	9309720	RP11-375N15.1	C	T	0.495	0.012951	0.001604	6.89E-16
rs2519093	9	136141870	ABO	C	T	0.188	0.016195	0.002008	7.31E-16
rs2267662	6	35250440	ZNF76	C	A	0.147	-0.0182	0.00226	8.09E-16
rs4741016	9	1036132	DMRT2	T	C	0.229	-0.01515	0.001909	2.11E-15
rs1421656	5	72928034	ARHGEF28	A	G	0.282	0.014139	0.001783	2.17E-15
rs10039241	5	176578072	LRG_512	G	A	0.149	0.017666	0.002237	2.86E-15
rs6968554	7	17287106	AC003075.4	A	G	0.365	-0.0131	0.001662	3.15E-15
rs39324	7	116969876	WNT2	A	G	0.367	-0.01333	0.001695	3.78E-15
rs11905856	20	46356466	SULF2	A	G	0.187	0.016521	0.002118	6.17E-15
rs1017547	15	63563303	RAB8B	G	A	0.47	0.012551	0.001611	6.77E-15
rs157521	5	55821031	AC022431.2	T	G	0.451	-0.01237	0.00159	7.23E-15
rs9367175	6	43787755	RP11-344J7.2	G	A	0.306	0.01335	0.001726	1.04E-14
rs11879735	19	48147444	GLTSCR1	A	T	0.113	-0.01944	0.002524	1.35E-14
rs8126001	20	62711459	OPRL1	C	T	0.482	0.013516	0.001759	1.54E-14
rs144164168	20	44476037	ACOT8	C	T	0.0203	-0.04483	0.005843	1.68E-14
rs2218584	10	118408642	PNLIPRP2	C	T	0.422	-0.01244	0.001625	1.88E-14
rs12462758	19	52439687	ZNF613	C	T	0.229	-0.01453	0.001904	2.36E-14
rs7298463	12	6663964	IFFO1	G	C	0.169	-0.01657	0.002174	2.50E-14
rs2090032	2	20372547	RN7SL140P	G	C	0.491	-0.01219	0.001601	2.70E-14
rs76086106	5	39491186	DAB2	G	A	0.0439	0.030818	0.004052	2.86E-14

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs1412234	9	28410683	LINGO2	T	C	0.316	-0.0131	0.001724	2.92E-14
rs2534600	7	38280224	TRGC2	A	C	0.418	0.012689	0.001671	3.12E-14
rs76533352	16	57150943	CPNE2	T	G	0.0696	0.02411	0.003181	3.45E-14
rs2455850	3	15782711	ANKRD28	A	C	0.313	0.013383	0.001771	4.18E-14
rs79612456	15	58766286	RP11-355N15.1	T	G	0.0349	0.035292	0.004676	4.42E-14
rs13734	20	17594729	RRBP1	G	A	0.181	-0.01559	0.002071	5.28E-14
rs116862304	16	56388893	GNAO1	C	T	0.0179	0.048287	0.006418	5.33E-14
rs4818226	21	42633065	BACE2	A	G	0.326	-0.01288	0.001713	5.39E-14
rs275182	15	39446211	RP11-624L4.1	A	G	0.17	0.016071	0.002145	6.70E-14
rs78424108	13	41673457	MIR3168	C	G	0.0694	-0.0236	0.00315	6.82E-14
rs1250530	10	81017467	ZMIZ1	T	C	0.304	-0.01327	0.001771	6.86E-14
rs1861410	2	58933591	LINC01122	C	T	0.451	0.012075	0.001618	8.41E-14
rs12888955	14	103256877	LRG_229	G	A	0.347	0.012446	0.001671	9.61E-14
rs72844048	11	3105622	AC108448.2	G	A	0.0147	-0.05353	0.007198	1.03E-13
rs12216486	6	160329616	MAS1	G	A	0.0909	-0.02062	0.002784	1.28E-13
rs2121138	19	4024891	PIAS4	G	T	0.071	-0.02424	0.003273	1.31E-13
rs4904550	14	89804856	FOXN3	G	A	0.197	0.014908	0.002043	2.93E-13
rs4794008	17	47048961	GIP	C	T	0.279	0.013444	0.001851	3.73E-13
rs11656743	17	65980289	BPTF	G	A	0.216	0.01428	0.001969	4.10E-13
rs67638406	8	106404215	ZFPM2	A	G	0.352	-0.01232	0.001701	4.43E-13
rs72805157	16	85888214	RP11-568J23.1	C	G	0.12	-0.02004	0.002776	5.13E-13
rs73068475	7	26397540	SNX10	G	A	0.124	-0.01748	0.002424	5.52E-13
rs13076933	3	12327431	PPARG	T	G	0.256	-0.0133	0.001844	5.55E-13
rs112200607	20	30167514	RNU6-384P	G	A	0.162	0.015745	0.002187	6.11E-13
rs55888929	6	139301177	REPS1	C	T	0.153	0.016132	0.002247	7.02E-13
rs2811466	3	129299265	PLXND1	G	A	0.104	0.018846	0.002626	7.13E-13
rs938587	3	185884216	DGKG	A	T	0.263	0.013111	0.001828	7.42E-13
rs2254069	10	122875589	RP11-159H3.1	G	A	0.116	-0.01751	0.002449	8.66E-13
rs9923623	16	11468267	CTD-3088G3.8	C	T	0.335	0.012621	0.001766	8.84E-13

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs6066371	20	46143914	NCOA3	A	T	0.332	0.015043	0.002105	8.85E-13
rs805317	2	54135004	PSME4	T	A	0.453	0.011507	0.001613	9.82E-13
rs1761662	7	80238463	CD36	C	T	0.463	0.011454	0.001607	1.01E-12
rs6750720	2	100670745	AFF3	C	G	0.336	0.012037	0.001689	1.02E-12
rs12931143	16	88006065	BANP	T	C	0.469	-0.01205	0.001691	1.04E-12
rs1438588	5	153444950	MFAP3	A	G	0.366	-0.01184	0.001666	1.21E-12
rs35662434	4	69339933	TMPRSS11E	A	G	0.214	0.014539	0.002054	1.45E-12
rs62328321	4	143319721	INPP4B	C	T	0.364	0.011703	0.001654	1.48E-12
rs2185454	14	65870594	FUT8	G	T	0.339	0.012164	0.00172	1.50E-12
rs10188334	2	653874	TMEM18	C	T	0.169	0.015234	0.002154	1.53E-12
rs76428106	13	28604007	LRG_457	T	C	0.0112	-0.05871	0.008304	1.55E-12
rs73029263	6	164113762	RP1-230L10.1	A	G	0.13	0.016763	0.002374	1.64E-12
rs2243616	12	121406370	HNF1A-AS1	G	T	0.352	-0.01207	0.001709	1.64E-12
rs72771232	5	82857040	VCAN-AS1	T	C	0.0744	0.021833	0.003095	1.74E-12
rs7132908	12	50263148	FAIM2	G	A	0.385	-0.01152	0.001637	1.93E-12
rs7107356	11	47676170	AGBL2	A	G	0.492	0.011455	0.001628	1.99E-12
rs12641981	4	45179883	RP11-362I1.1	C	T	0.427	-0.01129	0.001605	2.06E-12
rs1168040	1	62960575	DOCK7	T	C	0.318	0.012301	0.001751	2.14E-12
rs12434231	14	100349781	EML1	T	C	0.157	-0.01565	0.002229	2.20E-12
rs17326656	2	48962291	LHCGR	G	T	0.228	-0.01343	0.001913	2.24E-12
rs3814883	16	29994922	TAOK2	C	T	0.464	-0.01156	0.001648	2.33E-12
rs11870735	17	481604	VPS53	C	T	0.175	-0.01475	0.002108	2.61E-12
rs2613496	1	72754788	NEGR1	T	C	0.18	-0.01462	0.0021	3.37E-12
rs74625348	18	60846430	BCL2	G	C	0.231	0.013464	0.001935	3.48E-12
rs12143329	1	202197659	LGR6	C	T	0.403	0.011352	0.001634	3.73E-12
rs11651957	17	17497409	RP11-524F11.2	G	A	0.0598	-0.02359	0.0034	3.99E-12
rs10463352	5	141710499	AC005592.2	G	A	0.425	0.011348	0.001638	4.25E-12
rs10826337	10	61409469	SLC16A9	G	A	0.42	-0.01135	0.001641	4.63E-12
rs75246752	1	145630111	RNF115	G	C	0.00764	0.069414	0.010039	4.69E-12

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs16940712	15	59027139	ADAM10	A	G	0.0241	-0.03709	0.005375	5.17E-12
rs9604045	13	113927208	LDHBP1	G	T	0.234	0.01527	0.002214	5.27E-12
rs6506032	18	287534	RP11-705O1.8	G	A	0.0743	0.020694	0.003012	6.39E-12
rs34511042	14	70816750	SYNJ2BP-COX16	G	A	0.366	-0.01146	0.001671	6.95E-12
rs2172071	10	122968030	LINC01153	C	T	0.316	0.011798	0.001725	7.89E-12
rs7541031	1	28374815	EYA3	C	T	0.274	-0.01248	0.001827	8.44E-12
rs3803357	15	40751555	BAHD1	C	A	0.499	0.010942	0.001606	9.56E-12
rs7940105	11	68732026	MRGPRD	G	A	0.273	-0.01257	0.001846	9.68E-12
rs11171739	12	56470625	ERBB3	C	T	0.417	-0.01099	0.001616	1.04E-11
rs75772194	8	36846435	AC090453.1	A	G	0.166	-0.01446	0.002128	1.09E-11
rs1133483	4	100023923	RP11-696N14.1	G	C	0.43	-0.01098	0.001619	1.16E-11
rs11038700	11	45899791	CRY2	A	C	0.277	0.012181	0.0018	1.31E-11
rs62543565	9	16901067	BNC2	C	A	0.37	0.011185	0.001659	1.57E-11
rs147945015	18	47411162	MYO5B	C	T	0.0224	0.038807	0.005767	1.71E-11
rs1240808	6	16817368	RP1-151F17.2	A	G	0.344	0.011208	0.001673	2.08E-11
rs548824	9	14088789	NFIB	T	C	0.163	0.014529	0.002171	2.21E-11
rs10273527	7	139036598	LUC7L2	G	T	0.235	0.012748	0.001906	2.25E-11
rs9342045	6	86425148	RP11-33E24.3	C	T	0.451	0.010882	0.001638	3.07E-11
rs4838074	9	126467202	DENND1A	C	T	0.297	-0.01164	0.001754	3.22E-11
rs17577779	14	101200445	DLK1	T	C	0.478	0.011047	0.001665	3.28E-11
rs10842707	12	26471364	RP11-283G6.5	C	T	0.228	-0.01263	0.001906	3.47E-11
rs7586605	2	85561141	Y_RNA	C	T	0.268	0.012012	0.00182	4.07E-11
rs73702633	7	74288780	STAG3L2	G	A	0.0615	-0.02476	0.003752	4.17E-11
rs257377	7	106801088	PRKAR2B	G	T	0.198	0.013336	0.002022	4.27E-11
rs4788481	16	72936570	ZFHX3	T	G	0.288	-0.01213	0.00184	4.41E-11
rs61912867	12	20580220	PDE3A	G	A	0.127	-0.01603	0.002433	4.53E-11
rs35053761	2	227418737	MIR5702	C	A	0.0735	-0.02026	0.003082	4.96E-11
rs7732603	5	158614357	RNF145	A	C	0.278	0.011743	0.001788	5.17E-11

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs2911136	5	96177767	CTD-2260A17.2	C	G	0.367	0.010963	0.001671	5.34E-11
rs3767298	1	205037354	CNTN2	G	A	0.179	-0.0137	0.002091	5.67E-11
rs2254038	21	30745722	BACH1-IT2	C	T	0.411	-0.01085	0.001661	6.60E-11
rs6460894	7	12247330	TMEM106B	T	C	0.324	-0.01129	0.001734	7.43E-11
rs12347784	9	107689289	ABCA1	T	G	0.0983	-0.01775	0.002725	7.46E-11
rs12635093	3	152150885	MBNL1	C	T	0.244	0.012177	0.001871	7.61E-11
rs2294927	22	44382684	SAMM50	T	C	0.392	-0.01064	0.001642	9.21E-11
rs4984511	15	96704443	NR2F2-AS1	C	T	0.269	0.011597	0.00179	9.22E-11
rs9293683	5	75700310	IQGAP2	C	T	0.338	0.011225	0.001742	1.17E-10
rs2415142	15	73431600	NEO1	T	G	0.438	-0.0105	0.00163	1.20E-10
rs2290846	4	151199080	LRBA	G	A	0.271	-0.01155	0.001795	1.24E-10
rs254560	5	134443606	C5orf66	G	A	0.412	-0.01036	0.001612	1.30E-10
rs192154135	1	151279323	PI4KB	C	T	0.013	0.046842	0.00731	1.47E-10
rs11861132	16	28871860	SH2B1	A	G	0.403	-0.01054	0.001644	1.47E-10
rs17242395	19	11206969	LRG_274	G	A	0.113	0.016013	0.002499	1.48E-10
rs4855447	3	69803328	MITF	G	A	0.208	0.01261	0.001971	1.57E-10
rs5745110	5	172572954	BNIP1	C	T	0.125	0.015338	0.0024	1.66E-10
rs7732198	5	112379659	MCC	T	C	0.472	-0.01026	0.001614	2.08E-10
rs7686914	4	69537915	UGT2B15	T	C	0.471	0.01155	0.00183	2.79E-10
rs34796300	11	13315205	ARNTL	T	C	0.441	-0.01021	0.00162	2.96E-10
rs12270749	11	32469259	WT1-AS	T	A	0.312	-0.01083	0.001722	3.26E-10
rs1994721	11	29204531	RP11-466I1.1	G	A	0.149	0.01398	0.002226	3.40E-10
rs936018	17	42475962	GPATCH8	A	C	0.0274	-0.0324	0.005164	3.49E-10
rs12506484	4	38402147	RP11-83C7.1	C	T	0.0919	-0.01731	0.002761	3.61E-10
rs2502806	1	161135559	USP21	A	G	0.393	0.010505	0.00168	4.02E-10
rs10901852	10	126728522	CTBP2	G	C	0.261	-0.0114	0.001824	4.08E-10
rs75543218	16	57153441	CPNE2	C	G	0.0137	-0.04449	0.007127	4.30E-10
rs55841050	19	41750550	AXL	A	G	0.118	0.01548	0.002489	4.96E-10

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs62092076	18	40759172	RIT2	A	G	0.261	0.011307	0.001819	5.07E-10
rs12585801	13	81393264	ARF4P4	G	A	0.15	0.014069	0.002263	5.08E-10
rs61731455	5	665295	TPPP	A	G	0.185	0.01283	0.002067	5.40E-10
rs9867989	3	24291197	THRB	T	C	0.12	-0.01526	0.00246	5.60E-10
rs7848507	9	128593132	PBX3	G	T	0.365	-0.01053	0.001699	5.67E-10
rs2126606	12	89674799	RP11-13A1.3	T	C	0.379	-0.01058	0.001708	5.89E-10
rs546011672	8	19980352	AC100802.3	C	G	0.0191	-0.04297	0.006948	6.23E-10
rs11226111	11	103878499	PDGFD	T	C	0.152	-0.01405	0.002282	7.31E-10
rs6601924	10	5247302	AKR1C4	T	C	0.157	0.01368	0.002224	7.67E-10
rs1473580	3	114790058	ZBTB20	C	G	0.106	-0.01604	0.002608	7.81E-10
rs13066793	3	87037543	VGLL3	A	G	0.0913	0.017521	0.002861	9.07E-10
rs76166162	1	172380926	PIGC	G	A	0.0613	-0.0204	0.003332	9.27E-10
rs111731678	7	130418744	KLF14	T	A	0.196	-0.01345	0.002197	9.40E-10
rs854562	7	94947969	PON1	C	T	0.313	-0.01059	0.001731	9.45E-10
rs12934512	16	71965230	PKD1L3	C	G	0.268	0.011017	0.001804	1.02E-09
rs2325976	20	2837772	VPS16	C	T	0.364	0.010118	0.001657	1.02E-09
rs72629496	1	2953143	ACTRT2	G	A	0.166	-0.01327	0.002178	1.11E-09
rs13404250	2	188116196	AC007319.1	C	T	0.331	0.01042	0.001711	1.14E-09
rs4714902	6	46044237	CLIC5	C	A	0.343	-0.0109	0.001791	1.17E-09
rs144349177	5	180267079	LINC00847	A	C	0.00981	-0.05475	0.009008	1.22E-09
rs1955512	14	33175822	AKAP6	G	A	0.43	0.010137	0.001669	1.24E-09
rs4597680	3	57596105	RNU6ATAC26P	T	C	0.211	-0.01222	0.002015	1.33E-09
rs1007766	7	1034903	C7orf50	G	A	0.49	-0.0098	0.001617	1.34E-09
rs1292066	17	57932314	TUBD1	T	A	0.213	-0.01195	0.001971	1.35E-09
rs1128162	17	26721791	SLC46A1	C	A	0.446	-0.00975	0.001609	1.37E-09
rs34158262	11	58437534	GLYAT	C	T	0.0832	0.017844	0.002947	1.40E-09
rs1037117	15	102068658	PCSK6	G	A	0.252	0.01136	0.001876	1.41E-09
rs7656088	4	55499537	KIT	C	T	0.475	0.009826	0.001624	1.43E-09
rs3754749	2	173906149	RAPGEF4	A	G	0.17	0.01287	0.002129	1.49E-09

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs83921	9	33797056	RP11-133O22.6	A	T	0.384	-0.00999	0.001656	1.59E-09
rs17432483	2	63023836	EHBP1	T	C	0.0609	-0.0206	0.003414	1.61E-09
rs930457	4	143785362	INPP4B	T	C	0.352	-0.01026	0.001706	1.77E-09
rs62189714	2	162966841	AC008063.3	C	A	0.432	0.010086	0.001676	1.78E-09
rs800648	12	69966459	FRS2	C	T	0.362	-0.01003	0.001668	1.81E-09
rs13303368	1	914852	C1orf170	G	C	0.414	-0.01095	0.001822	1.85E-09
rs7658703	4	83778839	SEC31A	T	C	0.198	-0.012	0.001998	1.91E-09
rs116998829	8	19520669	CSGALNACT1	A	G	0.0361	0.026802	0.004466	1.95E-09
rs114317173	15	59479968	MYO1E	A	G	0.0101	0.053794	0.008995	2.22E-09
rs515939	11	57432629	ZDHHC5	G	A	0.271	-0.01071	0.001792	2.27E-09
rs6663226	1	66164748	LEPR	T	A	0.365	-0.00994	0.001664	2.28E-09
rs72729583	14	69149556	RAD51B	T	G	0.0718	0.019113	0.003213	2.71E-09
rs6772271	3	196193763	LRG_185	C	T	0.285	0.010658	0.001794	2.84E-09
rs664732	5	131281458	AC034228.4	G	C	0.0433	-0.02314	0.003906	3.11E-09
rs62544382	9	33109996	B4GALT1	G	T	0.431	0.009804	0.001657	3.28E-09
rs11624759	14	29746964	RP11-562L8.1	C	A	0.345	0.009972	0.001687	3.41E-09
rs6512256	19	18444360	RN7SL513P	T	G	0.217	0.01179	0.002	3.73E-09
rs7707821	5	52813375	FST	A	G	0.402	0.009746	0.001654	3.84E-09
rs17166	8	19969011	AC100802.3	C	T	0.0253	0.030998	0.005272	4.11E-09
rs11642733	16	30969061	SETD1A	G	A	0.107	-0.01515	0.002577	4.18E-09
rs1484917	2	146349854	AC092484.1	A	G	0.398	0.009624	0.00164	4.36E-09
rs17710668	17	3880546	RP11-459C13.1	C	G	0.192	-0.012	0.002047	4.59E-09
rs2364482	12	6502131	RP1-102E24.8	T	G	0.187	0.012082	0.002062	4.62E-09
rs11734408	4	23882519	PPARGC1A	A	G	0.298	0.010176	0.001738	4.79E-09
rs1949102	8	60654261	RP11-960H2.2	G	A	0.262	-0.01064	0.001818	4.87E-09
rs36137963	4	96150666	UNC5C	A	T	0.354	-0.00997	0.001706	5.09E-09
rs9537803	13	58366634	PCDH17	C	T	0.279	0.010451	0.001789	5.14E-09
rs9648555	7	46656869	HMGN1P19	A	G	0.454	0.009481	0.001626	5.45E-09

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs357409	7	138039643	AC008155.1	T	C	0.367	0.010368	0.001778	5.52E-09
rs7497973	15	23948571	NDN	G	A	0.487	0.009459	0.001623	5.56E-09
rs8130925	21	43725765	AP001623.1	G	T	0.496	0.009406	0.001617	6.01E-09
rs2928577	8	6603704	AGPAT5	G	A	0.387	0.009659	0.001662	6.12E-09
rs6992709	8	82414063	RP11-157I4.4	G	A	0.148	0.01324	0.002288	7.20E-09
rs1086246	9	13701224	RP11-536O18.1	G	A	0.33	0.009921	0.001718	7.65E-09
rs2531995	16	4013467	ADCY9	C	T	0.381	-0.00956	0.001656	7.87E-09
rs10958704	8	38328302	FGFR1	A	G	0.402	-0.00936	0.001623	7.94E-09
rs1988650	2	42689362	KCNG3	C	T	0.209	0.011528	0.002004	8.75E-09
rs73602128	20	566768	TCF15	T	C	0.0776	0.017647	0.003072	9.25E-09
rs13252421	8	19613947	CSGALNACT1	C	G	0.0129	-0.04087	0.007117	9.35E-09
rs335519	18	29790889	MEP1B	G	A	0.411	-0.00934	0.001626	9.44E-09
rs10778120	12	101875141	SPIC	A	T	0.336	0.009877	0.001724	1.01E-08
rs12023488	1	212498198	RP11-384C4.7	C	T	0.374	0.00946	0.001655	1.08E-08
rs3859027	16	88551465	ZFPM1	T	C	0.392	-0.00955	0.001671	1.12E-08
rs12155724	8	93205563	RP11-777J24.1	A	T	0.331	0.009848	0.001728	1.21E-08
rs11656686	17	74094664	EXOC7	C	T	0.426	0.009691	0.0017	1.21E-08
rs1149709	10	77038729	ZNF503-AS1	A	G	0.404	0.009313	0.001637	1.27E-08
rs12138866	1	224432260	NVL	A	G	0.039	-0.02407	0.004251	1.49E-08
rs6723650	2	30503722	LBH	A	C	0.455	-0.00913	0.001613	1.52E-08
rs9968060	3	62471282	CADPS	C	T	0.366	-0.0097	0.001719	1.67E-08
rs10911505	1	184049978	TSEN15	T	C	0.357	-0.00945	0.001676	1.70E-08
rs9787203	1	39288581	RRAGC	T	A	0.123	-0.01486	0.002639	1.79E-08
rs78377758	7	79775814	GNAI1	G	A	0.0861	0.016492	0.002941	2.05E-08
rs75130471	13	94085242	GPC6	G	A	0.246	-0.01051	0.001876	2.09E-08
rs2229741	21	16340289	NRIP1	C	T	0.422	-0.00901	0.001608	2.11E-08
rs1197674	15	42149952	SPTBN5	T	C	0.357	0.009501	0.001696	2.12E-08
rs59312889	19	5101405	KDM4B	C	G	0.0642	-0.01826	0.003263	2.21E-08

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs6944945	7	100795046	AP1S1	G	A	0.317	-0.00975	0.001743	2.25E-08
rs140391113	11	116434784	AP001891.1	T	C	0.00714	-0.05633	0.010081	2.30E-08
rs11161570	1	85721911	C1orf52	T	C	0.497	-0.00889	0.001593	2.40E-08
rs9447497	6	76219717	RPL26P20	T	C	0.364	0.009237	0.001662	2.71E-08
rs9637619	4	105845011	RP11-556I14.1	G	A	0.163	-0.01211	0.002178	2.73E-08
rs1149453	8	64532168	RN7SKP135	T	G	0.187	0.011431	0.002057	2.73E-08
rs11000759	10	75493726	GLUD1P3	G	A	0.423	-0.00924	0.001663	2.75E-08
rs4129968	13	99227770	STK24	A	G	0.356	-0.00967	0.001742	2.80E-08
rs2900177	9	123497645	AHCYP2	C	T	0.299	-0.00971	0.001751	2.95E-08
rs10733794	10	65355538	REEP3	A	G	0.328	0.009534	0.001723	3.15E-08
rs647016	1	42659669	FOXJ3	T	A	0.145	0.012498	0.002262	3.28E-08
rs117665650	9	82188040	TLE4	G	C	0.0978	0.014989	0.002714	3.34E-08
rs72934503	6	98583488	RP11-436D23.1	A	G	0.461	0.011174	0.002025	3.42E-08
rs17849502	1	183532580	NCF2	G	T	0.0458	-0.02105	0.003817	3.51E-08
rs12523657	5	173357406	CPEB4	C	G	0.173	0.011599	0.002105	3.57E-08
rs2080062	7	17697803	AC006482.1	T	C	0.101	-0.01495	0.002713	3.59E-08
rs2376180	4	37154148	RP11-103J17.1	T	A	0.489	0.008937	0.001624	3.74E-08
rs1714008	4	57945621	IGFBP7	G	A	0.414	0.008899	0.001618	3.81E-08
rs9956721	18	22621869	RP11-958F21.3	T	A	0.424	-0.00894	0.001629	4.03E-08
rs12028554	1	154223101	UBAP2L	G	A	0.437	-0.00892	0.001629	4.27E-08
rs980183	2	59311536	LINC01122	G	A	0.388	0.00894	0.001632	4.31E-08
rs891984	5	141465403	NDFIP1	T	A	0.366	0.00946	0.001729	4.43E-08
rs9608668	22	28545329	TTC28	T	C	0.26	-0.01045	0.001915	4.91E-08

Abbreviations. **HDL-C**: High-Density Lipoprotein Cholesterol; **SNP**: Single Nucleotide Polymorphism; **RSID**: Reference SNP Cluster ID; **GRCh37**: Genome Reference Consortium Human Build 37

Table S6. Triglycerides SNPs for European Populations used in MR analysis

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs624698	1	62900811	USP1	A	G	0.397	-0.0635	0.002002	1.00E-200
rs78980904	2	21121464	AC012361.1	T	C	0.236	-0.06338	0.001943	1.00E-200
rs4665378	2	27548038	MPV17	A	C	0.403	-0.06118	0.00165	1.00E-200
rs42129	7	72822848	Y_RNA	C	A	0.159	-0.07349	0.002295	1.00E-200
rs11986182	8	19727447	INTS10	A	T	0.162	-0.07127	0.002204	1.00E-200
rs2385114	8	126447308	TRIB1	C	T	0.493	0.052242	0.001693	1.00E-200
rs174533	11	61549025	TMEM258	G	A	0.347	0.052738	0.001693	1.00E-200
rs513533	11	116519358	AP000770.1	A	G	0.108	0.124243	0.00261	1.00E-200
rs61905078	11	116596309	BUD13	A	C	0.0745	0.202084	0.003046	1.00E-200
rs77697917	17	41840849	DUSP3	C	T	0.0275	0.156265	0.005071	1.00E-200
rs116843064	19	8429323	ANGPTL4	G	A	0.0201	-0.24476	0.005745	1.00E-200
rs72999033	19	19366632	AC138430.4	C	T	0.064	-0.1114	0.00331	1.00E-200
rs283810	19	45388241	CTB-129P6.4	T	G	0.066	0.110819	0.003347	1.00E-200
rs186369086	11	116607955	BUD13	C	A	0.00698	0.263578	0.010051	1.37E-151
rs4846913	1	230294715	GALNT2	C	A	0.402	-0.04222	0.001635	4.73E-147
rs150799387	8	19766707	LPL	G	A	0.0134	0.177191	0.007028	3.04E-140
rs55707100	15	43820717	MAP1A	C	T	0.0273	0.124697	0.00495	5.26E-140
rs6073958	20	44551855	PLTP	T	C	0.204	0.04525	0.001997	1.07E-113
rs41290120	19	45382675	PVRL2	G	A	0.0397	0.091629	0.004112	5.17E-110
rs6882076	5	156390297	TIMD4	T	C	0.365	0.035441	0.001666	2.23E-100
rs1077835	15	58723426	LIPC	A	G	0.215	0.041534	0.001957	5.97E-100
rs998584	6	43757896	VEGFA	C	A	0.478	0.035008	0.001676	7.38E-97
rs9461755	6	32441046	HLA-DRB9	G	A	0.0506	0.08134	0.004027	1.03E-90
rs1024137	2	227095876	AC068138.1	G	T	0.367	0.033501	0.001665	5.02E-90
rs13389219	2	165528876	COBLL1	C	T	0.403	-0.03206	0.001638	2.88E-85
rs3764261	16	56993324	AC012181.1	C	A	0.32	-0.03357	0.00172	7.46E-85

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs7924036	10	65191645	JMJD1C	G	T	0.483	-0.03058	0.00164	1.33E-77
rs6900444	6	31241182	HLA-C	C	T	0.483	0.031147	0.001687	4.40E-76
rs187429064	19	19380513	TM6SF2	A	G	0.0108	-0.15408	0.008431	1.30E-74
rs4418728	10	94839724	CYP26A1	G	T	0.453	-0.02943	0.001622	1.28E-73
rs77644716	2	27991368	MRPL33	G	C	0.0905	-0.05201	0.00287	2.16E-73
rs13108218	4	3443931	HGFAC	A	G	0.387	-0.03009	0.001715	6.64E-69
rs3843467	5	55856375	AC022431.2	G	T	0.19	0.035216	0.002033	3.14E-67
rs78058190	2	219699999	PRKAG3	G	A	0.0532	0.070969	0.004181	1.28E-64
rs672457	6	139836562	RP11-12A2.3	G	A	0.422	-0.02707	0.001624	2.22E-62
rs1532085	15	58683366	ALDH1A2	A	G	0.386	-0.02723	0.001641	7.37E-62
rs151450	4	88016514	AFF1	G	A	0.392	0.027402	0.001659	2.95E-61
rs12303671	12	124492610	FAM101A	T	G	0.337	-0.02666	0.001725	6.49E-54
rs76895963	12	4384844	CCND2	T	G	0.0167	-0.10196	0.006672	9.98E-53
rs12928099	16	15150505	PDXDC1	C	A	0.297	-0.02585	0.001774	4.59E-48
rs6602909	13	114551993	GAS6	T	C	0.332	0.025049	0.001767	1.32E-45
rs752446	8	10649767	PINX1	G	A	0.373	0.023787	0.001683	2.35E-45
rs10260148	7	130430969	KLF14	C	T	0.27	0.025479	0.001822	1.87E-44
rs41272086	6	161008646	LPA	G	A	0.0994	0.037482	0.002696	5.88E-44
rs577721086	6	127440047	RSPO3	T	C	0.0531	0.050632	0.003691	7.89E-43
rs112201728	6	160551486	SLC22A1	C	T	0.0684	0.042835	0.003126	9.98E-43
rs2908298	7	44192552	GCK	G	A	0.156	0.029666	0.002216	7.02E-41
rs34144542	11	116774447	SIK3	A	G	0.0685	-0.04386	0.003286	1.19E-40
rs901746	11	47260319	LRG_467	A	G	0.313	-0.02349	0.001765	2.02E-40
rs4804414	19	7223785	INSR	C	T	0.426	0.021505	0.001625	5.47E-40
rs117788606	7	72921771	BAZ1B	T	C	0.00839	-0.11878	0.009058	2.75E-39
rs2925979	16	81534790	CMIP	T	C	0.301	-0.02283	0.001749	6.11E-39
rs79395356	12	57738600	R3HDM2	T	G	0.25	-0.0247	0.00191	3.03E-38
rs4820325	22	38599978	MAFF	G	A	0.418	0.021318	0.001653	4.72E-38

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs148754782	8	19910576	AC100802.3	G	A	0.0313	-0.06013	0.004704	2.05E-37
rs9297994	8	59392324	CYP7A1	G	A	0.339	-0.02153	0.001699	7.98E-37
rs2540949	2	65284231	CEP68	A	T	0.387	-0.02077	0.001658	5.02E-36
rs62033400	16	53811788	FTO	A	G	0.404	0.020431	0.001636	9.01E-36
rs16971107	17	76394042	PGS1	G	A	0.385	-0.02053	0.001644	9.29E-36
rs13107325	4	103188709	SLC39A8	C	T	0.0595	0.040622	0.003319	1.91E-34
rs11656743	17	65980289	BPTF	G	A	0.216	-0.02406	0.001983	6.94E-34
rs2845885	11	63869062	RP11-21A7A.2	C	T	0.0574	-0.04236	0.003492	7.11E-34
rs1337101	1	219726100	RP11-95P13.2	G	T	0.311	-0.02131	0.001774	3.00E-33
rs4149056	12	21331549	SLCO1B1	T	C	0.162	0.0261	0.002174	3.23E-33
rs79287178	3	172294500	RP11-408H1.3	G	A	0.0285	0.058902	0.004994	4.11E-32
rs7786102	7	25965759	CTD-2227E11.1	G	A	0.286	-0.02118	0.001796	4.42E-32
rs684773	3	135956305	PCCB	A	C	0.198	0.02457	0.002125	6.21E-31
rs17184382	15	63792486	USP3	A	C	0.41	-0.01924	0.001668	8.90E-31
rs10755927	8	126629594	RP11-136O12.2	C	T	0.205	0.022731	0.001973	1.07E-30
rs874744	19	45513417	RELB	C	T	0.0231	0.067545	0.005889	1.88E-30
rs9817452	3	156795414	LINC00880	G	T	0.382	-0.01931	0.001685	2.05E-30
rs75679663	17	4667972	TM4SF5	C	A	0.00907	-0.09738	0.008504	2.31E-30
rs4804833	19	7970635	MAP2K7	A	G	0.395	-0.0189	0.001653	2.96E-30
rs2396083	6	43804808	RP11-344J7.2	G	C	0.305	0.020224	0.001777	5.34E-30
rs117664574	19	45427779	APOC1P1	C	G	0.0214	-0.06648	0.005845	5.66E-30
rs17036160	3	12329783	PPARG	C	T	0.125	-0.02668	0.002428	4.26E-28
rs2298058	13	95248566	TGDS	C	T	0.321	0.0191	0.00174	4.86E-28
rs114165349	1	27021913	ARID1A	G	C	0.0232	0.058748	0.005373	7.99E-28
rs72801474	5	132444128	HSPA4	G	A	0.0964	-0.03056	0.002818	2.13E-27
rs71352241	19	45443934	APOC4-APOC2	G	T	0.0291	-0.0547	0.00506	3.04E-27

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs62543565	9	16901067	BNC2	C	A	0.371	-0.018	0.001673	5.52E-27
rs2934	8	19692026	INTS10	A	C	0.468	-0.01749	0.00164	1.51E-26
rs2255400	10	113936244	GPAM	A	C	0.287	0.01876	0.001765	2.25E-26
rs2523608	6	31322559	HLA-B	G	A	0.424	-0.01756	0.001666	5.53E-26
rs9401908	6	127105492	RPS4XP9	A	G	0.445	-0.01724	0.001649	1.43E-25
rs6595187	5	118723609	TNFAIP8	A	G	0.283	-0.01868	0.001793	2.05E-25
rs16826087	1	39812518	MACF1	C	T	0.216	0.02037	0.001955	2.07E-25
rs4976033	5	67714246	CTC-537E7.3	A	G	0.408	0.017139	0.001667	8.45E-25
rs9604570	13	114631940	LINC00565	A	G	0.178	-0.0221	0.002153	1.03E-24
rs1800978	9	107665978	ABCA1	C	G	0.123	-0.02551	0.002488	1.14E-24
rs10832027	11	13357183	ARNTL	G	A	0.327	0.017187	0.001707	7.65E-24
rs79760705	5	53298716	ARL15	G	T	0.108	0.025907	0.00258	1.00E-23
rs2686189	8	11655229	FDFT1	C	T	0.324	-0.01737	0.001746	2.49E-23
rs2306363	11	65405600	SIPA1	G	T	0.198	-0.02016	0.002055	1.05E-22
rs78908794	11	116638854	BUD13	T	C	0.0195	0.055938	0.005736	1.82E-22
rs12953014	17	47453444	RP11-1079K10.2	A	G	0.38	0.016064	0.001681	1.23E-21
rs2963431	5	157970432	RP11-32D16.1	C	T	0.231	0.018209	0.001915	1.96E-21
rs12454712	18	60845884	BCL2	T	C	0.382	-0.01661	0.001753	2.61E-21
rs2642438	1	220970028	MARC1	A	G	0.293	-0.01683	0.001776	2.68E-21
rs1997833	20	39690342	TOP1	T	C	0.278	0.017212	0.001828	4.66E-21
rs1534696	7	26397239	SNX10	C	A	0.442	-0.01564	0.00167	7.64E-21
rs1652344	18	21128857	NPC1	T	C	0.492	-0.0151	0.001615	8.90E-21
rs12878001	14	64239629	RP11-804M7.2	T	G	0.162	0.02051	0.002195	9.28E-21
rs6499240	16	69686912	NFAT5	A	G	0.422	0.015601	0.00167	9.37E-21
rs6066138	20	45594711	EYA2	G	A	0.269	-0.01689	0.001812	1.14E-20
rs1056610	22	39080078	TOMM22	T	C	0.315	0.016239	0.001742	1.14E-20
rs1519480	11	27675712	BDNF-AS	C	T	0.319	-0.01579	0.001721	4.49E-20

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs7196161	16	31110981	VKORC1	G	A	0.381	0.015178	0.001664	7.29E-20
rs142122182	10	5263341	AKR1C4	A	G	0.155	-0.02068	0.002271	8.65E-20
rs28819812	4	157652753	PDGFC	C	A	0.327	-0.01552	0.001725	2.23E-19
rs9746779	16	987940	LMF1	A	G	0.371	-0.01497	0.001676	4.06E-19
rs115661421	2	26945012	KCNK3	A	G	0.00926	0.077647	0.00873	5.86E-19
rs6070138	20	56118558	PCK1	C	A	0.409	0.014799	0.001665	6.29E-19
rs7251733	19	56099951	FIZ1	G	A	0.156	0.019775	0.002226	6.57E-19
rs1292042	17	57934995	TUBD1	A	G	0.194	0.018076	0.002043	8.91E-19
rs11113118	12	107199142	RIC8B	G	A	0.228	0.016912	0.001916	1.08E-18
rs8126001	20	62711459	OPRL1	C	T	0.481	-0.01574	0.001783	1.11E-18
rs6977416	7	150542711	AOC1	G	A	0.337	0.015096	0.001711	1.12E-18
rs4253766	22	46623905	PPARA	C	T	0.105	0.02305	0.00262	1.41E-18
rs727428	17	7537792	SHBG	T	C	0.438	-0.01464	0.001672	1.99E-18
rs12952495	17	17687776	RAI1	C	T	0.392	-0.01474	0.001683	2.06E-18
rs6448429	4	26066863	RP11-324H7.1	C	T	0.163	0.019053	0.002181	2.45E-18
rs745578	8	72466324	RP11-1102P16.1	G	A	0.245	0.016398	0.001887	3.59E-18
rs74676173	1	149904474	MTMR11	A	C	0.0729	-0.02689	0.003095	3.70E-18
rs7898735	10	74712546	PLA2G12B	T	C	0.0848	0.025315	0.002921	4.45E-18
rs11046481	12	22747967	RP11-268P4.2	T	G	0.231	-0.01708	0.001987	8.23E-18
rs17052061	3	52346240	DNAH1	T	G	0.163	-0.01843	0.002177	2.56E-17
rs7485656	12	125315647	SCARB1	A	G	0.146	0.019375	0.002293	2.89E-17
rs79067298	7	73094267	DNAJC30	G	A	0.0381	-0.03576	0.004236	3.12E-17
rs2269869	19	19109545	SUGP2	G	A	0.243	-0.01647	0.001952	3.28E-17
rs35662434	4	69339933	TMPRSS11E	A	G	0.214	0.017303	0.002064	5.14E-17
rs7578604	2	121308660	AC073257.1	G	T	0.245	0.015793	0.001885	5.44E-17
rs57457691	19	33911705	PEPD	T	C	0.375	-0.014	0.001673	5.83E-17
rs34139543	16	72158477	PMFBP1	C	T	0.102	0.021988	0.002649	1.04E-16

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs117315666	11	116937074	SIK3	G	T	0.044	-0.03249	0.003935	1.50E-16
rs4938277	11	116397532	AP001891.1	C	T	0.279	0.015207	0.001844	1.62E-16
rs10958704	8	38328302	FGFR1	A	G	0.402	0.013488	0.001638	1.84E-16
rs1883783	1	54890956	RP4-705F19.1	T	G	0.414	0.013541	0.001652	2.48E-16
rs1326122	13	50722895	DLEU1	C	A	0.0238	-0.04454	0.005435	2.51E-16
rs2691583	7	17831806	SNX13	C	T	0.5	0.013332	0.001629	2.76E-16
rs11240358	1	205070573	RBBP5	G	A	0.399	0.013535	0.001654	2.78E-16
rs2129869	12	26457650	RP11-283G6.5	A	T	0.227	0.015658	0.001914	2.84E-16
rs35045014	6	107432157	BEND3	C	A	0.468	0.01398	0.001709	2.85E-16
rs6792725	3	24520283	THRB	A	G	0.325	-0.01485	0.001824	3.86E-16
rs60120444	3	170747667	SLC2A2	G	A	0.0971	0.022153	0.002721	3.86E-16
rs1126673	4	100045616	ADH4	C	T	0.308	0.014186	0.001745	4.35E-16
rs4717663	7	72071682	TYW1B	A	C	0.489	0.016441	0.002023	4.39E-16
rs9849171	3	12622623	MKRN2	G	C	0.358	-0.01358	0.001673	4.65E-16
rs7696969	4	143326714	INPP4B	T	G	0.362	-0.01359	0.001681	6.36E-16
rs4955426	3	49143438	QARS	C	A	0.22	0.015836	0.001962	6.92E-16
rs4768657	12	46086071	LINC00938	C	T	0.457	0.013011	0.001623	1.09E-15
rs17779355	10	103928374	NOLC1	G	A	0.0579	-0.02728	0.00342	1.48E-15
rs55875049	10	114012078	TECTB	A	G	0.185	0.016406	0.002059	1.64E-15
rs62397245	5	176750688	LMAN2	C	G	0.236	0.015261	0.001917	1.71E-15
rs78444298	1	184672098	EDEM3	G	A	0.0157	-0.05294	0.006653	1.76E-15
rs4916253	1	172361032	PIGC	G	T	0.428	-0.01295	0.001635	2.37E-15
rs154254	5	55820584	AC022431.2	G	C	0.419	0.012849	0.001626	2.70E-15
rs10846606	12	124620185	FAM101A	C	G	0.377	-0.0132	0.001676	3.31E-15
rs13112099	4	69537733	UGT2B15	T	G	0.471	0.014272	0.001813	3.51E-15
rs73029263	6	164113762	RP1-230L10.1	A	G	0.13	-0.01889	0.0024	3.52E-15
rs6029178	20	39178557	SNORD112	G	A	0.368	-0.01307	0.001663	3.82E-15
rs9969991	1	51058495	FAF1	T	A	0.426	-0.013	0.001656	4.08E-15

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs6142068	20	32557172	PIGPP3	T	C	0.389	0.013092	0.001674	5.20E-15
rs17036126	3	12287863	GSTM5P1	C	T	0.126	0.018793	0.002422	8.59E-15
rs71603401	4	18034463	LCORL	A	G	0.124	0.019063	0.002459	9.13E-15
rs12654983	5	131098220	FNIP1	G	A	0.0595	0.026836	0.003507	1.99E-14
rs10971957	9	34151463	RP11-537H15.4	C	T	0.404	-0.01274	0.001672	2.47E-14
rs6787781	3	155535351	SLC33A1	G	C	0.269	0.014092	0.001851	2.71E-14
rs12534129	7	117018135	ASZ1	T	C	0.399	0.012748	0.001694	5.17E-14
rs12960928	18	57897803	RP11-795H16.2	T	C	0.267	0.013788	0.001834	5.57E-14
rs117280021	17	64294581	PRKCA	T	C	0.00782	-0.07234	0.009658	6.90E-14
rs76306670	2	26970769	KCNK3	C	T	0.0312	-0.03548	0.004741	7.20E-14
rs1861410	2	58933591	LINC01122	C	T	0.451	-0.01215	0.00163	9.21E-14
rs7961572	12	111797013	FAM109A	A	G	0.248	-0.01418	0.001907	1.06E-13
rs9912287	17	1630992	RP11-961A15.1	G	A	0.218	0.014582	0.001984	1.97E-13
rs4683710	3	142661841	RP11-372E1.4	A	G	0.304	-0.0128	0.001744	2.16E-13
rs112144244	17	41920101	CD300LG	C	T	0.169	0.021551	0.002937	2.18E-13
rs10495702	2	20350751	AC098828.3	G	A	0.443	-0.0118	0.001614	2.69E-13
rs38205	7	15913588	AC006041.1	A	C	0.375	-0.01229	0.001691	3.71E-13
rs4818008	21	40611442	BRWD1	T	A	0.366	0.012422	0.001712	3.99E-13
rs165316	1	91533297	RPL5P6	A	G	0.201	-0.01452	0.002003	4.15E-13
rs12750321	1	154251626	LRG_64	A	G	0.289	-0.01306	0.001804	4.49E-13
rs72679249	4	111122162	ELOVL6	A	G	0.316	0.012997	0.0018	5.15E-13
rs12820906	12	123493123	PITPNM2	A	G	0.245	-0.0138	0.001912	5.30E-13
rs1133400	10	134459388	INPP5A	A	G	0.209	0.014463	0.002004	5.36E-13
rs3820897	2	3642361	LRG_350	T	C	0.198	0.014614	0.002034	6.75E-13
rs8029570	15	101905966	PCSK6	T	G	0.322	0.012401	0.001728	7.22E-13
rs275179	15	39447529	RP11-624L4.1	A	G	0.161	-0.0159	0.002218	7.64E-13
rs6060491	20	34175100	FER1L4	T	C	0.123	0.017547	0.00245	7.96E-13
rs9828193	3	66841702	RPL21P41	A	G	0.179	0.015253	0.002134	8.77E-13

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs77692382	1	40406715	Y_RNA	C	T	0.243	0.013469	0.001886	9.25E-13
rs4822458	22	24265659	AP000350.6	C	T	0.454	-0.0118	0.001653	9.41E-13
rs4770433	13	23903791	SACS	A	G	0.448	-0.01146	0.001606	9.66E-13
rs75772194	8	36846435	AC090453.1	A	G	0.166	0.0153	0.002151	1.15E-12
rs7608096	2	146373927	AC079163.1	T	A	0.438	-0.01154	0.001626	1.27E-12
rs72946301	2	111861838	AC096670.3	G	A	0.1	-0.01931	0.002722	1.30E-12
rs3850986	15	73091283	ADPGK-AS1	G	T	0.313	0.012344	0.001741	1.33E-12
rs147525635	15	43468698	TMEM62	G	A	0.447	0.012372	0.001751	1.62E-12
rs245479	5	57606869	PLK2	A	G	0.182	-0.01493	0.002118	1.78E-12
rs2063724	11	78133077	RP11-452H21.1	T	C	0.175	-0.01506	0.002152	2.66E-12
rs17713879	2	254215	SH3YL1	G	A	0.369	-0.01173	0.001679	2.74E-12
rs12084478	1	16467391	EPHA2	C	T	0.392	0.011852	0.001696	2.79E-12
rs3974807	11	49161967	RP11-107P7.4	C	T	0.184	0.01509	0.002161	2.89E-12
rs625476	9	13705973	RP11-536O18.1	C	T	0.318	-0.01202	0.001725	3.19E-12
rs4800116	18	19852666	snoU13	G	C	0.47	0.011267	0.001617	3.22E-12
rs12891399	14	104293533	PPP1R13B	T	C	0.339	0.011926	0.001714	3.44E-12
rs11835323	12	20463433	RP11-284H19.1	T	C	0.469	-0.01128	0.001621	3.55E-12
rs74328314	1	61670759	NFIA	A	G	0.0717	-0.02177	0.003156	5.28E-12
rs13280592	8	116686752	TRPS1	C	G	0.274	0.013036	0.001892	5.61E-12
rs111936426	9	107724707	RP11-217B7.3	G	A	0.085	0.019809	0.002878	5.91E-12
rs1907241	10	122897984	RP11-159H3.1	G	T	0.0258	0.037174	0.005402	5.94E-12
rs116994374	9	117084672	ORM1	G	A	0.0667	-0.02324	0.003383	6.45E-12
rs11242227	5	134135076	DDX46	A	G	0.112	0.017626	0.002576	7.79E-12
rs10215153	7	116399131	MET	G	A	0.308	0.01202	0.001757	7.92E-12
rs2613504	1	72835740	RPL31P12	G	A	0.186	0.014353	0.002106	9.36E-12
rs4352308	22	29131788	CHEK2	C	T	0.289	0.012199	0.001796	1.09E-11
rs35665085	22	17625915	CECR5	G	A	0.0556	0.024251	0.003571	1.12E-11

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs58301047	8	144296480	GPIHBP1	G	C	0.199	-0.01391	0.002051	1.16E-11
rs148566631	2	203447875	BMPR2	G	A	0.117	0.017035	0.002515	1.25E-11
rs10950655	7	17308367	AC003075.4	T	C	0.486	0.010963	0.001619	1.27E-11
rs10944134	6	86426468	RP11-33E24.3	A	C	0.36	-0.01284	0.001896	1.29E-11
rs10882089	10	94369060	KIF11	A	T	0.462	-0.01481	0.002204	1.84E-11
rs34162548	11	14556220	PSMA1	C	T	0.054	-0.02524	0.003758	1.86E-11
rs1491508	2	43754992	THADA	C	T	0.195	0.013714	0.002043	1.91E-11
rs74714416	15	57491156	TCF12	C	T	0.0404	0.028534	0.004259	2.09E-11
rs12154436	7	150343784	GIMAP6	G	A	0.194	0.01364	0.002037	2.15E-11
rs10763928	10	33632954	RP11-90B22.1	C	T	0.0889	0.019067	0.002852	2.29E-11
rs12601778	17	80191348	SLC16A3	C	T	0.43	0.01175	0.001761	2.49E-11
rs262695	1	2144788	RP11-181G12.4	A	G	0.312	0.012319	0.001846	2.50E-11
rs169080	19	4980864	KDM4B	T	C	0.362	-0.01131	0.001697	2.62E-11
rs10818471	9	123394045	MEGF9	A	G	0.298	0.01182	0.001774	2.69E-11
rs960230	21	42622479	BACE2	A	G	0.212	0.013249	0.001989	2.71E-11
rs7451008	6	20673880	CDKAL1	T	C	0.268	0.012035	0.001808	2.82E-11
rs6925800	6	161498311	MAP3K4	T	C	0.045	0.026683	0.004011	2.88E-11
rs2276808	3	36986992	TRANK1	C	T	0.455	-0.01085	0.001638	3.47E-11
rs36184164	6	43781092	VEGFA	T	G	0.109	0.017591	0.002656	3.52E-11
rs151235402	20	569164	TCF15	C	T	0.0169	0.04248	0.006448	4.46E-11
rs2182784	13	51133421	DLEU1	C	A	0.213	0.013	0.00198	5.18E-11
rs2302883	11	65824994	SF3B2	T	C	0.23	0.012759	0.001945	5.40E-11
rs57074291	2	227229344	AC068138.1	C	G	0.264	-0.0119	0.001832	8.30E-11
rs1377587	4	961191	DGKQ	A	G	0.453	-0.01088	0.00168	9.43E-11
rs35144691	9	92126375	SEMA4D	G	A	0.258	0.012054	0.001862	9.68E-11
rs12254046	10	135435981	FRG2B	C	T	0.0267	0.041296	0.006389	1.03E-10
rs114745683	1	226508335	YBX1P9	G	A	0.0422	0.025884	0.004011	1.10E-10

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs8054124	16	86434553	RP11-158I3.1	T	C	0.231	0.012365	0.00192	1.19E-10
rs16844848	2	161029592	ITGB6	A	G	0.195	0.013039	0.002033	1.41E-10
rs72815843	10	61462114	SLC16A9	G	A	0.198	-0.01307	0.002038	1.42E-10
rs3110452	17	28647752	TMIGD1	G	A	0.386	-0.01081	0.001685	1.43E-10
rs1571953	9	35121795	FAM214B	C	T	0.219	0.012417	0.001937	1.46E-10
rs74510325	22	50315382	CRELD2	C	G	0.0331	-0.02876	0.004492	1.53E-10
rs11075011	16	11828282	TXNDC11	T	A	0.415	0.011602	0.001813	1.57E-10
rs2808407	9	126295298	DENND1A	G	T	0.282	0.011528	0.001804	1.67E-10
rs1486000	8	82362153	RP11-157I4.4	A	G	0.161	-0.01421	0.00224	2.29E-10
rs148827772	4	110766064	RRH	A	G	0.0199	0.04096	0.006468	2.41E-10
rs547866	1	93725053	CCDC18	C	T	0.375	-0.01073	0.001694	2.43E-10
rs543928	6	32285248	C6orf10	G	A	0.375	0.011381	0.001799	2.54E-10
rs1759415	9	6674253	RP11-390F4.2	A	G	0.411	0.010293	0.001628	2.55E-10
rs62246310	3	9497884	SETD5	C	T	0.0989	0.018533	0.002938	2.84E-10
rs6924805	6	18747705	MIR548A1	G	T	0.406	-0.01039	0.001649	2.97E-10
rs72885907	11	30185850	RP4-562D20.2	G	A	0.367	-0.01065	0.00169	2.97E-10
rs34931250	17	66879927	ABCA8	C	T	0.0592	0.022052	0.003501	3.00E-10
rs12545916	8	106415737	ZFPM2	G	A	0.232	0.012061	0.001915	3.04E-10
rs1260752	6	109160802	ARMC2	T	C	0.189	-0.0133	0.002114	3.15E-10
rs61778886	1	23795181	ASAP3	G	C	0.13	0.017526	0.002788	3.26E-10
rs4245791	2	44074431	ABCG8	C	T	0.302	-0.01104	0.001759	3.43E-10
rs3828435	3	4743267	ITPR1	G	C	0.399	-0.01028	0.001637	3.43E-10
rs72683923	14	50735947	L2HGDH	T	C	0.0157	-0.04228	0.006753	3.83E-10
rs5011434	7	12268717	TMEM106B	C	T	0.408	0.010214	0.001632	3.90E-10
rs6773982	3	184894478	EHHADH-AS1	T	A	0.058	0.021302	0.003404	3.92E-10
rs1408579	10	101912194	ERLIN1	C	T	0.474	0.010032	0.001606	4.21E-10
rs115912456	5	82815158	VCAN	A	G	0.0427	-0.02489	0.003987	4.31E-10
rs34457392	10	63956395	RTKN2	A	G	0.101	-0.01664	0.002668	4.40E-10

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs849335	7	28223990	JAZF1-AS1	T	C	0.355	-0.01046	0.00168	4.72E-10
rs431420	19	54792079	AC010518.3	G	T	0.196	-0.01293	0.002077	4.81E-10
rs79639690	2	227605895	IRS1	T	C	0.0463	0.024247	0.003901	5.11E-10
rs34035326	10	120476990	CACUL1	A	G	0.33	0.010819	0.001741	5.14E-10
rs68165376	8	8820345	RP11-62H7.2	T	C	0.0756	0.019356	0.003117	5.28E-10
rs2235831	6	42883911	PTCRA	G	C	0.308	0.010841	0.00175	5.83E-10
rs4425336	4	89753225	FAM13A	A	G	0.223	-0.01192	0.001925	5.92E-10
rs139260949	17	63978436	CEP112	T	C	0.0551	-0.022	0.003553	5.92E-10
rs4916493	3	196139979	UBXN7	A	T	0.362	0.010372	0.001675	5.93E-10
rs2954011	8	126424379	TRIB1	C	T	0.46	-0.0107	0.001729	6.02E-10
rs13214458	6	140702211	MIR3668	G	T	0.253	-0.01154	0.001866	6.19E-10
rs4894128	2	180707310	ZNF385B	T	C	0.305	-0.01078	0.001744	6.40E-10
rs3744811	18	13649542	LDLRAD4	C	T	0.225	-0.01236	0.002002	6.66E-10
rs9924509	16	67573296	CTD-2012K14.2	T	C	0.0901	-0.01737	0.002814	6.71E-10
rs1061599	18	212756	USP14	A	C	0.272	-0.01115	0.001808	6.94E-10
rs72812817	5	173356748	CPEB4	A	G	0.316	0.010631	0.001724	7.03E-10
rs2857640	22	30056945	LRG_511	G	A	0.459	0.009963	0.001618	7.40E-10
rs118010373	11	14757444	PDE3B	A	C	0.0297	-0.03165	0.005149	7.92E-10
rs7312441	12	56941146	RBMS2	C	A	0.346	-0.01068	0.001743	8.95E-10
rs12962548	18	42599167	SETBP1	G	A	0.306	0.010596	0.001735	1.01E-09
rs560952	2	169704487	SPC25	T	C	0.33	-0.01033	0.001697	1.17E-09
rs13054650	22	50625079	TRABD	G	C	0.368	-0.01033	0.001705	1.38E-09
rs3120159	1	26238289	AL033528.1	A	G	0.406	0.010176	0.001681	1.42E-09
rs143076454	16	921179	LMF1	G	A	0.0163	0.041457	0.006859	1.50E-09
rs11666808	19	18383506	KIAA1683	T	C	0.362	-0.01022	0.001694	1.64E-09
rs1016565	9	1032567	DMRT2	G	A	0.312	0.010498	0.001748	1.91E-09
rs10118003	9	96918293	RP11-2B6.3	T	C	0.443	-0.00977	0.001628	1.94E-09

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs3825669	14	89804276	FOXN3	A	G	0.196	-0.01226	0.002043	1.97E-09
rs73729220	6	32342119	C6orf10	A	G	0.0149	0.040815	0.00681	2.06E-09
rs4391043	4	124835370	LINC01091	A	T	0.226	0.0117	0.001953	2.10E-09
rs7528419	1	109817192	CELSR2	A	G	0.219	-0.01154	0.001931	2.26E-09
rs603104	15	40757035	RP11-64K12.8	A	C	0.448	-0.00985	0.001647	2.27E-09
rs72969958	4	155364666	DCHS2	G	T	0.00872	0.055415	0.009287	2.42E-09
rs9431827	1	230440257	AL691479.1	C	T	0.33	-0.01045	0.001751	2.44E-09
rs13066793	3	87037543	VGLL3	A	G	0.0914	-0.01715	0.002883	2.68E-09
rs10052346	5	78472599	DMGDH	G	T	0.393	-0.01002	0.001685	2.80E-09
rs10076475	5	156422231	TIMD4	G	A	0.426	0.009717	0.001637	2.92E-09
rs11009319	10	33537248	NRP1	G	A	0.138	-0.01379	0.002324	2.93E-09
rs10784771	12	69621264	RP11-324P9.1	A	G	0.447	-0.00993	0.001673	3.00E-09
rs4671331	2	59312675	LINC01122	T	C	0.391	-0.00977	0.001647	3.03E-09
rs10985875	9	125813505	RABGAP1	T	G	0.112	-0.01793	0.003025	3.08E-09
rs2044755	3	24292926	THRB	T	A	0.122	0.014539	0.002458	3.30E-09
rs62130519	2	26564119	EPT1	C	T	0.0624	0.020294	0.003433	3.39E-09
rs7141420	14	79899454	NRXN3	C	T	0.476	0.009637	0.001638	4.02E-09
rs841229	16	4731155	MGRN1	G	A	0.468	-0.00976	0.001659	4.06E-09
rs144468328	20	57780137	ZNF831	G	A	0.0143	-0.04198	0.007146	4.22E-09
rs33554	5	112312736	DCP2	G	C	0.498	0.009528	0.001626	4.59E-09
rs3758199	9	139097645	QSOX2	C	T	0.5	-0.00982	0.001678	4.82E-09
rs78677597	6	100610044	RP3-344J20.1	A	C	0.23	-0.01123	0.001924	5.30E-09
rs11746307	5	141713797	AC005592.2	G	A	0.365	-0.00988	0.001695	5.51E-09
rs1035691	11	10658735	MRVI1	A	G	0.359	0.009927	0.001703	5.59E-09
rs11029441	11	26279145	ANO3	T	C	0.0872	0.017255	0.00296	5.59E-09
rs7007256	8	26203081	PPP2R2A	G	A	0.288	0.010318	0.001772	5.79E-09
rs12023563	1	210436110	ST13P19	G	A	0.298	-0.01031	0.001778	6.80E-09
rs1360485	13	31031884	HMGB1	C	T	0.307	0.010263	0.001776	7.48E-09

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs55673514	7	73021456	MLXIPL	A	G	0.0264	0.032502	0.005624	7.52E-09
rs113235777	13	114484903	TMEM255B	C	G	0.0136	-0.04285	0.007416	7.57E-09
rs62459120	7	44337834	CAMK2B	A	T	0.0241	-0.03116	0.005408	8.35E-09
rs113870241	20	30207379	MIR3193	C	T	0.0404	-0.0247	0.004287	8.38E-09
rs10030238	4	141808805	RNF150	A	G	0.406	-0.00944	0.00164	8.57E-09
rs17694094	19	33723705	SLC7A10	T	G	0.471	-0.00931	0.001619	8.98E-09
rs687914	2	45878760	PRKCE	G	T	0.238	0.011227	0.001953	9.02E-09
rs2947793	12	66199734	RPSAP52	G	A	0.0796	0.016901	0.002941	9.10E-09
rs12641981	4	45179883	RP11-362I1.1	C	T	0.428	0.00931	0.00162	9.11E-09
rs113268469	18	57862913	RP11-795H16.2	C	T	0.017	-0.03823	0.006653	9.11E-09
rs3806577	2	206546150	NRP2	T	C	0.461	-0.00945	0.00165	1.01E-08
rs1045279	3	141663683	TFDP2	C	T	0.155	0.012912	0.002254	1.02E-08
rs5750142	22	36036213	MB	G	A	0.497	0.009309	0.001629	1.11E-08
rs12069549	1	200036139	NR5A2	C	T	0.124	0.013844	0.002426	1.16E-08
rs2237650	7	106694366	PRKAR2B	A	C	0.461	0.009216	0.001618	1.22E-08
rs61900203	11	77759168	NDUFC2-KCTD14	A	G	0.149	0.014203	0.002494	1.23E-08
rs56184290	3	129328539	PLXND1	A	G	0.0996	-0.01532	0.002699	1.38E-08
rs35819975	21	40459532	RPL23AP12	G	C	0.328	0.009981	0.001761	1.44E-08
rs7200532	16	80756615	CDYL2	T	C	0.363	0.009637	0.001701	1.46E-08
rs6972751	7	149180500	ZNF746	T	C	0.304	-0.01032	0.001825	1.55E-08
rs4481184	3	185505787	IGF2BP2	C	T	0.309	0.009767	0.00173	1.64E-08
rs62252236	3	69893064	MITF	A	G	0.425	0.009358	0.001658	1.67E-08
rs12376215	9	112241370	PTPN3	T	A	0.28	-0.01019	0.00181	1.82E-08
rs6425364	1	171617708	MYOC	T	G	0.3	-0.00992	0.001763	1.85E-08
rs6998846	8	9193341	RP11-115J16.1	G	A	0.23	-0.01076	0.001919	2.03E-08
rs3768583	1	183538725	NCF2	T	A	0.473	0.008972	0.001603	2.20E-08
rs7068223	10	63681422	ARID5B	C	T	0.475	0.009305	0.001664	2.26E-08

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs7159513	14	29856951	RP11-260G13.1	G	A	0.172	0.012444	0.002227	2.32E-08
rs57314858	2	208031471	KLF7	C	T	0.4	0.009283	0.001662	2.35E-08
rs10791660	11	103871039	PDGFD	C	A	0.186	0.011827	0.002121	2.48E-08
rs4321388	2	239871065	AC114788.2	T	A	0.477	-0.00892	0.001603	2.63E-08
rs28510484	15	31637569	KLF13	G	C	0.164	0.01239	0.002229	2.72E-08
rs11664106	18	2846812	EMILIN2	A	T	0.36	-0.01055	0.001902	2.91E-08
rs1567518	19	57480893	CTC-258N23.3	A	G	0.469	0.009189	0.001658	2.96E-08
rs7662031	4	77205338	FAM47E-STBD1	T	C	0.202	0.011081	0.002004	3.22E-08
rs11123823	2	100916315	LONRF2	A	G	0.462	-0.00895	0.001618	3.24E-08
rs148636348	11	396294	PKP3	G	A	0.02	0.032378	0.005868	3.43E-08
rs2439649	9	111875463	TMEM245	G	A	0.465	-0.00903	0.001641	3.76E-08
rs11649885	17	76827880	USP36	A	G	0.477	-0.00904	0.001647	4.00E-08
rs62017272	14	40014604	RP11-111A21.1	T	C	0.199	0.011126	0.002027	4.04E-08
rs2237026	4	55539253	LRG_307	A	G	0.438	-0.00909	0.001658	4.17E-08
rs35603048	15	40391965	BMF	C	T	0.236	0.011067	0.00202	4.29E-08
rs2159359	17	49239036	NME2	C	A	0.182	-0.01139	0.002083	4.56E-08
rs680379	20	12969400	SPTLC3	A	G	0.392	-0.00894	0.001639	4.85E-08

Abbreviations. **SNP**: Single Nucleotide Polymorphism; **RSID**: Reference SNP Cluster ID; **GRCh37**: Genome Reference Consortium Human Build 37

Table S7. SBP SNPs for Chinese Populations used in MR analysis, adjusted for BMI

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs880315	1	10796866	CASZ1	T	C	0.373	0.9612	0.0962	1.50E-22
rs1078417	1	25327653	MIR4425	C	T	0.54	0.537	0.094	1.30E-08
rs12037987	1	113042822	WNT2B	T	C	0.696	0.8696	0.1013	4.90E-19
rs760077	1	155178782	MTX1	T	A	0.181	0.9872	0.1219	1.70E-16
rs1275923	2	26932796	KCNK3	T	C	0.758	1.1372	0.1096	1.80E-25
rs1260326	2	27730940	GCKR	C	T	0.525	0.5066	0.0936	4.40E-08
rs138938305	2	54924901	EML6	C	T	0.991	3.0559	0.5161	1.50E-09
rs142733694	2	86285211	POLR1A	T	TCTCTA	0.836	0.8705	0.1372	6.30E-11
rs10194493	2	165005726	AC092684.1	C	T	0.581	0.9279	0.0953	3.20E-22
rs6717108	2	225561035	DOCK10	C	T	0.471	0.5771	0.0964	3.00E-09
rs56137247	2	240195301	HDAC4	C	A	0.549	0.5289	0.0938	2.10E-08
rs67913572	3	14925876	FGD5	G	A	0.802	0.7505	0.1208	2.50E-09
rs9847595	3	27595519	Y_RNA	C	T	0.79	0.9385	0.115	4.40E-16
rs11713101	3	47338554	KLHL18	C	T	0.659	0.5682	0.0987	4.30E-09
rs3821843	3	53558012	CACNA1D	G	A	0.519	0.8857	0.0969	2.50E-20
rs12485510	3	53850209	CHDH	C	T	0.337	0.5882	0.1003	2.50E-09
rs11721038	3	168849576	MECOM	C	T	0.843	0.8855	0.1297	2.00E-11
rs7631972	3	169365354	MECOM	T	C	0.888	1.2136	0.1532	1.20E-15
rs61277036	4	54794343	FIP1L1	A	G	0.826	0.7066	0.125	4.50E-09
rs12509595	4	81182554	FGF5	T	C	0.6	1.7549	0.0955	4.10E-76
rs17010957	4	86719165	ARHGAP24	T	C	0.829	0.7749	0.1287	5.50E-10
rs11724647	4	156646340	GUCY1A3	A	T	0.807	0.7024	0.1183	1.70E-09
rs10059884	5	32832474	AC026703.1	C	A	0.343	0.7507	0.0994	9.30E-15
rs6886162	5	122482973	PRDM6	A	G	0.439	0.7987	0.0968	8.80E-17
rs7713256	5	127813874	FBN2	C	T	0.794	0.6391	0.1161	3.50E-08
rs10477399	5	148347543	SH3TC2	A	G	0.769	0.7806	0.1147	3.00E-12

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs2963479	5	157990431	RP11-32D16.1	C	G	0.711	0.5693	0.103	2.20E-08
rs707930	6	31741659	VWA7	C	T	0.573	0.6247	0.0954	3.20E-11
rs2254303	6	43276390	ZNF318	A	G	0.29	0.6376	0.1038	7.50E-10
rs78461345	6	126151100	NCOA7	A	G	0.978	1.8875	0.3374	3.40E-08
rs11154369	6	127100379	RPS4XP9	A	C	0.51	0.6099	0.0939	2.80E-11
rs17080091	6	150997401	PLEKHG1	T	C	0.936	1.2612	0.1919	1.30E-11
rs199987302	7	1076944	AC073957.15	G	A	0.734	0.7951	0.1061	5.80E-14
rs35427695	7	19041171	HDAC9	T	TA	0.623	0.58	0.0998	7.90E-09
rs5883064	7	27241878	HOTTIP	ACT	A	0.4	0.674	0.0962	1.20E-12
rs33979592	7	45959517	IGFBP3	TC	T	0.269	0.6873	0.1105	1.30E-09
rs11986274	8	38259481	LETM2	T	C	0.66	0.5968	0.0988	3.00E-09
rs10086846	8	143997131	CYP11B2	T	C	0.666	1.0137	0.1022	2.60E-23
rs1110061	9	127262802	NR5A1	C	G	0.676	0.6015	0.1023	9.20E-09
rs10900130	10	45423160	TMEM72	C	T	0.662	0.5993	0.0998	1.40E-09
rs7392557	10	60283524	BICC1	G	A	0.532	0.6076	0.0992	1.30E-09
rs7080472	10	96012950	PLCE1	T	G	0.433	1.0028	0.0953	1.40E-26
rs145231815	10	104639651	AS3MT	G	GTATA	0.726	1.6683	0.1051	4.60E-56
rs11405687	10	105718428	SLK	C	CA	0.918	0.9365	0.175	2.80E-08
rs6585252	10	115723678	NHLRC2	C	G	0.848	0.8975	0.1302	4.60E-12
rs1907228	10	122978944	LINC01153	G	A	0.486	0.5763	0.0942	2.80E-09
rs569550	11	1887068	AC051649.12	T	G	0.269	0.859	0.1066	2.40E-16
rs12800049	11	16248894	SOX6	C	T	0.732	0.6848	0.1071	2.30E-10
rs11230729	11	61278918	LRRC10B	T	C	0.626	0.8781	0.098	7.10E-20
rs4930295	11	65390554	PCNXL3	G	C	0.729	0.5857	0.106	2.10E-08
rs7928576	11	100598981	ARHGAP42	C	T	0.508	0.8222	0.0937	7.90E-19
rs5795058	11	116848929	SIK3	CA	C	0.828	0.8416	0.1241	5.90E-12
rs202077443	12	12845637	GPR19	GT	G	0.66	0.5785	0.0999	3.10E-09
rs4759320	12	54433011	HOXC4	C	G	0.591	0.5326	0.0996	4.20E-08

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs2681485	12	90025622	ATP2B1	G	A	0.377	1.0624	0.0966	1.70E-28
rs11329920	12	93519878	RP11-511B23.2	G	GC	0.677	0.6291	0.1026	8.10E-10
rs11065836	12	111671076	CUX2	A	G	0.574	0.5398	0.0975	1.70E-08
rs10774792	12	115552598	RP11-25E2.1	T	C	0.393	0.6363	0.0955	4.10E-12
rs780872	12	116196721	RP11-110L15.1	C	A	0.428	0.6011	0.102	4.00E-09
rs2071384	15	91430617	FES	A	G	0.213	0.7682	0.1185	1.50E-10
rs12596482	16	81503540	CMIP	T	C	0.638	0.5443	0.0987	1.50E-08
rs75268076	16	89714864	CHMP1A	A	T	0.494	0.5767	0.0959	1.30E-09
rs1991401	17	62502435	DDX5	G	A	0.631	0.5906	0.0984	4.00E-10
rs8102624	19	2161443	AP3D1	G	A	0.708	0.7418	0.1029	2.90E-13
rs167479	19	11526765	RGL3	T	G	0.518	0.7696	0.0932	4.00E-17
rs16981598	19	17204327	MYO9B	A	G	0.637	0.5442	0.0983	6.40E-09
rs144306528	20	10691127	RP11-103J8.1	A	AT	0.525	0.6139	0.0937	1.30E-11
rs6108787	20	10967214	RP11-103J8.1	T	G	0.463	0.577	0.0939	2.50E-10
rs6021247	20	50108980	NFATC2	G	A	0.529	0.7005	0.0936	1.10E-13
rs72630038	X	71354380	RP11-262D11.1	G	A	0.858	0.6663	0.112	8.60E-10

Abbreviations. **SBP**: Systolic Blood Pressure; **SNP**: Single Nucleotide Polymorphism; **RSID**: Reference SNP Cluster ID; **GRCh37**: Genome Reference Consortium Human Build 37

Table S8. SBP SNPs for Chinese Populations used in MR analysis, unadjusted for BMI

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs880315	1	10796866	CASZ1	T	C	0.373	1.0178	0.1006	4.90E-23
rs12037987	1	113042822	WNT2B	T	C	0.696	0.8373	0.1059	2.00E-16
rs760077	1	155178782	MTX1	T	A	0.181	1.0315	0.1274	1.70E-16
rs1275923	2	26932796	KCNK3	T	C	0.758	1.2162	0.1146	2.30E-26
rs117262811	2	54898273	SPTBN1	G	C	0.987	2.644	0.4603	7.50E-09
rs1186699	2	61662555	USP34	A	C	0.78	0.6488	0.118	2.50E-08
rs142733694	2	86285211	POLR1A	T	TCTCTA	0.836	0.8821	0.1434	3.30E-10
rs10194493	2	165005726	AC092684.1	C	T	0.581	0.864	0.0997	4.50E-18
rs62175706	2	174215709	AC092573.2	A	G	0.586	0.5861	0.1021	4.20E-09
rs6717108	2	225561035	DOCK10	C	T	0.471	0.6338	0.1008	3.60E-10
rs56137247	2	240195301	HDAC4	C	A	0.549	0.6154	0.0981	4.60E-10
rs9847595	3	27595519	Y_RNA	C	T	0.79	0.9557	0.1203	1.50E-15
rs11713101	3	47338554	KLHL18	C	T	0.659	0.6896	0.1032	8.50E-12
rs3821843	3	53558012	CACNA1D	G	A	0.519	0.9074	0.1013	9.70E-20
rs4687588	3	53850210	CHDH	C	G	0.667	0.6418	0.1051	5.60E-10
rs11721038	3	168849576	MECOM	C	T	0.843	0.935	0.1356	2.00E-11
rs6791024	3	169367152	MECOM	A	G	0.888	1.2442	0.1603	5.50E-15
rs61277036	4	54794343	FIP1L1	A	G	0.826	0.7187	0.1307	1.10E-08
rs12509595	4	81182554	FGF5	T	C	0.6	1.7265	0.0998	4.90E-67
rs17010957	4	86719165	ARHGAP24	T	C	0.829	0.7605	0.1346	6.20E-09
rs4691620	4	156338219	RP11-27G13.5	C	T	0.134	0.8051	0.1473	3.60E-08
rs11724647	4	156646340	GUCY1A3	A	T	0.807	0.7239	0.1237	2.40E-09
rs10059884	5	32832474	AC026703.1	C	A	0.343	0.7444	0.1039	1.80E-13
rs6886162	5	122482973	PRDM6	A	G	0.439	0.8035	0.1012	1.10E-15
rs10477399	5	148347543	SH3TC2	A	G	0.769	0.7662	0.1199	4.50E-11
rs4713467	6	31437275	HCP5	T	C	0.761	0.7091	0.1138	2.00E-10

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs112042348	6	43260011	SLC22A7	A	AC	0.708	0.7177	0.1088	6.60E-11
rs11154369	6	127100379	RPS4XP9	A	C	0.51	0.6539	0.0981	7.30E-12
rs17080091	6	150997401	PLEKHG1	T	C	0.936	1.2924	0.2007	2.00E-11
rs199987302	7	1076944	AC073957.15	G	A	0.734	0.7122	0.111	9.40E-11
rs5883064	7	27241878	HOTTIP	ACT	A	0.4	0.7099	0.1005	7.50E-13
rs2392929	7	106414069	RP5-884M6.1	T	G	0.865	0.7692	0.1428	4.70E-08
rs12544694	8	25898316	EBF2	T	C	0.632	0.5896	0.1018	4.70E-09
rs2411257	8	38293477	FGFR1	T	C	0.635	0.5792	0.103	2.40E-08
rs12543431	8	76782791	snoU13	T	C	0.545	0.6763	0.0989	2.10E-11
rs10086846	8	143997131	CYP11B2	T	C	0.666	1.0143	0.1069	3.20E-21
rs1110061	9	127262802	NR5A1	C	G	0.676	0.5982	0.107	3.40E-08
rs72240788	10	45418763	TMEM72	T	TTC	0.674	0.6384	0.1058	1.50E-09
rs7392557	10	60283524	BICC1	G	A	0.532	0.634	0.1037	1.60E-09
rs7080472	10	96012950	PLCE1	T	G	0.433	0.9796	0.0997	3.10E-23
rs145231815	10	104639651	AS3MT	G	GTATA	0.726	1.4357	0.11	4.60E-39
rs6585252	10	115723678	NHLRC2	C	G	0.848	0.9252	0.1362	1.50E-11
rs569550	11	1887068	AC051649.12	T	G	0.269	0.8267	0.1115	3.00E-14
rs12800049	11	16248894	SOX6	C	T	0.732	0.6876	0.112	2.00E-09
rs11230729	11	61278918	LRRC10B	T	C	0.626	0.8936	0.1025	1.70E-18
rs1355921	11	78139279	RP11-452H21.4	T	C	0.632	0.5746	0.1036	3.00E-08
rs6590805	11	100546036	AP000620.1	C	T	0.503	0.8042	0.099	1.90E-16
rs5795058	11	116848929	SIK3	CA	C	0.828	0.7548	0.1298	4.60E-09
rs66561220	12	12878765	APOLD1	C	T	0.553	0.6387	0.0988	1.80E-10
rs11105377	12	90088829	ATP2B1	C	T	0.374	1.089	0.1015	2.30E-27
rs11329920	12	93519878	RP11-511B23.2	G	GC	0.677	0.6228	0.1073	4.70E-09
rs77768175	12	112736118	HECTD4	G	A	0.79	0.7871	0.121	4.60E-11
rs35434	12	115554676	RP11-25E2.1	A	G	0.765	0.7568	0.1152	5.90E-12

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs374072727	15	53097337	RP11-209K10.2	C	CT	0.559	0.6023	0.1057	1.50E-08
rs4932374	15	91431255	FES	A	G	0.211	0.7751	0.1239	3.90E-10
rs11642015	16	53802494	FTO	C	T	0.875	1.0374	0.1484	3.00E-12
rs75268076	16	89714864	CHMP1A	A	T	0.494	0.6125	0.1003	5.50E-10
rs34879232	17	26722039	SLC46A1	GA	G	0.285	0.6085	0.1096	2.40E-08
rs17650301	17	62479273	POLG2	C	A	0.63	0.5518	0.1013	1.20E-08
rs8102624	19	2161443	AP3D1	G	A	0.708	0.763	0.1076	4.90E-13
rs167479	19	11526765	RGL3	T	G	0.518	0.8264	0.0974	8.20E-18
rs144306528	20	10691127	RP11-103J8.1	A	AT	0.525	0.6527	0.098	7.60E-12
rs6108787	20	10967214	RP11-103J8.1	T	G	0.463	0.6752	0.0981	1.20E-12
rs6021247	20	50108980	NFATC2	G	A	0.529	0.6957	0.0979	2.10E-12
rs5927879	X	31851923	LRG_199	C	T	0.324	0.4874	0.0874	3.00E-08
rs72630038	X	71354380	RP11-262D11.1	G	A	0.858	0.685	0.1172	2.10E-09

Abbreviations. **SBP**: Systolic Blood Pressure; **SNP**: Single Nucleotide Polymorphism; **RSID**: Reference SNP Cluster ID; **GRCh37**: Genome Reference Consortium Human Build 37

Table S9. SBP SNPs from Biobank Japan, used as comparison for CKB analyses

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs17035646	1	10796547	CASZ1	A	G	0.676279	-0.03277	0.003575	4.80E-20
rs3790604	1	113046879	WNT2B	A	C	0.295301	-0.02732	0.003691	1.40E-13
rs2586886	2	26932031	KCNK3	T	C	0.226052	0.041033	0.004062	5.40E-24
rs17046344	2	55023600	EML6	A	G	0.065779	-0.03884	0.006746	8.50E-09
rs73029563	2	165008166	AC092684.1	G	C	0.571312	-0.02536	0.003404	9.30E-14
rs4684242	3	14895105	FGD5	C	G	0.353439	-0.0237	0.003482	1.00E-11
rs35225531	3	27593064	Y_RNA	T	C	0.290471	-0.02821	0.004045	3.10E-12
rs9814480	3	53590465	CACNA1D	T	C	0.123892	-0.03873	0.005078	2.40E-14
rs10857147	4	81181072	FGF5	T	A	0.297126	-0.04264	0.003654	1.80E-31
rs13112725	4	106911742	NPNT	C	G	0.644787	-0.01967	0.003539	2.70E-08
rs10059884	5	32832474	AC026703.1	A	C	0.598781	-0.02236	0.003391	4.30E-11
rs2523559	6	31330969	DHFRP2	A	C	0.856226	0.027029	0.004754	1.30E-08
rs16896398	6	43262704	SLC22A7	T	A	0.346998	-0.02501	0.003525	1.30E-12
rs34370526	6	127171454	RPS4XP9	T	C	0.450462	-0.01911	0.003414	2.20E-08
rs200908540	7	1077117	AC073957.15	C	A	0.199134	-0.02448	0.004421	3.10E-08
rs60772526	7	27245996	HOTTIP	T	C	0.590113	-0.0227	0.00344	4.20E-11
rs67529638	8	76829893	snoU13	G	C	0.393283	0.018907	0.003411	3.00E-08
rs9692989	8	144042624	RP11-273G15.2	T	A	0.359305	0.023743	0.003477	8.50E-12
rs112913898	10	104958900	NT5C2	A	G	0.255801	0.034841	0.003871	2.30E-19
rs59448728	10	115710663	NHLRC2	A	T	0.212554	-0.02387	0.004162	9.70E-09
rs4980386	11	1895708	LSP1	A	C	0.759601	-0.02308	0.00394	4.70E-09
rs360140	11	9776567	RP11-540A21.2	A	C	0.665291	-0.0219	0.003526	5.20E-10
rs1847150	11	100579833	ARHGAP42	C	T	0.449433	-0.02225	0.003345	2.90E-11
rs1401982	12	89989599	ATP2B1	A	G	0.618249	-0.03725	0.003452	3.70E-27
rs671	12	112241766	ALDH2	A	G	0.247611	0.06159	0.003884	1.20E-56
rs35493	12	115513566	RP11-25E2.1	T	A	0.448561	0.021536	0.003347	1.20E-10
rs74971075	12	133099570	FBRSL1	A	G	0.108846	-0.03145	0.005671	2.90E-08

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs7166599	15	91403674	RN7SL363P	G	A	0.136642	-0.03076	0.005235	4.20E-09
rs8073894	17	59241155	BCAS3	C	T	0.741155	0.020994	0.003815	3.70E-08
rs6506469	18	715714	ENOSF1	C	T	0.466777	0.020949	0.003336	3.40E-10
rs12459507	19	2224387	DOT1L	C	A	0.269469	-0.02085	0.003753	2.80E-08
rs7360000	19	19266848	MEF2B	T	C	0.352887	-0.02155	0.003593	2.00E-09
rs6031431	20	42795152	LRG_394	G	A	0.320507	-0.02038	0.003623	1.80E-08

Abbreviations. **SBP**: Systolic Blood Pressure; **SNP**: Single Nucleotide Polymorphism; **RSID**: Reference SNP Cluster ID; **GRCh37**: Genome Reference Consortium Human Build 37

Table S10. SBP SNPs for European Populations used in MR analysis

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs7796	1	1752730	GNB1	C	G	0.4878	-0.3146	0.0297	3.68E-26
rs263533	1	2232129	SKI	C	T	0.4202	-0.1606	0.0246	6.71E-11
rs2493296	1	3410468	MEGF6	C	T	0.1437	0.3286	0.0348	4.12E-21
rs2232460	1	6599445	NOL9	G	A	0.335	-0.1788	0.0258	4.32E-12
rs488834	1	10707845	RP4-734G22.3	C	T	0.7573	-0.3328	0.0301	2.19E-28
rs2236796	1	11832558	C1orf167	G	C	0.1604	-0.7974	0.0327	4.29E-131
rs2861484	1	15486170	TMEM51	G	T	0.2012	-0.2296	0.0302	2.99E-14
rs1010069	1	16026442	PLEKHM2	G	A	0.4553	0.1625	0.0243	2.36E-11
rs12752228	1	25077759	CLIC4	C	T	0.1804	0.1799	0.0314	9.68E-09
rs114165349	1	26695422	ZNF683	G	C	0.0257	0.5759	0.0827	3.39E-12
rs74606343	1	27523564	CHCHD3P3	T	C	0.0877	-0.2852	0.0433	4.60E-11
rs143167197	1	28407861	EYA3	A	G	0.072	0.3931	0.0492	1.43E-15
rs1408945	1	41899206	RNA5SP45	G	T	0.4228	-0.2818	0.0243	4.16E-31
rs1209384	1	43299418	ERMAP	A	G	0.6177	-0.2248	0.0258	3.18E-18
rs688540	1	47536775	CYP4Z1	G	A	0.8667	0.2249	0.0365	6.96E-10
rs778124	1	56140534	RP11-466L17.1	G	A	0.3759	0.2513	0.0248	4.65E-24
rs61772592	1	56514009	RP1-158P9.1	A	G	0.1241	0.2693	0.037	3.29E-13
rs2092867	1	61411773	RP4-668G5.1	C	A	0.6472	0.1452	0.0251	7.40E-09
rs7547570	1	66542812	PDE4B	G	A	0.3818	0.1897	0.0249	2.30E-14
rs640935	1	78083962	ZZZ3	T	C	0.2396	0.1724	0.0283	1.13E-09
rs12035750	1	88866416	RN7SL583P	T	C	0.3713	0.2483	0.0252	6.88E-23
rs7514579	1	93585793	MTF2	A	C	0.2298	-0.1917	0.0289	3.07E-11
rs10776752	1	112501706	LRG_445	G	T	0.078	0.7251	0.046	7.28E-56
rs57748895	1	115283548	CSDE1	A	T	0.0186	0.7931	0.0921	7.00E-18
rs10159340	1	116292367	CASQ2	C	T	0.7313	-0.1531	0.0273	1.96E-08
rs11585169	1	150599561	ENSA	T	A	0.5706	0.141	0.0246	9.66E-09
rs35479618	1	153689947	RP11-216N14.1	G	A	0.0168	0.9244	0.1153	1.08E-15

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs2279470	1	164808130	PBX1	C	G	0.2155	-0.1719	0.0295	5.55E-09
rs12138150	1	169129500	NME7	C	T	0.3938	-0.2076	0.0247	4.50E-17
rs1043069	1	180890232	KIAA1614	T	G	0.382	-0.1817	0.0252	5.56E-13
rs4651224	1	184616048	C1orf21	C	T	0.4496	0.1914	0.0247	9.25E-15
rs817140	1	193302396	LINC01031	T	C	0.2758	0.1608	0.027	2.52E-09
rs12042924	1	197328287	CRB1	T	C	0.4665	0.1658	0.0245	1.19E-11
rs2075865	1	207065831	IL24	C	A	0.4086	-0.1489	0.0246	1.47E-09
rs10863593	1	207844729	CR1L	T	C	0.3231	-0.1891	0.0257	2.09E-13
rs12134698	1	209814456	LAMB3	C	A	0.1948	-0.2091	0.0302	4.56E-12
rs924653	1	217162129	ESRRG	G	A	0.5208	-0.1551	0.024	1.05E-10
rs11117882	1	217563075	GPATCH2	A	G	0.2356	0.2333	0.0285	2.44E-16
rs12088448	1	218372828	RNU1-141P	A	C	0.3527	0.1767	0.0256	5.65E-12
rs2494196	1	219589239	RP11-392O17.1	C	A	0.2821	-0.1673	0.0268	4.28E-10
rs1745417	1	228016578	PRSS38	C	T	0.5196	0.1834	0.0243	4.76E-14
rs699	1	230710048	RP11-543E8.1	A	G	0.4117	0.327	0.0244	8.01E-41
rs4926499	1	248861710	XX-CR54.1	G	C	0.8245	0.2955	0.0368	9.46E-16
rs6723772	2	12854566	RP11-333O1.1	C	T	0.1048	-0.2272	0.0402	1.55E-08
rs55881012	2	19519299	AC092594.1	A	C	0.4993	0.1939	0.0241	7.89E-16
rs11125853	2	24939798	NCOA1	A	G	0.2323	-0.2733	0.0287	1.50E-21
rs1275980	2	26694101	OTOF	C	T	0.6114	-0.5164	0.0246	3.31E-98
rs4666076	2	28429662	BRE	G	A	0.3455	-0.1522	0.0252	1.45E-09
rs6734118	2	37332212	EIF2AK2	C	A	0.2167	-0.3217	0.0347	1.87E-20
rs12471513	2	37765767	AC006369.2	C	T	0.3247	0.1626	0.0271	2.01E-09
rs2160236	2	40330136	SLC8A1	G	C	0.3846	-0.231	0.0258	2.94E-19
rs76326501	2	42940738	MTA3	A	C	0.0889	-0.5594	0.0432	2.40E-38
rs2115859	2	43158953	AC016735.1	T	C	0.6377	-0.214	0.0252	2.32E-17
rs6745430	2	55594133	CCDC88A	C	A	0.1761	-0.2002	0.0316	2.49E-10
rs2868985	2	55964335	PNPT1	G	A	0.8038	-0.1712	0.0303	1.67E-08
rs10188388	2	59783328	AC007131.2	C	T	0.4123	-0.198	0.0248	1.42E-15

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs10198613	2	64072136	UGP2	T	C	0.3527	0.1478	0.0259	1.17E-08
rs2540949	2	65057097	Y_RNA	A	T	0.3805	-0.2606	0.0248	9.71E-26
rs4433986	2	66533179	AC092669.6	A	G	0.6046	-0.1883	0.0246	1.97E-14
rs11676702	2	68183455	AC079112.1	T	C	0.0813	0.3051	0.045	1.14E-11
rs7578166	2	71402911	AC007881.4	A	C	0.617	-0.1404	0.0248	1.51E-08
rs10171459	2	73207600	SFXN5	A	C	0.4746	0.1337	0.024	2.57E-08
rs2003931	2	86229713	POLR1A	G	A	0.3493	-0.2124	0.0254	5.68E-17
rs17030997	2	105503421	RP11-13J10.1	A	T	0.0913	-0.2335	0.0423	3.34E-08
rs3811640	2	112019402	MIR4435-1HG	C	A	0.2783	-0.2005	0.0268	8.05E-14
rs62158170	2	113324598	POLR1B	A	G	0.2167	-0.1933	0.0295	6.11E-11
rs58132852	2	120680873	PTPN4	C	T	0.079	-0.2708	0.047	8.22E-09
rs6541771	2	121218943	LINC01101	G	A	0.4364	-0.1672	0.025	2.18E-11
rs11123059	2	124671429	RN7SKP102	G	A	0.5683	0.1342	0.0243	3.47E-08
rs72844527	2	137575819	THSD7B	C	A	0.1652	0.2051	0.0326	2.95E-10
rs7575637	2	145038684	GTDC1	G	A	0.4581	-0.2211	0.0241	5.13E-20
rs79909779	2	146270182	AC092484.1	G	A	0.1382	0.2228	0.0361	6.93E-10
rs79523138	2	160511702	AC009961.3	A	G	0.1162	0.2713	0.039	3.30E-12
rs4667454	2	164011216	RPL7P61	A	G	0.3257	-0.2178	0.0259	3.84E-17
rs73029563	2	164151656	RNU6-627P	C	G	0.5381	0.4536	0.0244	2.72E-77
rs16849728	2	164763795	AC092684.1	A	G	0.0604	-0.3474	0.0516	1.71E-11
rs2362866	2	176673538	EXTL2P1	G	A	0.6626	0.1495	0.0256	4.97E-09
rs17400834	2	177762179	RNU6-187P	T	C	0.1428	0.1905	0.0347	3.83E-08
rs55875326	2	179875463	CCDC141	G	C	0.2673	-0.208	0.0278	6.61E-14
rs6732308	2	182131461	AC104820.2	A	G	0.2028	-0.2877	0.0302	1.81E-21
rs2370735	2	187015349	AC104058.1	A	G	0.186	0.1762	0.031	1.26E-08
rs1894774	2	190726820	PMS1	A	G	0.5092	-0.2219	0.0242	5.08E-20
rs17474056	2	203122831	NOP58	C	T	0.5627	-0.2152	0.0243	9.23E-19
rs3845811	2	207656788	FASTKD2	C	G	0.4295	0.2496	0.0247	6.05E-24
rs1250259	2	215435759	VWC2L	T	A	0.7383	-0.3031	0.0326	1.32E-20

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs1863798	2	217906060	RP11-574O16.1	A	G	0.3791	0.2073	0.0251	1.36E-16
rs10171839	2	218186591	DIRC3	G	A	0.4931	-0.1469	0.0248	2.98E-09
rs2943646	2	226234818	AC067961.1	A	G	0.6433	0.2709	0.025	1.94E-27
rs1044822	2	229764422	PID1	C	T	0.1472	-0.2291	0.0343	2.29E-11
rs6543523	2	239306568	TRAF3IP1	A	G	0.2547	0.1647	0.028	4.25E-09
rs13001766	2	241406119	GPC1	T	C	0.2567	0.1655	0.0291	1.26E-08
rs79349366	2	241526862	CAPN10	C	T	0.0326	-0.5136	0.0714	6.21E-13
rs7632040	3	11385981	ATG7	A	G	0.6031	0.2942	0.0247	9.71E-33
rs1623112	3	14851365	RP11-95M5.1	T	C	0.6372	-0.2635	0.025	5.23E-26
rs2623079	3	20069953	KAT2B	C	G	0.381	-0.1961	0.0248	2.57E-15
rs2643826	3	27521497	SLC4A7	C	T	0.4522	0.4177	0.0247	5.54E-64
rs7624539	3	30252315	AC137674.2	C	T	0.1591	0.2549	0.0362	1.91E-12
rs7427249	3	37530998	ITGA9	G	A	0.579	-0.1998	0.0244	2.62E-16
rs73075656	3	41064456	RP11-520A21.1	G	T	0.1339	-0.2552	0.0357	8.78E-13
rs6442094	3	48038087	MAP4	C	G	0.679	0.2846	0.0262	1.96E-27
rs3821843	3	53523985	CACNA1D	G	A	0.6828	0.3071	0.027	6.29E-30
rs1599230	3	53615459	CACNA1D	A	G	0.4905	0.1446	0.0244	2.97E-09
rs6445819	3	56719634	FAM208A	C	T	0.3127	-0.2479	0.0261	1.89E-21
rs7618284	3	66371822	SLC25A26	G	C	0.3377	-0.1732	0.0266	7.66E-11
rs1852922	3	70887561	RNU6-281P	G	A	0.6845	0.1911	0.0262	3.31E-13
rs12494446	3	85573513	CADM2	G	A	0.6288	0.2166	0.0249	3.15E-18
rs13090602	3	115399319	GAP43	T	A	0.4986	-0.1697	0.0242	2.37E-12
rs9842387	3	121898148	CASR	A	G	0.4633	0.1329	0.0242	4.15E-08
rs6438857	3	124838796	SLC12A8	T	C	0.4252	-0.2145	0.0245	2.29E-18
rs4854572	3	133082395	TMEM108	G	A	0.4557	-0.1664	0.0245	1.04E-11
rs9809556	3	134231130	CEP63	G	A	0.3925	0.2923	0.0247	2.21E-32
rs34905952	3	138385793	PIK3CB	G	A	0.16	0.2401	0.0329	2.78E-13
rs62271373	3	150348753	SELT	T	A	0.0592	0.4389	0.0539	4.06E-16
rs447429	3	154192157	RP11-656A15.1	T	C	0.8606	-0.2355	0.0348	1.35E-11

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs9833313	3	157859002	RSRC1	A	T	0.757	0.1579	0.0285	3.06E-08
rs73156495	3	158702758	IQCJ	C	T	0.1646	0.2083	0.0324	1.29E-10
rs74516251	3	169112989	MECOM	G	A	0.079	-0.4591	0.0456	7.63E-24
rs988398	3	169381073	MECOM	C	T	0.4424	0.3037	0.0242	3.28E-36
rs1421002	3	169611888	KRT18P43	T	C	0.4736	0.1548	0.0243	2.00E-10
rs12638862	3	169759718	GPR160	A	G	0.263	-0.2807	0.0275	1.73E-24
rs263017	3	183785229	HTR3C	A	G	0.508	-0.1907	0.0241	2.60E-15
rs6778977	3	185584744	IGF2BP2	T	G	0.3421	0.1797	0.0256	2.25E-12
rs67822851	3	196441176	PIGX	A	G	0.39	-0.1434	0.0252	1.34E-08
rs6777317	3	197344088	AC024560.3	G	A	0.2955	0.1714	0.0268	1.71E-10
rs61789367	4	2250166	MXD4	G	A	0.0451	0.3355	0.0599	2.10E-08
rs1290933	4	2666490	FAM193A	C	A	0.6902	-0.237	0.0261	1.20E-19
rs3752440	4	3445027	HGFAC	G	A	0.0843	0.3282	0.0455	5.67E-13
rs2315518	4	14900544	LINC00504	G	C	0.3377	0.1467	0.0253	6.99E-09
rs2011603	4	18023861	LCORL	G	A	0.7365	0.2464	0.0275	3.14E-19
rs35107212	4	26797768	TBC1D19	A	G	0.4021	0.2168	0.0248	2.53E-18
rs2291433	4	38385454	RP11-83C7.1	T	G	0.5348	-0.2327	0.0242	7.13E-22
rs12511987	4	46593606	RNU6-412P	T	G	0.178	0.207	0.0321	1.08E-10
rs955683	4	48718177	FRYL	T	C	0.5046	0.1793	0.0257	2.91E-12
rs871606	4	53933078	SCFD2	T	C	0.1063	-0.3572	0.0401	5.56E-19
rs10015278	4	62930432	LPHN3	T	C	0.3634	-0.1474	0.025	3.60E-09
rs4859681	4	76478011	THAP6	C	A	0.5296	-0.2064	0.0241	1.02E-17
rs12509595	4	80261400	NAA11	T	C	0.2915	0.7041	0.0266	5.97E-155
rs60909079	4	82909091	RP11-689K5.3	G	C	0.2492	-0.1998	0.0279	8.02E-13
rs72976750	4	85804531	WDFY3	T	C	0.1386	0.4688	0.0352	1.41E-40
rs10028284	4	88831762	HSP90AB3P	A	T	0.1803	-0.2553	0.0319	1.09E-15
rs10155222	4	95164035	SMARCAD1	A	T	0.5505	-0.1559	0.0241	1.03E-10
rs13107325	4	102267552	PPP3CA	C	T	0.0764	-0.7456	0.0461	6.00E-59
rs7677290	4	105988160	RP11-556I14.1	G	C	0.3902	0.166	0.0248	2.08E-11

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs9992763	4	108137562	DKK2	G	T	0.5692	-0.1728	0.0244	1.35E-12
rs13118687	4	110485340	CCDC109B	G	A	0.4636	-0.2634	0.0244	4.13E-27
rs112835657	4	119660826	SEC24D	G	A	0.0332	-0.3891	0.0688	1.54E-08
rs12639622	4	123891299	SPATA5	T	C	0.1853	-0.1918	0.0316	1.34E-09
rs9330353	4	137518476	AC093875.1	T	A	0.4227	0.2414	0.0247	1.55E-22
rs72719160	4	143130123	INPP4B	A	T	0.3199	0.2288	0.0258	8.54E-19
rs6820938	4	147474751	RNU1-44P	A	G	0.1855	0.1801	0.0314	1.03E-08
rs4691666	4	155477536	RP11-158C21.3	C	T	0.5325	-0.2843	0.0243	1.05E-31
rs3796584	4	155719895	RBM46	G	A	0.3623	-0.2753	0.0249	2.28E-28
rs17035181	4	156757359	ASIC5	T	G	0.1456	-0.2949	0.0342	7.22E-18
rs869396	4	168766849	RP11-310I9.1	C	A	0.4696	-0.197	0.0245	8.44E-16
rs3756712	5	308981	AHRR	A	C	0.6241	-0.1508	0.0252	2.16E-09
rs10069690	5	1279675	LRG_343	C	T	0.2585	0.2762	0.0286	4.43E-22
rs4702102	5	15698810	FBXL7	A	T	0.6983	-0.1805	0.0329	4.02E-08
rs7733331	5	32828740	AC026703.1	T	C	0.5996	0.5493	0.0245	4.23E-111
rs10941043	5	33194645	CTD-2066L21.3	T	G	0.2879	0.192	0.0265	4.37E-13
rs10072115	5	38758999	CTD-2127H9.1	A	C	0.648	0.1378	0.0251	4.01E-08
rs10039350	5	51441463	CTD-2203A3.1	T	C	0.3848	-0.1777	0.025	1.21E-12
rs1694068	5	53987800	AC112198.1	T	A	0.6161	0.2463	0.0248	3.28E-23
rs10054208	5	56393165	RP11-101B14.4	C	T	0.3578	0.1471	0.0265	2.72E-08
rs464605	5	56511543	GPBP1	C	T	0.7447	0.2119	0.0276	1.49E-14
rs13156073	5	62395530	RP11-586E1.1	A	G	0.0751	0.3597	0.0458	4.33E-15
rs6414805	5	67026772	RP11-83M16.6	T	C	0.2573	-0.1779	0.0274	8.48E-11
rs10066799	5	68412563	CTC-498J12.3	G	T	0.2808	0.1841	0.0269	7.96E-12
rs7703932	5	68748704	RN7SL476P	C	T	0.4776	0.1701	0.0241	1.62E-12
rs10076149	5	78579031	JMY	C	G	0.4645	-0.2198	0.0241	7.93E-20
rs258951	5	92718160	NR2F1-AS1	T	A	0.5491	0.153	0.0245	4.19E-10
rs709668	5	96838483	RP11-1E3.1	A	G	0.7968	0.2584	0.0302	1.08E-17
rs1871190	5	98618015	RP11-93O17.1	G	T	0.3327	0.1444	0.0262	3.68E-08

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs41508849	5	108872579	RP11-231G15.1	T	C	0.2137	-0.1745	0.0296	3.58E-09
rs9122	5	113022434	YTHDC2	T	C	0.507	-0.1508	0.0241	3.69E-10
rs11241305	5	115048720	CTC-505O3.3	C	A	0.4894	0.2515	0.0241	1.58E-25
rs6890251	5	123150725	KRT18P16	C	T	0.1579	0.3641	0.0332	5.52E-28
rs7703751	5	123496287	HMGB1P29	A	T	0.2554	-0.2518	0.0277	1.13E-19
rs17677603	5	128521800	ISOC1	A	G	0.383	0.3128	0.0248	2.36E-36
rs28421137	5	139657545	PFDN1	A	G	0.3776	-0.1447	0.0253	1.03E-08
rs1962648	5	140650721	PCDHB15	C	T	0.4198	0.1859	0.0244	2.51E-14
rs111304635	5	142392423	ARHGAP26	C	T	0.2494	-0.1826	0.0285	1.48E-10
rs79376802	5	149105397	RN7SL868P	T	C	0.0518	-0.3901	0.0555	2.17E-12
rs17609994	5	158045536	EBF1	A	G	0.2002	0.3069	0.0311	5.78E-23
rs12716337	5	158387220	EBF1	T	C	0.3806	-0.4027	0.0251	5.79E-58
rs10052777	5	158842073	AC008703.1	T	C	0.3961	0.2841	0.0248	2.73E-30
rs67909753	5	173879055	RP11-267A15.1	G	A	0.2894	-0.2251	0.0266	2.44E-17
rs12153395	5	179984477	CNOT6	G	A	0.1097	-0.288	0.0398	4.79E-13
rs2569882	6	1619912	GMDS	T	C	0.4349	-0.1982	0.0251	2.66E-15
rs6921610	6	6996928	RP11-556O15.1	C	T	0.5371	-0.1616	0.0242	2.23E-11
rs1630736	6	12295754	EDN1	C	T	0.4622	-0.1664	0.0246	1.39E-11
rs9368222	6	20686765	CDKAL1	C	A	0.2676	0.2276	0.0271	4.12E-17
rs885769	6	22060353	CASC15	T	C	0.6378	0.1637	0.0252	8.16E-11
rs2744133	6	22392031	RP3-404K8.2	A	G	0.2742	-0.2621	0.0271	3.65E-22
rs79782817	6	25882450	HIST1H2APS2	G	T	0.1034	0.4271	0.0396	3.55E-27
rs17189980	6	31018263	HCG22	T	C	0.0371	0.4136	0.0663	4.37E-10
rs2753960	6	31795067	HSPA1B	G	T	0.424	0.3651	0.0245	2.04E-50
rs116793716	6	31924533	NELFE	T	G	0.0169	-0.5871	0.0972	1.55E-09
rs116839142	6	34290794	RPS10-NUDT3	T	G	0.029	0.4261	0.0778	4.40E-08
rs2815063	6	39294759	KIF6	C	A	0.1314	0.232	0.0367	2.75E-10
rs1214761	6	43386693	ABCC10	A	G	0.6655	-0.3278	0.0257	2.18E-37
rs6905288	6	43791136	RP11-344J7.2	G	A	0.5671	0.1641	0.0246	2.50E-11

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs61559025	6	50983277	FTH1P5	T	G	0.0626	0.3034	0.0515	3.94E-09
rs1474698	6	56199399	COL21A1	C	T	0.5503	-0.1934	0.0243	1.59E-15
rs12195276	6	72947991	RIMS1	C	T	0.723	-0.1601	0.0273	4.45E-09
rs1984195	6	78947674	SNORD112	G	A	0.4883	0.235	0.0244	5.68E-22
rs9344215	6	81668946	RP11-486E2.1	C	T	0.2353	0.1635	0.0294	2.60E-08
rs6921291	6	96618366	FUT9	C	T	0.1906	0.3343	0.0316	3.99E-26
rs9486916	6	108692727	LACE1	C	T	0.2007	0.2064	0.0312	3.84E-11
rs6926566	6	111762696	REV3L	C	T	0.4277	-0.1554	0.0242	1.45E-10
rs961764	6	117200993	RFX6	C	G	0.5738	0.1753	0.0243	5.14E-13
rs17574500	6	117496392	VGLL2	C	T	0.1366	0.1998	0.0361	3.22E-08
rs10782230	6	125907366	RP11-624M8.1	G	A	0.4826	0.192	0.0243	2.56E-15
rs35973496	6	126841981	MIR588	C	T	0.4407	0.4228	0.0242	2.67E-68
rs2327429	6	133888699	RP3-323P13.2	T	C	0.2919	-0.1779	0.0267	2.61E-11
rs9403140	6	140062358	RP11-15H7.2	G	A	0.6807	0.1633	0.0261	3.85E-10
rs6910072	6	142846719	RP11-440G9.1	A	G	0.1882	-0.188	0.0311	1.46E-09
rs62434125	6	150681235	IYD	T	C	0.0691	-0.6935	0.0475	2.65E-48
rs12199746	6	151575388	AKAP12	G	A	0.595	-0.1528	0.0247	6.16E-10
rs13203975	6	152011969	ESR1	G	A	0.1164	0.2871	0.0381	4.83E-14
rs633880	6	159263552	OSTCP1	T	C	0.8643	-0.2528	0.0359	1.83E-12
rs12661036	6	163316444	PACRG	T	C	0.2267	0.1671	0.0304	3.96E-08
rs12192157	6	169320808	RP3-495K2.3	C	T	0.5487	0.1772	0.0248	9.23E-13
rs2180052	6	170280901	RP1-182D15.2	G	A	0.1522	-0.2126	0.0344	6.49E-10
rs73033340	7	1156056	C7orf50	A	G	0.0302	-0.7233	0.0897	7.56E-16
rs6959688	7	1927195	MAD1L1	A	G	0.4022	0.197	0.0249	2.80E-15
rs2906157	7	2479369	AC004840.9	A	G	0.6421	-0.237	0.0255	1.72E-20
rs73049928	7	4630318	FOXK1	A	G	0.1913	0.1949	0.0315	6.08E-10
rs3807925	7	18503627	HDAC9	A	G	0.3497	0.1475	0.0254	6.21E-09
rs28688791	7	18999982	HDAC9	T	C	0.2006	0.3414	0.0303	1.69E-29
rs2023844	7	27203619	RP1-170O19.20	G	A	0.9245	0.8082	0.0457	7.34E-70

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs6961048	7	27288568	EVX1	C	G	0.1046	0.5082	0.0395	7.03E-38
rs1000597	7	30897563	AQP1	T	C	0.0755	0.3312	0.048	5.04E-12
rs17171688	7	40330306	SUGCT	G	A	0.0461	-0.4274	0.0619	5.23E-12
rs11977526	7	45968511	IGFBP3	G	A	0.4047	-0.3294	0.0246	9.67E-41
rs4948099	7	56004557	ZNF713	C	A	0.7504	-0.1583	0.0284	2.61E-08
rs37624	7	74472240	WBSCR16	G	A	0.6731	-0.1504	0.0258	5.46E-09
rs7796089	7	77882781	MAGI2	G	C	0.3387	-0.1981	0.0256	1.10E-14
rs42032	7	92608112	RN7SL7P	G	A	0.2658	-0.2964	0.0273	1.88E-27
rs2392929	7	106773623	CTA-360L10.1	T	G	0.2021	0.7111	0.0301	1.59E-123
rs1621	7	116797552	ST7	G	A	0.6685	-0.1696	0.0257	4.36E-11
rs13234407	7	130753455	LINC-PINT	G	A	0.4867	-0.2282	0.0239	1.39E-21
rs35680304	7	131288736	PODXL	C	T	0.5899	0.239	0.0248	5.86E-22
rs75672964	7	131636251	AC009518.4	C	T	0.043	0.4836	0.0654	1.43E-13
rs12535680	7	131658446	AC009518.4	A	G	0.3357	0.1807	0.0267	1.27E-11
rs10952729	7	140539703	BRAF	A	G	0.428	-0.1501	0.0243	6.15E-10
rs3918226	7	150993088	SMARCD3	C	T	0.0804	0.5655	0.0467	8.94E-34
rs10224210	7	151716108	GALNTL5	T	C	0.2805	0.353	0.0269	3.16E-39
rs62480403	7	155951558	Y_RNA	C	G	0.3177	-0.1943	0.0262	1.21E-13
rs1527575	7	156506577	LMBR1	A	G	0.6414	-0.1535	0.0253	1.38E-09
rs71499040	8	1763752	MIR596	G	C	0.7075	0.1818	0.0268	1.13E-11
rs7814142	8	10780042	XKR6	G	A	0.5865	-0.329	0.0244	2.26E-41
rs2872717	8	23545911	NKX3-1	T	C	0.2444	0.2266	0.0281	7.95E-16
rs4871964	8	26051281	RNA5SP258	A	G	0.2508	-0.3124	0.0279	3.63E-29
rs113814035	8	26162975	PPP2R2A	C	T	0.0983	0.2706	0.0408	3.25E-11
rs73226179	8	26614448	ADRA1A	G	A	0.0598	0.2863	0.0514	2.49E-08
rs1906672	8	38272507	FGFR1	G	A	0.2323	0.2515	0.0286	1.44E-18
rs1976450	8	48544382	SPIDR	G	T	0.1628	-0.1941	0.0326	2.68E-09
rs28601620	8	50812452	RP11-401N18.1	C	T	0.1974	0.2789	0.0305	5.27E-20
rs11988716	8	56240944	XKR4	A	G	0.1344	0.211	0.0358	3.62E-09

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs2354862	8	63589186	NKAIN3	A	C	0.3576	-0.1798	0.0253	1.09E-12
rs7006340	8	68024107	CSPP1	A	G	0.3054	0.1624	0.0264	7.19E-10
rs7819550	8	73294347	RNA5SP271	G	A	0.1922	-0.1805	0.0308	4.56E-09
rs2920926	8	75788830	RP11-758M4.4	A	G	0.7306	0.259	0.0276	6.36E-21
rs16939357	8	76778327	snoU13	T	C	0.158	-0.2717	0.0334	4.44E-16
rs4739832	8	81899784	PAG1	A	C	0.4118	-0.1946	0.0247	3.04E-15
rs62526122	8	91757341	TMEM64	G	A	0.2769	0.155	0.0284	4.99E-08
rs74900445	8	94256675	LINC00535	T	C	0.1196	0.2805	0.0372	4.58E-14
rs4582532	8	94957029	PSMA2P2	G	A	0.5073	-0.1559	0.0241	9.51E-11
rs1786345	8	100662523	VPS13B	A	C	0.4352	-0.1921	0.0244	3.77E-15
rs35783704	8	104954030	RIMS2	G	A	0.1009	-0.4395	0.0417	6.07E-26
rs7003445	8	105712958	RP11-127H5.1	T	C	0.4088	0.162	0.0248	6.72E-11
rs7341594	8	119397237	AC023590.1	G	A	0.2197	0.2841	0.0291	1.55E-22
rs6986368	8	125500955	RP11-158K1.3	A	T	0.3313	0.1809	0.0262	5.01E-12
rs4260863	8	128374367	RP11-382A18.2	C	G	0.39	-0.1878	0.025	5.36E-14
rs7012866	8	134604716	RP11-629O1.2	T	G	0.5004	0.2012	0.024	5.38E-17
rs4374967	8	140051700	RP11-324F11.1	A	G	0.59	-0.1518	0.0244	5.33E-10
rs4961230	8	140711985	KCNK9	G	A	0.299	0.2034	0.0269	3.78E-14
rs3802228	8	142910802	AC104417.1	A	G	0.5447	-0.2455	0.0242	4.09E-24
rs950196	9	9361571	PTPRD	T	A	0.1424	0.2811	0.0351	1.25E-15
rs7021547	9	13893247	LINC00583	A	G	0.1964	-0.1732	0.0303	1.08E-08
rs9886665	9	22942771	AL391117.1	T	C	0.7259	-0.1629	0.0275	3.36E-09
rs4553000	9	34223555	UBAP1	C	T	0.5127	-0.1471	0.0239	7.82E-10
rs76452347	9	35906474	HRCT1	C	T	0.2053	-0.2204	0.0309	9.93E-13
rs10869543	9	68930602	AL353763.2	T	A	0.7595	-0.1702	0.0291	5.15E-09
rs2224858	9	80817190	RP11-336N8.4	A	G	0.815	0.1789	0.0309	7.37E-09
rs10512183	9	87258815	NTRK2	T	G	0.2277	-0.1618	0.0289	2.15E-08
rs7045409	9	92439258	RP5-1050E16.2	T	A	0.3705	-0.1766	0.0252	2.25E-12
rs2297603	9	99015983	HSD17B3	G	T	0.1086	0.2387	0.0395	1.48E-09

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs10980141	9	110004372	RAD23B	G	C	0.1735	-0.2021	0.0335	1.51E-09
rs10759426	9	110391367	RNU6-996P	T	C	0.1275	0.3446	0.0362	1.69E-21
rs10817535	9	113929779	RP11-202G18.1	T	G	0.5933	-0.198	0.0247	1.21E-15
rs2416800	9	120781725	TPT1P9	A	G	0.5553	-0.1765	0.0249	1.26E-12
rs700080	9	123099416	CDK5RAP2	T	C	0.8756	0.2836	0.0437	8.40E-11
rs7023828	9	125736315	RABGAP1	C	T	0.4165	-0.2538	0.0246	4.91E-25
rs2243955	9	127541361	OLFML2A	A	G	0.6361	-0.1826	0.0255	7.88E-13
rs3861882	9	129703025	RALGPS1	T	C	0.2809	-0.1465	0.0268	4.79E-08
rs3847320	9	134041630	RP11-544A12.4	G	A	0.2712	-0.1718	0.0281	9.48E-10
rs12378485	9	136619086	VAV2	G	T	0.6519	-0.185	0.0262	1.79E-12
rs3814500	9	136975638	LL09NC01-251B2.3	C	T	0.7035	-0.1477	0.0266	2.81E-08
rs11252324	10	4082376	RP11-433J20.1	G	T	0.0781	-0.3242	0.0454	9.13E-13
rs56352451	10	5762902	RP11-336A10.2	C	T	0.1328	0.2009	0.0353	1.31E-08
rs11256837	10	10798572	RP1-186E20.1	G	A	0.2031	0.2024	0.0309	5.58E-11
rs1779240	10	18187384	MRC1	G	A	0.7662	-0.3861	0.0285	7.24E-42
rs12258967	10	18439030	CACNB2	C	G	0.2972	-0.5456	0.027	1.43E-90
rs11006779	10	27941657	MKX	T	C	0.5279	0.1821	0.0241	4.14E-14
rs61841111	10	30003641	SVIL	A	G	0.3874	-0.1452	0.0256	1.45E-08
rs211307	10	32060310	Y_RNA	A	G	0.2462	0.1782	0.028	1.87E-10
rs7099692	10	44896515	RP11-256K7.1	G	C	0.7118	-0.1796	0.0268	2.21E-11
rs34130368	10	47327566	RP11-144G6.10	A	C	0.117	0.2515	0.0387	7.87E-11
rs2574985	10	50369726	C10orf128	A	G	0.718	0.1493	0.0269	2.80E-08
rs57541197	10	61751548	C10orf40	G	A	0.14	-0.5336	0.0346	1.39E-53
rs2236295	10	62805132	LINC00845	G	T	0.3942	-0.2203	0.0249	8.47E-19
rs58413618	10	63474579	C10orf107	T	C	0.3308	0.1786	0.0258	4.28E-12
rs4746687	10	67547960	RP11-222A11.1	C	T	0.2184	0.1617	0.0292	3.07E-08
rs4746139	10	73647891	RP11-472G21.2	A	C	0.1417	0.4099	0.0349	7.00E-32
rs16931177	10	74111132	DNAJB12	C	G	0.6049	-0.1909	0.0247	1.22E-14
rs4120852	10	80258838	LINC00856	G	T	0.6098	-0.145	0.0247	4.13E-09

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs117472357	10	87855059	GRID1	C	T	0.0413	0.4611	0.0634	3.60E-13
rs2689690	10	94139949	MARCH5	C	T	0.3629	-0.2085	0.0253	1.82E-16
rs2274224	10	94279840	IDE	G	C	0.4366	-0.3594	0.0242	5.39E-50
rs11190709	10	100792906	HPSE2	G	A	0.8869	0.493	0.0383	5.59E-38
rs11191531	10	103047141	LBX1-AS1	G	C	0.0818	-1.0058	0.0446	1.74E-112
rs61863693	10	103960975	ELOVL3	G	A	0.2805	-0.1575	0.0278	1.41E-08
rs74926649	10	104517677	WBP1L	G	A	0.0163	0.6648	0.0972	7.81E-12
rs4506565	10	112996282	BTBD7P2	A	T	0.3146	0.2257	0.0259	3.20E-18
rs1801253	10	114045297	TECTB	G	C	0.732	0.399	0.0278	9.76E-47
rs11592107	10	121209450	GRK5	G	A	0.3079	0.2384	0.0261	6.13E-20
rs36213074	10	122456881	WDR11-AS1	T	C	0.1257	0.2603	0.0364	8.25E-13
rs7912283	10	131959515	GLRX3	G	A	0.6473	-0.2001	0.0256	5.88E-15
rs1133400	10	132645884	MIR378C	A	G	0.2132	0.2622	0.0298	1.24E-18
rs569550	11	1865838	TNNI2	T	G	0.3964	0.4891	0.0253	2.01E-83
rs78140146	11	2151575	IGF2	G	T	0.0454	0.3513	0.0611	8.79E-09
rs360153	11	9740727	SWAP70	T	C	0.5848	0.279	0.0244	3.49E-30
rs11042901	11	10633364	MRVI1	T	C	0.3343	0.1746	0.0269	8.72E-11
rs10766318	11	16289349	SOX6	C	T	0.7951	-0.4225	0.0298	1.32E-45
rs378769	11	16885459	PLEKHA7	G	A	0.2959	0.2471	0.0263	5.87E-21
rs11026590	11	22497332	CTD-2140G10.1	A	C	0.0746	0.3386	0.0473	8.18E-13
rs10501082	11	27249431	RP11-1L12.3	C	T	0.405	-0.1423	0.0245	6.64E-09
rs10835313	11	28359014	METTL15	G	A	0.505	-0.2029	0.0245	1.31E-16
rs11031051	11	30334160	ARL14EP	A	C	0.3126	0.1833	0.0261	2.01E-12
rs77175710	11	31881378	RCN1	A	G	0.1505	0.1944	0.0346	1.86E-08
rs5030178	11	32427023	WT1	A	C	0.364	0.1803	0.0253	1.07E-12
rs11037809	11	44020418	RP11-613D13.4	A	G	0.3035	-0.1532	0.0263	5.51E-09
rs3740781	11	45234375	PRDM11	C	T	0.7259	0.2059	0.0276	7.98E-14
rs12419342	11	47446993	PSMC3	C	T	0.7029	0.1909	0.0266	7.23E-13
rs7107356	11	47654618	MTCH2	A	G	0.5041	0.3874	0.024	2.08E-58

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs7101864	11	58285716	LPXN	A	G	0.603	0.1468	0.0247	2.77E-09
rs7125196	11	61505093	DAGLA	T	C	0.1181	-0.3417	0.0374	7.19E-20
rs2306363	11	65638129	EFEMP2	G	T	0.2013	-0.4046	0.0301	4.44E-41
rs7395791	11	69448148	CCND1	G	A	0.4415	-0.1922	0.0246	5.26E-15
rs10501410	11	72377762	PDE2A	G	A	0.0703	0.2713	0.0486	2.44E-08
rs7927515	11	76414286	GUCY2EP	C	A	0.3432	0.1886	0.026	3.87E-13
rs2289124	11	89491309	RP11-313I2.11	G	A	0.1683	-0.2968	0.0337	1.41E-18
rs604723	11	100739815	ARHGAP42	T	C	0.7223	0.5377	0.0273	2.12E-86
rs17653824	11	100870458	TMEM133	T	A	0.1207	0.2219	0.0372	2.56E-09
rs17737180	11	101195522	RP11-788M5.3	C	T	0.5571	0.1424	0.0244	5.08E-09
rs998328	11	107188572	RP11-819C21.1	A	G	0.291	-0.206	0.0271	2.80E-14
rs4938353	11	117175481	BACE1	G	A	0.8235	-0.2864	0.0318	2.07E-19
rs512730	11	117360341	DSCAML1	T	A	0.194	0.2124	0.0304	3.01E-12
rs2156805	11	122739860	CRTAM	A	G	0.5117	-0.172	0.0242	1.20E-12
rs11222084	11	130403335	RP11-535N6.2	A	T	0.3607	0.3286	0.0255	5.08E-38
rs4140962	12	366413	RP11-283I3.4	A	T	0.5633	-0.1882	0.0253	1.11E-13
rs11571376	12	950390	WNK1	C	G	0.2951	0.1985	0.0267	1.10E-13
rs3819536	12	2327832	LRG_334	G	A	0.7053	0.1755	0.0262	2.25E-11
rs10845556	12	12484415	MANSC1	A	G	0.4207	-0.1535	0.0243	2.73E-10
rs2024385	12	12735504	RP11-158N24.1	T	A	0.425	-0.2345	0.0252	1.26E-20
rs1010064	12	19847381	AEBP2	A	C	0.1848	-0.3232	0.031	1.97E-25
rs73075659	12	20220607	RP11-664H17.1	A	G	0.3315	-0.3431	0.0259	5.56E-40
rs2129869	12	26304717	SSPN	A	T	0.222	0.2044	0.0289	1.56E-12
rs61915422	12	27779629	PPFIBP1	C	G	0.129	-0.2034	0.0369	3.55E-08
rs7963397	12	32505703	BICD1	C	T	0.3951	0.1396	0.0246	1.41E-08
rs7316233	12	47811021	RP1-90J4.1	A	G	0.1453	0.243	0.0342	1.27E-12
rs150857355	12	48815557	RP11-370I10.6	G	C	0.0217	0.8301	0.0895	1.71E-20
rs60121503	12	49694964	RP11-161H23.9	G	C	0.0843	-0.2384	0.0433	3.68E-08
rs7302981	12	50144032	TMBIM6	A	G	0.623	-0.3553	0.0246	2.10E-47

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs10876531	12	54049934	ATP5G2	A	C	0.294	-0.3653	0.0268	3.59E-42
rs7979246	12	57698771	R3HDM2	A	G	0.6717	-0.1418	0.0257	3.65E-08
rs7306710	12	65982311	RP11-230G5.2	T	C	0.5238	0.2095	0.0243	7.20E-18
rs775634	12	67253624	RP11-123O10.4	A	G	0.7713	-0.1906	0.0293	7.76E-11
rs1245811	12	79243673	SYT1	T	C	0.6035	-0.1584	0.0254	4.22E-10
rs11614730	12	79532022	SYT1	G	A	0.6703	-0.2157	0.0265	3.68E-16
rs11105375	12	89694803	RP11-13A1.3	T	A	0.1679	-0.7405	0.0325	2.94E-115
rs10777213	12	89956222	RP11-981P6.1	G	A	0.5252	-0.1695	0.024	1.59E-12
rs56725874	12	94267735	CRADD	T	G	0.1194	-0.2194	0.0381	8.68E-09
rs17797222	12	102520168	PARPBP	A	G	0.2225	-0.182	0.0297	9.13E-10
rs597808	12	111535554	CUX2	A	G	0.5094	-0.5614	0.0292	1.57E-82
rs551510	12	114676231	RP11-139B1.1	C	T	0.5037	0.1378	0.0244	1.65E-08
rs1896326	12	114905151	RN7SKP216	G	A	0.2283	-0.2306	0.0297	7.61E-15
rs35429	12	115118062	TBX3	A	G	0.3874	-0.3911	0.0249	2.11E-55
rs4767332	12	115494467	RP11-25E2.1	C	A	0.5725	0.2184	0.0241	1.53E-19
rs117037316	12	132499688	EP400	C	G	0.0686	0.3414	0.0511	2.28E-11
rs682681	13	21719923	SAP18	T	C	0.668	0.2217	0.0258	9.46E-18
rs7982886	13	27302192	GPR12	G	A	0.1368	-0.2271	0.0351	9.85E-11
rs866573	13	29583115	MTUS2	C	T	0.1553	0.2868	0.0333	6.55E-18
rs9532243	13	31617271	WDR95P	A	C	0.5185	-0.2679	0.024	5.58E-29
rs190533862	13	40097000	LHFP	T	A	0.0641	0.2875	0.0517	2.75E-08
rs112684153	13	41275595	MIR320D1	T	C	0.0735	-0.3396	0.0498	9.24E-12
rs56249339	13	42191956	VWA8	G	T	0.1132	0.2245	0.0384	4.81E-09
rs1886220	13	46585349	ZC3H13	A	G	0.7556	0.204	0.0282	4.88E-13
rs770379	13	50565635	TRIM13	A	G	0.5465	-0.141	0.0241	5.23E-09
rs17245822	13	72557556	RNA5SP32	A	C	0.3685	0.1456	0.0254	1.04E-08
rs78474310	13	73252764	RPL21P110	A	G	0.0435	0.4824	0.0593	4.18E-16
rs9530914	13	79389116	LINC00331	G	T	0.3962	0.148	0.0253	4.95E-09
rs9515270	13	110699726	LINC00396	C	T	0.1779	0.1815	0.0318	1.19E-08

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs72661946	13	112974382	LINC01043	G	C	0.0868	0.3121	0.0456	7.89E-12
rs7320104	13	114294290	TFDP1	A	G	0.1529	0.3805	0.0342	1.02E-28
rs365990	14	23392602	PRMT5	A	G	0.3681	-0.1717	0.0248	4.57E-12
rs8904	14	35402011	RP11-85K15.2	G	A	0.3699	0.2819	0.0249	1.01E-29
rs7493678	14	38931713	RP11-14N4.1	A	T	0.3454	0.1569	0.0255	7.70E-10
rs72683923	14	50269229	NEMF	T	C	0.0202	-0.9677	0.0912	2.57E-26
rs3015451	14	50604373	SOS2	C	A	0.7226	-0.1486	0.027	3.52E-08
rs35413927	14	52953640	TXNDC16	A	G	0.3043	0.2462	0.0266	2.10E-20
rs67468157	14	54862556	CDKN3	C	T	0.3119	0.1571	0.026	1.51E-09
rs57140819	14	67551530	GPHN	C	G	0.173	-0.2022	0.0321	2.91E-10
rs2239268	14	72002874	SIPA1L1	G	A	0.6984	0.1859	0.0261	1.09E-12
rs3815460	14	72955551	RGS6	C	G	0.1032	0.2287	0.0396	7.51E-09
rs2165197	14	74613001	LIN52	T	C	0.4895	0.1833	0.0241	2.99E-14
rs11622562	14	77050805	RP11-187O7.3	C	T	0.3304	-0.1572	0.0256	7.71E-10
rs10134573	14	88424561	GALC	C	T	0.0766	-0.2572	0.046	2.19E-08
rs78154752	14	98130216	RP11-76E12.1	T	G	0.2075	0.2609	0.03	3.84E-18
rs17562391	14	99666913	BCL11B	C	T	0.418	0.2067	0.0244	2.31E-17
rs10149239	14	100371670	EML1	G	A	0.7386	0.1653	0.0275	1.88E-09
rs12881760	14	100709998	YY1	G	C	0.6575	0.155	0.026	2.39E-09
rs17679259	14	103428965	CDC42BPB	A	G	0.1331	-0.3583	0.0356	7.17E-24
rs8030856	15	40022766	FSIP1	C	G	0.4	0.1378	0.0248	2.66E-08
rs12101934	15	41060881	DNAJC17	C	T	0.426	0.2261	0.0248	7.69E-20
rs17687755	15	41937640	MGA	G	C	0.2956	0.1737	0.0265	5.65E-11
rs689304	15	48630163	DUT	C	T	0.0974	-0.3961	0.0408	3.00E-22
rs3098186	15	50518424	SLC27A2	C	T	0.5174	-0.202	0.0244	1.28E-16
rs7174174	15	63080128	TLN2	T	C	0.2259	0.1998	0.0292	8.02E-12
rs28429256	15	66639279	TIPIN	G	A	0.3341	0.1986	0.0262	3.76E-14
rs117539635	15	69390577	RP11-809H16.5	A	G	0.0237	-0.9659	0.1115	4.47E-18
rs11636952	15	74821981	ARID3B	T	C	0.6807	-0.4329	0.026	3.11E-62

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs7171632	15	80711475	ARNT2	G	C	0.4461	0.2988	0.0242	5.92E-35
rs28515053	15	85105323	UBE2Q2P1	G	A	0.0299	0.4581	0.0716	1.58E-10
rs2046341	15	85497641	SLC28A1	G	A	0.1915	-0.209	0.0304	6.68E-12
rs77032376	15	89467549	MFGE8	C	T	0.1487	-0.2059	0.0351	4.28E-09
rs8032315	15	90875067	GABARAPL3	T	A	0.3009	0.5716	0.0287	2.90E-88
rs12906962	15	94768842	MCTP2	T	C	0.326	0.2322	0.0261	5.48E-19
rs2589254	15	96247180	RNU2-3P	A	G	0.2525	0.1863	0.0275	1.30E-11
rs11853701	15	97759615	RP11-315L6.1	G	C	0.3338	0.1427	0.026	4.17E-08
rs11852739	15	99559266	RP11-654A16.3	G	A	0.7387	0.1711	0.0277	6.99E-10
rs2281226	16	1314766	PRSS29P	A	C	0.608	-0.1651	0.025	3.96E-11
rs2286472	16	2020567	TBL3	A	C	0.3197	0.2352	0.0261	2.29E-19
rs66768878	16	3972637	RP11-462G12.2	T	C	0.1533	0.2482	0.0341	3.67E-13
rs1982784	16	4616536	C16orf96	A	G	0.8105	-0.1723	0.0313	3.68E-08
rs80100774	16	4628546	C16orf96	T	G	0.0771	-0.3178	0.0512	5.54E-10
rs12102426	16	4872449	GLYR1	T	G	0.428	-0.2764	0.0243	4.72E-30
rs72779543	16	19148142	CTD-2349B8.1	G	C	0.212	0.2012	0.0331	1.25E-09
rs9928003	16	20346926	UMOD	A	C	0.1788	-0.2744	0.0315	2.66E-18
rs9937801	16	21076809	DNAH3	T	C	0.4295	-0.1755	0.0241	3.30E-13
rs73551577	16	24833822	TNRC6A	G	A	0.1905	-0.3383	0.0309	6.84E-28
rs11646118	16	29915150	ASPHD1	T	C	0.083	-0.2671	0.0446	2.17E-09
rs72793375	16	30726584	SRCAP	G	A	0.2163	0.18	0.0308	4.91E-09
rs10468291	16	49734135	ZNF423	C	A	0.5695	-0.164	0.0245	2.31E-11
rs12149258	16	50502915	SNORD112	A	G	0.1591	-0.2351	0.0331	1.23E-12
rs34869093	16	51723331	RN7SKP142	A	G	0.3707	0.1758	0.0252	2.83E-12
rs2042364	16	65222226	RP11-256I9.2	A	G	0.5358	0.1524	0.0241	2.80E-10
rs72792102	16	66782091	DYNC1LI2	A	G	0.0415	0.4059	0.0623	7.34E-11
rs62050280	16	70622878	IL34	G	A	0.048	0.6397	0.0733	2.53E-18
rs4888408	16	75398926	CFDP1	G	A	0.5851	0.3305	0.0245	1.70E-41
rs12926550	16	81476550	CMIP	G	A	0.3171	-0.1998	0.026	1.66E-14

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs8046196	16	82967743	CDH13	G	T	0.5131	0.15	0.0242	5.95E-10
rs8052276	16	87970986	CA5A	T	C	0.6498	-0.164	0.0261	3.32E-10
rs453528	16	89627353	RPL13	A	C	0.4794	-0.2098	0.0254	1.37E-16
rs9303175	17	1469693	SLC43A2	T	G	0.6548	0.2097	0.0261	9.97E-16
rs11653927	17	2108800	SMG6	C	T	0.3835	-0.2475	0.0249	3.02E-23
rs113086489	17	7268037	TMEM95	C	T	0.5508	0.3023	0.0249	7.60E-34
rs4511593	17	7552219	ATP1B2	C	T	0.6518	-0.2584	0.0255	3.96E-24
rs117285318	17	7967324	RP11-599B13.3	T	C	0.0763	-0.4159	0.0472	1.25E-18
rs3935505	17	17978706	GID4	T	G	0.2509	0.155	0.028	3.19E-08
rs4925159	17	18282196	EVPLL	G	A	0.4252	0.2184	0.0243	2.71E-19
rs7211535	17	20019051	SPECC1	A	G	0.5208	0.1411	0.0246	9.37E-09
rs4795637	17	31700433	ASIC2	T	C	0.2354	0.1869	0.0283	4.30E-11
rs9899540	17	32450906	ASIC2	A	T	0.5969	-0.1699	0.0253	1.82E-11
rs7213273	17	45078546	RP11-156P1.3	G	A	0.6562	-0.3234	0.0252	1.42E-37
rs17608766	17	46935905	CALCOCO2	T	C	0.1403	0.5723	0.0348	7.46E-61
rs9221	17	48598451	MYCBPAP	C	T	0.9114	0.3715	0.0426	2.86E-18
rs1962412	17	48892897	RP11-294J22.5	T	C	0.6975	0.2003	0.0265	4.40E-14
rs9897429	17	49441016	RP11-1018N14.2	G	A	0.5173	0.2201	0.0254	4.57E-18
rs1292053	17	59886176	BRIP1	A	G	0.4399	0.1687	0.0242	3.01E-12
rs2286526	17	61394696	AC037445.1	C	T	0.7105	0.3706	0.0277	7.53E-41
rs3826537	17	62692442	SMURF2	A	G	0.4186	0.1986	0.0245	4.60E-16
rs6504213	17	64304354	PRKCA	T	C	0.5844	0.2417	0.0251	6.25E-22
rs1436138	17	77320798	RBFOX3	A	G	0.3669	-0.2478	0.0251	6.39E-23
rs113393496	17	78372197	CTD-2047H16.4	G	A	0.1955	-0.172	0.0308	2.47E-08
rs1057040	17	78803713	RPTOR	G	A	0.528	-0.2048	0.0239	1.23E-17
rs10459965	17	81394216	AC139099.3	A	T	0.268	0.1939	0.0319	1.25E-09
rs9675039	17	83078468	AC139099.3	G	A	0.3673	0.1459	0.0255	1.07E-08
rs12955644	18	771046	YES1	C	G	0.1792	-0.3147	0.0324	2.59E-22
rs880132	18	7131619	SLC25A51P2	T	C	0.5733	-0.1821	0.0268	1.04E-11

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs62082230	18	25096107	RP11-739N10.1	T	A	0.2779	-0.1656	0.0277	2.38E-09
rs163049	18	26943536	CTD-2515C13.2	G	A	0.6043	0.1568	0.0249	3.11E-10
rs4637011	18	44410733	PIAS2	G	C	0.306	-0.2949	0.0266	1.24E-28
rs4890500	18	45019955	CTD-2130O13.1	A	G	0.2983	0.2383	0.0262	8.39E-20
rs7236548	18	45517785	ZBTB7C	C	A	0.1856	0.2819	0.031	9.74E-20
rs36010659	18	50757579	DCC	T	C	0.1408	-0.2318	0.0347	2.42E-11
rs10048404	18	56911251	GRP	C	T	0.3661	-0.2161	0.0266	4.63E-16
rs74625348	18	63179197	RP11-453M23.1	G	C	0.2311	-0.1794	0.0294	1.06E-09
rs1898165	18	75328421	RP11-63N3.1	G	A	0.5247	-0.1939	0.0241	8.22E-16
rs117777118	18	79401324	PARD6G	G	A	0.0402	-0.4587	0.0764	1.97E-09
rs141958336	19	2165384	DOT1L	G	A	0.0419	0.7392	0.0633	1.77E-31
rs12150907	19	4940618	UHRF1	G	A	0.1937	0.2031	0.031	6.05E-11
rs7248104	19	7224420	INSR	G	A	0.4123	-0.2167	0.0243	5.24E-19
rs67427762	19	7269891	INSR	C	T	0.1551	-0.347	0.0351	4.78E-23
rs10409243	19	10222312	P2RY11	C	T	0.5866	-0.2457	0.0253	2.34E-22
rs167479	19	11416089	CTC-510F12.4	G	T	0.5234	-0.494	0.0258	5.50E-82
rs7253742	19	17060879	CPAMD8	T	C	0.4077	0.1624	0.0243	2.52E-11
rs4319878	19	21741650	RP11-678G14.2	C	T	0.5615	0.1356	0.0246	3.61E-08
rs8108027	19	21933099	ZNF100	G	C	0.2949	0.1861	0.0267	2.95E-12
rs8108717	19	31421008	CTC-400I9.3	A	G	0.6119	-0.2271	0.0248	5.39E-20
rs41119	19	33524998	RHPN2	G	T	0.3489	0.1583	0.0268	3.69E-09
rs7412	19	44908822	CTC-512J12.4	C	T	0.081	-0.3921	0.0452	4.03E-18
rs11673093	19	45238836	snoZ6	G	A	0.2607	0.2221	0.028	2.24E-15
rs602662	19	48703728	CARD8	G	A	0.5178	0.1668	0.0241	4.45E-12
rs6085527	20	6386343	CASC20	A	G	0.5938	0.195	0.0246	2.44E-15
rs2423514	20	10712689	RP11-103J8.1	A	G	0.4572	-0.2601	0.0243	7.99E-27
rs1327235	20	10988382	RP11-103J8.2	A	G	0.4689	0.3732	0.0239	5.73E-55
rs6078093	20	11188021	RP4-734C18.1	G	A	0.4244	-0.1738	0.0243	8.77E-13
rs8125763	20	17902887	RNU6-192P	C	A	0.4731	0.1459	0.0242	1.54E-09

RSID	Chromosome	Position (GRCh37)	Nearest gene	Reference allele	Effect allele	Effect allele frequency	Beta	Standard error	P value
rs17812022	20	19026455	AL136090.1	C	T	0.0954	-0.3153	0.043	2.26E-13
rs6088749	20	31608553	BPIFB2	C	G	0.165	-0.2376	0.0324	2.22E-13
rs79384779	20	32627142	RALY	C	T	0.1549	0.3045	0.0344	8.16E-19
rs4812536	20	41630070	PTPRT	A	G	0.3422	-0.2155	0.0258	6.53E-17
rs6017279	20	44173331	EPPIN-WFDC6	T	C	0.1345	0.3864	0.0351	3.85E-28
rs4325821	20	48608249	SNAI1	T	C	0.489	-0.1444	0.0246	4.67E-09
rs6021246	20	51492425	RP11-80K6.2	T	C	0.3231	0.1758	0.0261	1.66E-11
rs2801008	20	53172179	DOK5	T	G	0.3191	0.1757	0.0262	2.01E-11
rs6026739	20	59164414	MIR548AG2	A	T	0.1236	0.562	0.0369	2.27E-52
rs3795154	20	63887088	LINC00266-1	C	T	0.3621	0.1871	0.0253	1.33E-13
rs6062344	20	64064671	LINC00266-1	C	T	0.4969	-0.2631	0.0245	7.76E-27
rs2178895	21	14961085	OR4K12P	C	T	0.6391	0.1755	0.0252	3.65E-12
rs2823139	21	15204463	CNN2P7	G	A	0.3386	0.2282	0.0259	1.33E-18
rs6517423	21	15421720	AP001347.6	G	A	0.2957	0.1478	0.0264	2.17E-08
rs2298359	21	32298829	KRTAP11-1	T	C	0.0622	-0.4544	0.0668	1.00E-11
rs2833834	21	32442070	UBE3AP2	C	A	0.2797	0.1908	0.027	1.63E-12
rs12627514	21	43339560	C2CD2	C	G	0.2879	0.2717	0.0276	6.55E-23
rs34487963	21	43418450	ZBTB21	C	A	0.0224	-0.6916	0.1011	7.95E-12
rs7278003	21	43546188	UMODL1	T	C	0.5648	0.1858	0.0242	1.84E-14
rs5992134	22	17951228	CECR2	G	T	0.2383	0.1557	0.0285	4.69E-08
rs2238787	22	19988883	ARVCF	G	A	0.2901	0.2381	0.0266	4.12E-19
rs12321	22	29057205	TTC28	G	C	0.432	-0.1852	0.0243	2.39E-14
rs5994376	22	31167569	OSBP2	A	C	0.5621	0.1443	0.0253	1.15E-08
rs5756813	22	37779470	ELFN2	G	T	0.6129	-0.1506	0.025	1.81E-09
rs28578714	22	50289492	ALG12	T	C	0.3918	-0.1844	0.0255	4.71E-13

Abbreviations. **SBP**: Systolic Blood Pressure; **SNP**: Single Nucleotide Polymorphism; **RSID**: Reference SNP Cluster ID; **GRCh37**: Genome Reference Consortium Human Build 37