

Genetic and Observational Associations of Nuclear Magnetic Resonance-measured Lipid-related Traits and Atherosclerotic Cardiovascular Mortality in the Mexico City Prospective Study

Clinical Trial Service Unit & Epidemiological Studies Unit

Nuffield Department of Population Health

Yunhe Wang, B.S.Med, MSc

Brasenose College, University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Michaelmas 2024

Abstract

Background: Blood lipids, particularly low-density lipoproteins (LDL) and other apoB-containing particles, play a central role in the pathophysiology of atherosclerosis and atherosclerotic cardiovascular diseases (ASCVD). However, previous evidence is mainly based on European populations and traditional lipid panels. This thesis examines the genetic and observational associations of NMR-measured lipid traits with ASCVD mortality in a population of Latin American ancestry.

Methods: In 1998-2004, the Mexico City Prospective Study (MCPS) recruited 150,000 adults aged >35 years from two districts of Mexico City, with follow up for cause-specific death until December 31, 2020. 120,000 participants had measures of NMR metabolomics by the time of analysis. Cross-sectional associations of the NMR lipid panel with a traditional chemistry lipid panel and baseline participant characteristics were assessed by Pearson correlation and linear regression. A genome-wide association study (GWAS) assessed the genetic structure of major composite lipid traits (LDL-C, HDL-C, TG, ApoB, and ApoA1). Cox regression yielded adjusted hazard ratios (HRs) for the association of both composite lipid traits and lipid subfractions with ASCVD mortality (defined as ischaemic heart disease, ischaemic stroke, or peripheral arterial disease) among participants aged 35-84 years without chronic diseases (except diabetes) and not taking lipid-lowering drugs (<1%). Mendelian randomization (MR) examined whether composite lipid traits had a causal effect on risk of ASCVD.

Results: 107,504 eligible participants aged 35-84 years were included in the main analyses (33% male; mean [SD] age, 51 [12] years). In cross-sectional analyses, complex correlation patterns within and across the same lipoprotein category within NMR lipid panels were observed, in line with their biochemical structures and interactions in lipid metabolism. NMR lipid traits were also associated with a range of baseline sociodemographic and lifestyle characteristics. GWAS of five major lipid traits revealed hundreds of conditionally independent signals, including multiple previously known

genetic loci involved in lipid metabolism (e.g., variants in *APOE, CELSR2, PCSK9, APOB, HMGCR, LPA, LPL, LDLR* for LDL-C) plus several potentially novel ones. In observational analyses of clinically relevant composite lipid traits assessed individually, *usual* level of ApoB and TG, but *not* of LDL-C, were positively associated with risk, while *usual* level of ApoA1 and HDL-C were inversely associated with risk in women but positively associated with risk in men. After accounting for the effect of other lipids, given LDL-C, usual ApoB was strongly *positively* associated with risk (per 2SD change: HR, 1.76, 1.71-1.81) while, given ApoB, usual LDL-C was strongly *negatively* associated with risk (0.58, 0.56-0.59). These results both implicate small dense LDL particles as being particularly atherogenic. The strong, log-linear association with TG was decreased after accounting for ApoB. Further investigation of lipid-related subfractions confirmed that the concentration of LDL particles (rather than the cholesterol in them *per se*) was more closely related to ASCVD mortality risk, with each ApoB-containing particle (VLDL-P, L-LDL-P and S-LDL-P) that has exactly one ApoB molecule related to similar risk of ASCVD. Smaller HDL particles (S-HDL-P) and their cholesterol contents (S-HDL-C) were negatively associated with risk, whereas larger ones (XL-HDL-P and XL-HDL-C) were positively associated with risk. These results suggest that for specific types of lipoprotein particles (LDL and HDL), their subfractions with varying sizes differ in their relevance to risk. Further MR analyses supported that only ApoB was independently and significantly related to an increased risk of ASCVD mortality in a multivariable setting and suggested ApoB, as the accurate measure of apoB particles, is causally related to the risk of ASCVD, independent of lipid contents such as TG and cholesterol, and HDL or ApoA1. MR findings did not support TG as a causal factor in the development of ASCVD.

Conclusion: Given the totality of the complementary evidence, ApoB is the strongest causal lipid factor and should be the primary measure of the atherogenic risk of the apoB lipoproteins, as well as the primary measure of the effectiveness and adequacy of therapy to lower the apoB lipoproteins in the prevention of ASCVD. Given the disproportionately high burden of ASCVD and related risk factors, such as diabetes and

obesity, in Mexico, the underuse of lipid-lowering drugs is a critical concern. Enhancing lipid management and control of ASCVD through the use of such medications, particularly statins targeting ApoB particles, is urgently needed. This approach could help alleviate the growing burden of ASCVD and related chronic diseases on a large scale, improving overall population well-being.

Statement of contribution to work

The initial idea of this project based on MCPS was proposed by my supervisors Jonathan Emberson and Jason Torres before my admission. I was responsible for shaping the general ideas of the project into thesis by the design, conception, and development of research questions with the continuous support from my supervision team (Jonathan Emberson, Jason Torres, and Diego J Aguilar-Ramirez) and other colleagues from the MCPS group and NDPH. I conducted the systematic literature reviews and planned the study with input from my supervisors. I conducted all statistical analyses, and generated all the tables and figures presented (part of the genetic analyses used the GWAS pipeline developed by Jason Torres and part of the observational analyses used established statistical programmes used and developed by MCPS team researchers including Jonathan Emberson, Rachel Wade, and Diego J Aguilar-Ramirez). I summarized and interpreted the findings and drafted the current thesis, which was further refined with the guidance and support from my supervisors.

Acknowledgments

I am deeply grateful to the participants from Mexico City Prospective Study for their altruistic contributions, as well as to the survey and project development and management teams in both Mexico and Oxford.

I am profoundly thankful to my supervisors, Jonathan Emberson, Jason Torres, and Diego J Aguilar-Ramirez for their trust, dedication, and unwavering support and guidance during my DPhil.

Thanks are due, also, to other members of the MCPS team and to other staff in the CTSU, including Robert Clarke, Rory Collins, Louisa Gnatiuc, Mike Hill, Mike Turner, Sarah Lewington, Richard Peto and Rachel Wade.

I would like to thank my friends and collaborators from Oxford and around the world with whom I've had the pleasure of working, including Junqing Xie, Daniel Prieto-Alhambra, and Michael V Vitiello.

I would like to extend my deepest gratitude to my family for their unwavering support throughout this journey. To my parents, your love, encouragement, and belief in me have been the foundation upon which I've built this achievement. You have always been my greatest source of strength, especially during the most challenging moments of this thesis.

Last but not least, I would also like to thank the scholarships from the Clarendon Fund, NDPH, and Brasenose college for enabling me to undertake this DPhil training.

Publications

*First author #Corresponding

1. **Wang, Y**.*, Su, B., Xie, J. et al. Long-term risk of psychiatric disorder and psychotropic prescription after SARS-CoV-2 infection among UK general population. ***Nature Human Behaviour*** (2024)
2. **Wang, Y**.*, Su, B., Alcalde-Herraiz, M. et al. Healthy lifestyle for the prevention of post-COVID-19 multisystem sequelae, hospitalization, and death. ***Nature Communications*** (2024)
3. **Wang, Y**.*, Alcalde-Herraiz, M, López-Güell, K. et al. Refinement of post-COVID condition core symptoms, subtypes, determinants, and health impacts: A cohort study integrating real-world data and patient-reported outcomes. ***eBioMedicine*** (2024, in press)
4. **Wang, Y**.*# et al. Proteomic aging clock across organ systems predicts psychiatric and neurodegenerative diseases and links distinct pathogenic pathways (***Nature Aging***, 2024 under peer-review)
5. **Wang, Y**.*, Xuan, Y., Su, B. et al. Predicting recurrence and survival in patients with non-metastatic renal-cell carcinoma after nephrectomy: a prospective population-based study with multicenter validation. ***International Journal of Surgery*** (2024)
6. Qin, H., … **Wang, Y**.# Trajectories of general and central obesity beyond middle age in relation to late-life cognitive decline and dementia. ***Obesity*** (2024, in press)
7. Zhu, Y., ... **Wang, Y**#. & Jin, L#. Transition in metabolic health phenotypes across general adiposity categories and association with the risk of depression: a prospective analysis. ***European Psychiatry*** (2024)
8. Hu, G., Qin, H., … **Wang, Y**# Composite healthy lifestyle, socioeconomic deprivation, and mental well-being during the COVID-19 pandemic: a prospective analysis. ***Molecular Psychiatry*** (2023)
9. Sun, Y.*, Zhang, Z.*, Wu, S.*, **Wang, Y**.*, et al. Burden of mental and substance use disorders in China and its provinces, 1990-2021: a systematic analysis for the global burden of disease study 2021. (Under consideration)
10. Xie, J., Prats-Uribe, A., Feng, Q., **Wang, Y**., Gill, D., Paredes, R., & Prieto-Alhambra, D. Clinical and genetic risk factors for acute incident venous thromboembolism in ambulatory patients with COVID-19. ***JAMA internal medicine*** (2022)
11. Xie, J., Mothe, B., Alcalde Herraiz, M., Li, C., Xu, Y., Jödicke, A. M., **Wang, Y**.... &Alhambra, D. P. Relationship between HLA genetic variations, COVID-19 vaccine antibody response, and risk of breakthrough outcomes. ***Nature Communications*** (2024)
12. **Wang, Y**.#, Shi, L., Que, J. et al. The impact of quarantine on mental health status among general population in China during the COVID-19 pandemic. ***Molecular psychiatry*** (2021)

List of Abbreviations

ApoA1	Apolipoprotein A-I
ApoB	Apolipoprotein B
ApoC3	Apolipoprotein C-III
BMI	Body-mass index
C	Cholesterol
HMG-CoA	β-Hydroxy β-methylglutaryl coenzyme A
CVD	Cardiovascular disease
SMCs	Smooth muscle cells
CE	Esterified cholesterol
CI	Confidence interval
FC	Free cholesterol
TG	Triglycerides
PL	Phospholipids
L	Total lipids
D	Mean particle size
TotCho	Total cholines
PC	Phosphatidylcholine
SM	Sphingomyelin
Lac	Lactate
Cit	Citrate
Glc	Glucose
Ala	Alanine
Gln	Glutamine
His	Histidine
Ile	Isoleucine
Leu	Leucine
Val	Valine
Phe	Phenylalanine
eGFR	Estimated glomerular filtration rate
HR	Hazard ratio
Tyr	Tyrosine
Ace	Acetate
AcAce	Acetoacetate
bOHBut	Beta-hydroxy-butyrate
Alb	Albumin
Crea	Creatinine
GlycA	Glycoprotein-A
PUFA	Poly-unsaturated fatty acids
MUFA	Mono-unsaturated fatty acids
SFA	Saturated fatty acids
DHA	Docosahexaecoic fatty acid
LA	Linoleic acid
FAw3	Omega-3 fatty acids
FAw6	Omega-6 fatty acids
TotFA	Total fatty acids
CKD	Chronic kidney disease
CTSU	Clinical Trial Service Unit
EDTA	Ethylene-diamine-tetra-acetic acid
FFA	Free fatty acids
HC	Hip circumference
HDL	High-density lipoprotein
IDL	Intermediate-density lipoprotein
L	Large
TRL	Triglyceride-rich lipoprotein
LDL	Low-density lipoprotein
LPL	Lipoprotein lipase
CETP	Cholesteryl ester transfer protein
PCSK9	Proprotein convertase subtilisin/kexin type 9
M	Medium
MCPS	Mexico City Prospective Study
MS	Mass spectrometry
nm	Nanometres
NMR	Nuclear magnetic resonance
OWOB	Overweight and obesity
P	Particle
QC	Quality control
S	Small
SD	Standard deviation
VLDL	Very low-density lipoprotein
W	Weight
WC	Waist circumference
WHR	Waist-hip ratio
WHTR	Waist-height ratio
XL	Very large
XS	Very small
FH	Familial hypercholesterolaemia
HbA_{1c}	Glycosylated haemoglobin
IHD	Ischaemic heart disease
US	United States
UK	United Kingdom
PC	Principal component
FDR	False discovery rate
ERFC	Emerging risk factors collaboration
MI	Myocardial infarction
OR	Odds ratio
PAD	Peripheral arterial disease
CAD	Coronary artery disease
IS	Ischaemic stroke
CHD	Coronary heart disease
ICD-10	International Classification of Diseases 10^{th} Revision
SBP	Systolic blood pressure
DBP	Diastolic blood pressure

Table of contents

Abstract **2**
Statement of contribution to work **5**
Acknowledgments **6**
Publications **7**
List of Abbreviations **8**
Table of contents **9**
List of Tables **15**
List of Figures **16**
CHAPTER 1: Introduction **18**
1.1 Atherosclerotic cardiovascular disease (ASCVD) 19
1.1.1 Epidemiology of ASCVD *19*
1.1.2 Pathophysiology of atherosclerosis *19*
1.2 Lipids, lipoproteins, apolipoproteins and ASCVD 21
1.2.1 Components and metabolism of lipoproteins *21*
1.2.2 LDL receptor and lipid-lowering therapies *22*
1.2.3 Causal effect of LDL on ASCVD *23*
1.2.4 Variability in apoB-containing lipoproteins, apoB, and ASCVD *25*
1.2.5 HDL, HDL-targeted therapies and ASCVD *27*
1.2.6 TG, LPL, and ASCVD *29*
1.2.7 Lp(a), LPA, and ASCVD *32*
1.3 Study setting 34
1.4 Thesis objectives 35
References 41
CHAPTER 2: Literature review **52**
2.1 Background and objectives 53
2.2 Methods of literature search and systematic review 54
2.2.1 Search strategy *54*

2.2.2 Eligibility criteria *54*
2.2.3 Study selection and data extraction *56*
2.3 Findings of observational study 56
2.3.1 Study inclusion *56*
2.3.2 Nested case-control and case-cohort studies *57*
2.3.3 Cohort studies *59*
2.4 Summary of observational studies and gap in current evidence 61
2.4.1 ApoB-containing lipoproteins, LDL/VLDL particle size, and ASCVD *61*
2.4.2 Concentration and size of HDL particle and ASCVD *62*
2.4.3 Lipids contained in lipoprotein particles and ASCVD *63*
2.4.4 Additional value of NMR lipid traits to traditional lipid panel in ASCVD risk prediction *63*
2.4.5 Gap in current evidence from observational studies and implications for future study *64*
2.5 Findings of Mendelian randomization studies 66
2.5.1 Study inclusion *66*
2.5.2 Three sample MR studies using traditional univariable and multivariable analyses *68*
2.5.3 Three sample MR studies using MR-BMA analyses *70*
2.5.3 Two sample MR studies using traditional univariable and multivariable analyses *72*
2.5.4 Gap in current evidence from MR studies and implications for future study *73*
2.6 Findings of genome-wide association studies 75
2.6.1 Study inclusion *75*
2.7 Conclusions 77
References 101
CHAPTER 3: Overall study design and methods 105
3.1 Study design, procedures, and follow-up 106
3.1.1 Overall study design and recruitment of participants *106*
3.1.2 Baseline assessment procedures *106*
3.1.3 Physical measurements *107*
3.1.4 Blood sample *107*

3.1.5 Follow-up and adjudication *108*
3.1.6 Funding and ethics approval *109*
3.2 Measurement of NMR lipids, lipoproteins, and apolipoproteins 109
3.2.1 Proton NMR spectroscopic method *109*
3.2.2 Sample management and quality control *110*
3.2.3 NMR lipid-related traits *110*
3.2.4 Statistical quality control *111*
3.3 General statistical methods 112
3.3.1 Selection of variables and variable derivation *112*
3.3.2 Selection of participants *113*
3.4 Baseline characteristics 114
Reference 115
CHAPTER 4: Cross-sectional associations of NMR lipid traits with demographic and lifestyle factors 125
4.1 Background 126
4.2 Methods 128
4.2.1 Study design and study population *128*
4.2.2 NMR lipid-related traits *128*
4.2.3 Cross-sectional analyses *129*
4.3 Results 130
4.3.1 Correlations of NMR lipid traits at baseline *130*
4.3.2 Correlations of NMR- and clinical chemistry-measured composite lipid traits *131*
4.3.3 Correlations of NMR-measured composite lipid traits and baseline characteristics *131*
4.4 Discussion 133
4.4.1 Correlations of NMR lipid traits at baseline *133*
4.4.2 Correlations of NMR-measured and clinical chemistry-measured composite lipid traits *134*
4.4.3 Correlations of NMR-measured composite lipid traits and baseline characteristics *134*
4.4.4 Strength and limitations *137*

4.5 Conclusion 138
Reference 153
CHAPTER 5: Genetics of major NMR lipid traits in MCPS: Genome-wide association studies 156
5.1 Background 157
5.2 Methods 160
5.2.1 Genotyping 160
5.2.2 Genetic quality control and exclusions 160
5.2.3 Phenotype definitions 161
5.2.4 Genotype-based principal components 162
5.2.5 Primary association analyses 162
5.2.6 Genomic inflation factor and Covariate-adjusted LD score regression 164
5.2.7 Power calculation 164
5.3 Results 165
5.3.1 Associations in GWAS of LDL-C 165
5.3.2 Associations in GWAS of HDL-C 166
5.3.3 Associations in GWAS of TG 167
5.3.4 Associations in GWAS of ApoA1 168
5.3.5 Associations in GWAS of ApoB 168
5.4 Discussion 169
5.5 Conclusion 172
Reference 201
CHAPTER 6: Observational associations between major lipid traits and atherosclerotic cardiovascular mortality: Prospective cohort study 204
6.1 Background 206
6.2 Methods 207
6.2.1 Study design and study population 207
6.2.2 Definitions of exposure and outcome 208
6.2.3 Statistical analysis 209
6.3 Results 213
6.3.1 Relationship of usual levels of lipid-related traits to ASCVD mortality risk . 213

6.3.2 Relationship of usual levels of other composite lipid traits to ASCVD mortality risk *216*
6.3.3 Secondary analyses and sensitivity analyses *216*
6.4 Discussion 217
6.4.1 Main findings in context *217*
6.4.2 Relevance of ApoB-containing lipoproteins to ASCVD *218*
6.4.3 Relevance of TG to ASCVD *221*
6.4.4 Relevance of HDL-related traits to ASCVD *222*
6.4.5 Heterogeneity of association with ASCVD between sex *224*
6.4.6 Strength and limitations *226*
6.5 Conclusion 228
Reference 242
CHAPTER 7: Genetic associations between major lipid traits and atherosclerotic cardiovascular mortality: Mendelian randomization study 249
7.1 Background 250
7.2 Methods 253
7.2.1 Genotyping *253*
7.2.2 Genetic quality control and exclusions *254*
7.2.3 Phenotype definitions *255*
7.2.4 SNP selection and construction of genetic risk scores for lipid traits *256*
7.2.5 MR analyses *258*
7.2.6 Sensitivity analyses *261*
7.2.7 Power calculations *262*
7.3 Results 263
7.3.1 Associations of genetic risk scores with baseline lipid related traits *263*
7.3.2 One-sample univariable and multivariable MR between lipid traits and ASCVD *265*
7.3.3 Sensitivity analyses using pleiotropy-robust two-sample MR approaches . *266*
7.4 Discussion 266
7.4.1 Main findings in context *266*
7.4.2 Causal relevance of ApoB and LDL-C on ASCVD *267*

7.4.3 Relevance of HDL-C and ApoA1 on ASCVD 268
7.4.4 Relevance of TG on ASCVD 270
7.4.5 Strength and limitations 271
7.5 Conclusion 273
References 296
CHAPTER 8: Discussion 300
8.1 Summary of main findings 301
8.1.1 Complexity of plasma lipidome and its link with participant characteristics 301
8.1.2 Genetic discovery of major lipid traits 301
8.1.3 Association of LDL-C with ASCVD 302
8.1.4 Association of TG with ASCVD 303
8.1.5 Association of HDL-C and ApoA1 with ASCVD 304
8.1.6 Causality of ApoB on ASCVD 305
8.1.7 Overall strengths and limitations 306
8.2 Clinical and public health implications 307
8.2.1 Lipid management and ASCVD control in Mexico 307
8.2.2 ApoB vs LDL-C as the primary measures of risk related to apoB-containing particles 308
8.2.3 Lowering the burden of cardiovascular diseases by targeting apoB-containing particles 311
8.3 Implications for future research 311
8.4 Conclusion 313
Reference 314
Supplementary appendix 319
List of Supplementary Tables and Figures 320

List of Tables

Description
Table 1.1 Classification and characteristics of normally occurring plasma lipoproteins
Table 2.1 Search terms used for literature review of observational or MR studies
Table 2.2 Search terms used for literature review of GWAS of NMR lipid related traits
Table 2.2 Characteristics of included observational studies investigating the association between NMR lipid traits and the risk of ASCVD
Table 2.3 Characteristics of included Mendelian randomization studies investigating the association between NMR lipid traits and the risk of ASCVD
Table 2.4 Characteristics of included GWAS of NMR lipid traits and representative GWAS of traditional lipids
Table 3.1 Information collected at baseline by questionnaire survey
Table 3.2 Description of NMR lipid measures used in the main analyses
Table 3.3 Selected socio-demographic and lifestyle characteristics of participants by sex
Table 3.4 Selected biological and medical characteristics of participants by sex
Table 4.1. NMR lipid traits by sex
Table 4.2. NMR lipid traits by age, region, and education level
Table 4.3. NMR lipid traits by smoking, alcohol intake, fasting duration, and obesity
Table 5.1. Conditionally independent signals from fine-mapped results of the LDL-C GWAS
Table 5.2. Conditionally independent signals from fine-mapped results of the HDL-C GWAS
Table 5.3. Conditionally independent signals from fine-mapped results of the TG GWAS
Table 5.4. Conditionally independent signals from fine-mapped results of the ApoA1 GWAS
Table 5.5. Conditionally independent signals from fine-mapped results of the ApoB GWAS
Table 6.1. Baseline characteristics of 107 504 participants without prior chronic disease (other than diabetes) and aged 35-84 at recruitment by groups of LDL-C
Table 7.1 Power calculation for two-stage least squares mendelian randomization of continuous exposure and binary outcome

List of Figures

Description

Figure 1.1 Initiation and progression of atherosclerosis

Figure 1.2 Lipid metabolism and targets for pharmacotherapy

Figure 3.1 Mexico City Prospective Study recruitment sites and dates, 1998-2004

Figure 3.2 Distribution of major NMR lipid measures before data transformation

Figure 3.3 Flowchart of participant selection

Figure 5.1. Manhattan plot for the GWAS of NMR-measured LDL-C

Figure 5.2. QQ plot for GWAS of NMR-measured LDL-C

Figure 5.3. Manhattan plot for the GWAS of NMR-measured HDL-C

Figure 5.4. QQ plot for GWAS of NMR-measured HDL-C

Figure 5.5. Manhattan plot for the GWAS of NMR-measured TG

Figure 5.6. QQ plot for GWAS of NMR-measured TG

Figure 5.7. Manhattan plot for the GWAS of NMR-measured ApoA1

Figure 5.8. QQ plot for GWAS of NMR-measured ApoA1

Figure 5.9. Manhattan plot for the GWAS of NMR-measured ApoB

Figure 5.10. QQ plot for GWAS of NMR-measured ApoB

Figure 6.1. The association between ApoB and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period

Figure 6.2. The association between LDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period

Figure 6.3. The association between ApoB given LDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period

Figure 6.4. The association between LDL-C given ApoB and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period

Figure 6.5. Age- and sex-specific association of usual LDL-related subfractions and ASCVD mortality during full follow-up period

Figure 6.6. The association between TG and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period

Figure 6.7. The association between TG given ApoB and LDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period

Figure 6.8. The association between ApoA1 and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period

Figure 6.9. The association between HDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period

Figure 6.10. The association between ApoA1 given ApoB and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period

Figure 6.11. The association between HDL-C given ApoB and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period

Figure 6.12. The association between HDL-related subfractions and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period

Figure 7.1 Illustration and assumptions of Mendelian randomization

Figure 7.2 Mean baseline NMR-measured LDL-C by GRS for LDL-C using SNPs from GLGC 2013, GLGC 2017, and GLGC 2021

Figure 7.3 Mean baseline NMR-measured HDL-C by GRS for HDL-C using SNPs from GLGC 2013, GLGC 2017, and GLGC 2021

Figure 7.4 Mean baseline NMR-measured TG by GRS for TG using SNPs from GLGC 2013, GLGC 2017, and GLGC 2021

Figure 7.5 Mean baseline NMR-measured ApoB by GRS for ApoB using SNPs from UKB 2020

Figure 7.6 Mean baseline NMR-measured ApoB by GRS for ApoB using SNPs from UKB 2020

Figure 7.7 Comparison of effect sizes of genetic variants associated with LDL-C in MCPS and GLGC GWAS

Figure 7.8 Comparison of effect sizes of genetic variants associated with HDL-C in MCPS and GLGC GWAS

Figure 7.9 Comparison of effect sizes of genetic variants associated with TG in MCPS and GLGC GWAS

Figure 7.10 Comparison of effect sizes of genetic variants associated with ApoB in MCPS and UKB GWAS

Figure 7.11 Comparison of effect sizes of genetic variants associated with ApoA1 in MCPS and UKB GWAS

Figure 7.12 Mean change in lipid-related traits per 1-SD higher weighted genetic risk scores

Figure 7.13 Cross−sectional associations of GRS−LDL−C and NMR biomarkers at recruitment and the effect of adjustment for confounders

Figure 7.14 Cross−sectional associations of GRS−HDL−C and NMR biomarkers at recruitment and the effect of adjustment for confounders

Figure 7.15 Cross−sectional associations of GRS−TG and NMR biomarkers at recruitment and the effect of adjustment for confounders

Figure 7.16 Cross−sectional associations of GRS−ApoB and NMR biomarkers at recruitment and the effect of adjustment for confounders

Figure 7.17 Cross−sectional associations of GRS−ApoA1 and NMR biomarkers at recruitment and the effect of adjustment for confounders

Figure 7.19 One-sample univariable and multivariable MR of major lipids and apolipoproteins and risk of ASCVD

Figure 7.20 Sensitivity analyses of two-sample MR using pleiotropy-robust methods

CHAPTER 1: Introduction

Chapter summary

Atherosclerotic cardiovascular diseases (ASCVD) are the leading cause of morbidity and premature death worldwide. LDL and other apoB-containing particles play a particularly central role in the pathophysiology of atherosclerosis and incident ASCVD events. Based on the totality of evidence from genetic studies, prospective epidemiologic cohort studies, Mendelian randomization (MR) studies, and randomized trials of LDL-modifying therapies, this chapter systematically introduces the association between major blood lipid-related traits (LDL-C, Lp(a), HDL-C, TG, apoB, and apoA1) and ASCVD. Consistent evidence from numerous and diverse types of clinical and genetic studies supports that LDL causes ASCVD, with an emerging consensus that apoB *per se*, as a proxy for the number of apoB-containing particles, is more accurately and directly related to ASCVD risk than the cholesterol or TG contents within these particles. The clinical benefits of lowering lipid contents, whether TG or cholesterol, should be proportional to the absolute change in apoB. However, direct evidence supporting these findings from large Latin American populations remains limited. The combination of detailed NMR metabolomics and genomic characterization from the Mexico City Prospective Study – a large long-term blood-based prospective study of Mexican adults – provides a unique opportunity to better understand the biological mechanisms of lipid metabolism and their relevance to the pathogenesis of ASCVD in a large Latin American population. The availability of detailed lipid-related measures across lipoprotein subclasses could also provide more information than the composite lipid measures by clinical chemistry assay. The main objectives of this thesis are to investigate the association of NMR-measured plasma lipid traits, including major composite lipids, their subfractions, and apolipoproteins, with the risk of ASCVD mortality using both observational regression analyses and MR analyses to strengthen causal inference.

1.1 Atherosclerotic cardiovascular disease (ASCVD)

1.1.1 Epidemiology of ASCVD

Atherosclerotic cardiovascular disease (ASCVD) and its principal clinical manifestations, including coronary heart disease (CHD) and ischaemic stroke, is the leading cause of morbidity, premature death and rising health costs worldwide.[1,2] The total number of deaths due to cardiovascular disease (CVD) has increased steadily since 1990 and reached 18.6 million worldwide in 2019, corresponding to 380 million years of life lost and 34.4 million years of disability.[2] Although the global age-standardized mortality rate of CVD has declined over the past three decades, the rate is increasing in many low- and middle-income countries, indicating that intensified regional efforts to control potential risk factors are needed to address this persistent global health threat.[2] A variety of modifiable risk factors have been identified to be associated with an increased risk of cardiovascular events, such as high level of blood pressure, fasting glucose, and low density lipoprotein (LDL) cholesterol, obesity, smoking, alcohol consumption, and low physical activity.[3,4] Among these risk factors, LDL plays a particularly central role in the pathophysiology of atherosclerosis and incident ASCVD events.[5,6] In 2019, it was estimated that about two-thirds of global CVD deaths were caused by atherosclerotic conditions in which LDL and other cholesterol-rich apolipoprotein B (apoB)-containing lipoproteins are the primary drivers.[2] Accordingly, the comprehensive assessment of the effect of plasma lipids on ASCVD risk and the management of dyslipidemias, accompanying other public health interventions, are critical to reduce the burgeoning burden of CVD in the world.

1.1.2 Pathophysiology of atherosclerosis

Atherosclerosis is a lipid-driven inflammatory process characterized by the accumulation of lipid and/or fibrous components and the formation of atherosclerotic plaque in the arterial intima at predisposed sites, such as curvatures and branch points (**Figure 1.1**).[7] Atherosclerotic plaques consist of lipids, (immune) cells, and cellular debris covered by

a fibrous cap. The pathogenesis of atherosclerosis generally includes three phases: initiation, progression and complications. The key events in the initiation of atherosclerosis are activation, dysfunction, and structural alterations of endothelial cells leading to subendothelial retention and accumulation of LDL and other cholesterol-rich apoB-containing lipoproteins <70 nm in diameter, which then trigger a number of events that propagate lesion development.[8,9] Retained atherogenic lipoprotein particles such as LDL are susceptible to oxidation (e.g. by reactive oxygen species) and other modifications (e.g. by myeloperoxidase), leading to the generation of oxidized LDL that is pro-inflammatory and immunogenic and trigger the recruitment of monocytes.[10] Recruited monocytes differentiate into macrophages and can further promote the oxidation of LDL, which are then recognized and internalized by scavenger receptors resulting in the formation of cholesterol-laden foam cells.[11] Moreover, smooth muscle cells (SMC) may also take up cholesterol-rich lipoproteins and contribute to the burden of foam cells in the lesions.[12] As the lesion evolves, macrophages and SMCs can undergo cell death by apoptosis or necrosis. The accumulation of debris from dead and dying cells and impaired function of efferocytosis (clearance of apoptotic cells) under the environment of chronic inflammation can contribute to the formation of a necrotic core containing cellular debris and cholesterol crystals.[5] Depending on the composition and architecture of plaque, segments of the atherosclerotic artery may remain quiescent, exhibit chronic stenosis, or precipitate thrombosis that can cause occlusion of the vessel and lead to acute ischaemic events. The mechanisms of thrombotic complications of atheroma typically include plaque rupture and plaque erosion. Atherosclerotic plaques with a thin fibrous cap, large necrotic core, pronounced inflammation, and low SMC density, are at risk of plaque rupture. During the evolution, the presence of spotty pattern of calcification in plaques is also considered to increase the tendency to rupture.[13] Plaque rupture accounts for the majority of acute thrombosis of coronary arteries that cause myocardial infarction (MI).[14,15] However, lesions with lower lipid cores or with thick fibrous caps can still precipitate thrombus in response to plaque erosion, and are characterized

by the subendothelial accumulation of hyaluronan and neutrophil.[15,16] LDL-lowering therapies such as statins mitigate the mechanisms of plaque rupture (e.g. formation of lipid core and LDL induced inflammation), but are less effective in reducing the risk of thrombosis caused by plaque erosion.[17] Besides LDL and other atherogenic lipoprotein particles, other risk factors for atherosclerosis and its thrombotic complications include smoking, diabetes mellitus, hypertension and other systemic prothrombotic factors.[5,7]

1.2 Lipids, lipoproteins, apolipoproteins and ASCVD

1.2.1 Components and metabolism of lipoproteins

Lipoproteins are lipid-protein complexes involved in the transportation and metabolism of lipids in the body.[18] They are spherical particles consisting of a hydrophobic core of triglycerides (TG) and cholesterol (cholesteryl esters, CE) surrounded by a hydrophilic monolayer of free cholesterol (FC), phospholipids, and apolipoproteins. The apolipoproteins function in maintaining the structure of lipoproteins, and also serve as ligands for specific receptors or as enzyme activators. Lipoproteins are generally divided into classes primarily based on hydrated density (measured by ultracentrifugation) or electrophoretic mobility (**Table 1**). The function and composition of lipoproteins (characterized by the lipid content and different apolipoproteins) vary by such classifications. The density of lipoprotein particles is inversely associated with their size, reflecting differences in the relative amount of low-density lipid core and high-density surface protein. The two largest and most buoyant lipoprotein are the triglyceride-rich chylomicrons, and very low-density lipoprotein (VLDL) that contain one single apoB molecule when secreted into plasma by the liver. Denser and more cholesterol-rich lipoproteins are intermediate density lipoproteins (IDL) and LDL. High-density lipoproteins (HDL_2 and HDL_3) are denser and richer in phospholipids and proteins, such as apoA1, lecithin–cholesterol acyltransferase (LCAT) that esterifies plasma free cholesterol (FC), and cholesterol ester transfer protein (CETP) that mediates the transfer of CE from HDL to LDL and VLDL for the exchange of TG (reverse cholesterol transport).

Lipoprotein (a) [Lp(a)] is larger but denser than LDL and consists of an LDL particle with apo(a) bound to the apoB-100 molecule. These denser and smaller lipoprotein classes including VLDL remnants, IDL, LDL, and Lp(a), referred to as apoB-containing particles, represent products formed during the lipolysis and of triglyceride-rich lipoproteins (TGRLs) by the enzyme lipoprotein lipase (LPL), hepatic lipase (HL), and CETP. ApoB-containing lipoproteins less than 70 nm in diameter can flux across the endothelium and can be trapped resulting in the initiation and progression of atherosclerotic plaque.

In most fasting individuals, LDL particles comprise ~90% of the apoB-containing lipoproteins in the circulation. However, in routine clinical practice, rather than LDL, it is the total amount of cholesterol contained in LDL (LDL-C), which is measured to estimate the plasma LDL level. As a result, LDL-C has been the standard lipid measure and focus for CVD risk assessment and for evaluating the benefit of lipid-lowering treatment in randomized clinical trials.[6]

1.2.2 LDL receptor and lipid-lowering therapies

Lipid metabolism and targets for pharmacotherapy are illustrated in **Figure 1.2**. The LDL receptor (LDLR) is a cell surface receptor capable of binding and internalizing lipoproteins containing apoB-100 or apoE. Most LDL are cleared from circulation by liver through LDLR pathway and reduced level and function of LDLR could lead to hypercholesterolaemia. The expression of the LDL receptor is closely entrained to cellular demand for cholesterol.[19] Current recommend lipid-lowering therapies such as statins, ezetimibe and proprotein convertase subtilisin/kexin type 9 (PCSK9) inhibitors which have been shown to reduce the ASCVD risk in randomized trials,[20-22] reduce the number of circulating LDL particles through upregulation of the LDLR and LDL clearance. The most widely used drug class for lowering cholesterol levels has been statins, which inhibit HMG-CoA reductase, the rate-limiting enzyme in cholesterol biosynthesis, and reduce intrahepatic cholesterol levels causing increased expression of LDLR and enhancing the uptake of exogenous cholesterol to the liver. Ezetimibe could also

upregulate hepatic LDLR by reducing intestinal absorption of cholesterol, which is a adjunctive medication to statin therapy because of the complementary mechanism of action. The protein PCSK9 also plays a major role in cholesterol homeostasis by inducing LDLR degradation. Therefore, PCSK9 inhibitors have been developed as a novel therapy to increase LDLR density and decrease plasma cholesterol level. In contrast, therapies that reduce cholesterol level by mechanisms other than reducing circulating LDL particles through upregulation of the LDLR or that mainly reduce plasma triglyceride levels, are not recommended by current guidelines because of uncertain clinical benefit.[20]

1.2.3 Causal effect of LDL on ASCVD

Extensive and consistent evidence from genetic studies, prospective cohort studies, Mendelian randomization studies, and randomized clinical trials of LDL-lowering therapies have established that LDL have a causal effect on ASCVD.[6] As previously mentioned, retention and accumulation of apoB-containing lipoproteins are the necessary step for the initiation and progression of atherosclerotic plaque.[5] As the concentration of LDL-C increases above physiologically required levels (~0.5-1.0 mmol/L),[6] the probability of intimal retention of LDL increases in a dose-dependent manner.[23] Familial hypercholesterolaemia (FH), caused by loss-of-function (LOF) mutations in the *LDLR* or *APOB* genes or gain-of-function (GOF) mutations in the *PCSK9* gene leading to decreased function and number of LDLR, results in a markedly increased level of LDL particles (and cholesterol within) and premature atherosclerosis.[24,25] Evidence from genetic studies of FH indicate that the extent of atherosclerosis and the risk of ASCVD in FH patients is proportional to the absolute magnitude and the duration of exposure to increased circulating LDL-C levels.[26,27] Evidence from large meta-analyses of prospective cohort studies have reported a consistent association between the absolute magnitude of exposure to LDL-C and the risk of ASCVD. The Emerging Risk Factors Collaboration reported an individual participant data meta-analysis including 165,544 participants without baseline CVD from 37 cohorts with 15,126 incident

fatal and nonfatal CVDs recorded during median follow-up of 10.4 years. In this study, plasma LDL-C was log-linearly associated with increased risk of non-fatal MI or CHD death.[28] Consistently, the Prospective Studies Collaboration reported a meta-analysis of 892,337 participants without baseline CVD from 61 cohorts with more than 55,000 vascular deaths (including 33,744 CHD deaths) and found a graded log-linear association between plasma total cholesterol and the risk of CHD death.[29]

Although observational studies are vulnerable to reverse causality, confounding and other sources of bias, strong associations reported in these cohort studies were confirmed by both Mendelian randomization (MR) studies (naturally randomized genetic studies) and randomized clinical trials that provide the most compelling evidence for causality.[6,22] A meta-analysis of individual participants from 26 randomised trials of statin therapy including almost 170,000 individuals demonstrated that lowering LDL-C reduces the risk of CVD events proportionate to the absolute reduction in LDL-C, and independently of other risk factors (22% reduction in CVD risk per millimole per litre reduction in LDL-C over a median treatment period of 5 years).[21,30] Importantly, in a meta-analysis comparing different LDL-C lowering therapies, both statin and non-statin therapies (such as ezetimibe, bile acid sequestrants, and ileal bypass) were associated with similar RRs of vascular events per unit lower LDL-C,[22] indicating that the effect of LDL-C on ASCVD risk reduction is approximately identical per unit change in LDL-C levels for any mechanism that acts via upregulation of LDLR. These findings were confirmed by MR studies that have demonstrated that the effects of genetic variants (e.g. *HMGCR*, *NPC1L1, PCSK9, LDLR*) encoding for current LDL-C lowering therapies on the risk of cardiovascular events were approximately the same after adjusting for reduction in LDL-C.[31,32] MR studies have also confirmed a dose-dependent and log-linear causal relationship between reduction in LDL-C and lifetime risk of ASCVD independent of the mechanism by which LDL-C is lowered.[33-35] However, the effect of long term exposure to lower LDL-C on the risk reduction of ASCVD per unit change in LDL-C as estimated in MR studies is substantially greater than the effect of lowering LDL-C beginning later

in life after atherosclerotic plaques have developed in statin trials, indicating that the risk of ASCVD is determined not only by the absolute magnitude of exposure to LDL-C but also the cumulative duration of exposure to LDL-C.[33] Altogether, consistent evidence from clinical and genetic studies that used LDL-C as surrogate for LDL particle concentration establishes that LDL particles cause ASCVD.

1.2.4 Variability in apoB-containing lipoproteins, apoB, and ASCVD

In most cases, LDL-C concentration, LDL particle number, and apoB level are highly correlated and provide similar information about cardiovascular risk.[6,36,37] Therefore, plasma LDL-C is widely used as an acceptable surrogate for LDL and apoB and the majority of studies have used LDL-C as a target to evaluate the clinical benefit of lipid-lowering therapies in trials and to predict the ASCVD risk in observational studies. However, increasing evidence from clinical and genetic studies suggests that apoB is a more accurate marker of cardiovascular risk attributable to the atherogenic apoB-containing lipoproteins than LDL-C or non-HDL-C.[37] In addition, the clinical benefit of lowering LDL-C may be determined by the absolute reduction in apoB-containing particles especially for treatments that lower LDL-C level without proportionally reducing apoB concentration (e.g. combination use of CETP inhibitors and statins).[38,39] The proposed superiority of apoB over cholesterol as a marker arises from the variability in apoB-containing lipoproteins and the observed discordance between changes in LDL-C and apoB levels. This is particularly important in cases where the reduction in ASCVD risk is more strongly associated with apoB levels than with LDL-C levels.

ApoB-containing lipoproteins consist of a spectrum of particle subclasses with heterogeneous composition and size, as well different metabolic and functional characteristics.[40-43] Depending on hepatic TG availability, the liver may secrete a range of lipoproteins varying in size from large TG-enriched VLDL1, smaller medium-sized VLDL2 that contain less TG, to IDL/LDL.[44] The profile of VLDL particles and their concentration, as well as the activity of LPL and CETP, and the binding affinity of LDL to

LDLR, principally determine the profile of LDL particle subclasses and their lipid composition. For example, CETP-mediated transfer of CE from HDL and transfer of TG from VLDL and their remnants largely determine variability in the lipid composition of LDL. Exchanges of core CE in LDL for TG in VLDL1 and VLDL2 increase TG content of LDL and decreased CE content. Subsequent hydrolysis of TG within LDL particles by HL activity shrinks LDL to smaller, denser and CE-depleted LDL particles (LDL-III or/and even LDL-IV in individuals with increased TG levels). In certain conditions where TG levels exceed 150 mg/dL (e.g. the metabolic syndrome, Type 2 diabetes, and hypertriglyceridaemia), smaller and denser cholesterol-depleted LDL particles predominate in LDL subclasses leading to the discordance between LDL-C, apoB and LDL particle concentration, and therefore LDL-C may not accurately reflect LDL particle number or its effect on ASCVD.[45]

Most studies using traditional analytical methods such as the C statistic to test whether a marker improves risk prediction suggest that apoB is better than LDL-C in a risk model,[37] with some studies reporting non-HDL-C and apoB to be equivalent predictors (such as in the Emerging Risk Factor Study[28] and Copenhagen Heart Study).[46] Conventional epidemiological analyses have also yielded mixed results regarding whether LDL-C, non-HDL-C, or apoB is the best marker of therapeutic effectiveness.[47-49] Additional studies using discordance analysis have consistently demonstrated that apoB is a more accurate marker of ASCVD risk than LDL-C and non-HDL-C. In the UK Biobank Study and the Women's Health Study, discordance analysis but not conventional analysis demonstrated that apoB was more strongly associated with CVD than LDL-C or non-HDL-C.[36,50,51] Discordant groups defined by multiple methods suggested that 20% to 60% of individuals have discordant apoB and LDL-C.[37]

MR analysis further adds to the evidence identified by discordance analysis that circulating number of apoB particles is the most direct and accurate measure of atherogenic injury attributable to the apoB-containing particles.[39] In the ACCELERATE trial, combination therapy of a CETP inhibitor with a statin caused large decrease in LDL-

C but did not result in significant clinical benefit (hazard ratio, 1.01; 95% confidence interval, 0.91-1.11; for the composite outcome of cardiovascular death, stroke, and myocardial infarction),[38] which seems inconsistent with the results that the clinical benefit of lipid-lowering therapies acting through upregulation of LDLR is proportional to the absolute reduction in LDL-C level. MR analyses using variants in the *CETP* and *HMGCR* genes to mimic the effects of CETP inhibitors and statins found that combination use was associated with discordant reductions in LDL-C and apoB concentration (with the expected reduction in LDL-C but a lower reduction in apoB level) and a risk of cardiovascular events that was proportional to the reduction in apoB but significantly less than that expected per unit change in LDL-C.[39] These findings indicate that therapies that alter the lipid content contained in apoB particles without necessarily lowering the number of those particles, may have decreased clinical benefit on ASCVD risk reduction.

Taken together, apoB-containing particles are atherogenic and the risk of subintimal retention and accumulation of these particles leading to atherosclerosis is more closely associated with the number (concentration) of circulating apoB particles than with the variable mass of cholesterol within them.

1.2.5 HDL, HDL-targeted therapies and ASCVD

Current evidence from epidemiological, genetic and clinical studies on the relevance of HDL particles and their cholesterol content (HDL-C) in the development of ASCVD has changed dramatically over recent years and remains inconsistent. Several observational studies report that plasma HDL-C level is inversely associated with the risk of ASCVD independent of LDL-C.[52-55] In the Framingham Heart Study,[54] 1mg/dL higher HDL was associated with about a 2% to 4% decrease in coronary artery disease risk. However, the highest levels of HDL-C reported in most of these studies were in the range of 80 to 90 mg/dL. Studies exploring larger range of HDL-C concentration have indicated a U-shaped relationship between HDL-C, ASCVD and all-cause mortality, with an increased CVD risk at high level of HDL-C.[56-59]

The inverse association between HDL-C and ASCVD observed in observational studies has prompted the hypothesis that increasing plasma HDL-C can prevent ASCVD and has provided the rationale for trials of interventions to increase plasma HDL-C levels. However, despite increasing HDL-C levels, these drugs tested in trials did not reduce the risk of ASCVD events, especially when used jointly with a background of statin therapy. The major pharmacological agents that can lower TG level and increase HDL-C level include niacin and fibrates, while CETP inhibitors can profoundly increase HDL-C level and slightly decrease LDL-C level. The addition of niacin to a background statin therapy in patients with ASCVD did not reduce the risk of major vascular events,[60,61] all-cause mortality, and ASCVD-specific mortality.[62] Fibrates also failed to produce clinical benefits. Although a subgroup of patients with hypertriglyceridemia (> 200 mg/dL) and low HDL-C (< 35 mg/dL) seemed to have a lower risk of ASCVD events if given a fibrate, the clinical benefit of fibrates in this subgroup was largely explained by the reduction in apoB containing particles rather than the change in HDL-C.[63] Fibrates failed to show benefit in all patient groups partly because VLDL apoB lowered by fibrates only account for a relatively small amount of apoB particles (LDL apoB makes up most of apoB in circulation).[37] However, in patients with elevated levels of TG, apoB in VLDL may rise to about 30% of total apoB and the reduction in apoB by fibrates could reach clinical significance.[37] In a meta-analysis investigating the impact of raising HDL-C on cardiovascular endpoints including 108 randomized clinical trials and about 300,000 individuals on statins, fibrates, niacin, and combinations of lipid-modifying drugs, no significant association was found between elevation in HDL-C and risk of non-fatal and fatal ASCVD, or all-cause mortality after adjusting for change in LDL-C.[64] CETP inhibitors also failed to show clinical benefit although they greatly increased the level of HDL-C by about 100%. Their effect of ASCVD events were either non-significant, negligible,[38,65-67] or small and occurred by a reduction in non-HDL-C.[68,69] A meta-analysis of randomized controlled trials investigating the effects of CETP inhibitors on ASCVD and all-cause mortality showed a non-significant reduction in the risk of non-fatal MI (7%) and ASCVD-

specific death (8%).[70] Additionally, evidence from genetic studies showed that genetically reduced HDL-C is not associated with an increased risk of atherosclerotic cardiovascular events.[39,71-74]

The above mentioned evidence challenges the traditional notion that HDL is protective against CVD risk through reverse cholesterol transport, a metabolic process where HDL removes cholesterol from macrophages in the arterial wall and delivers it to the liver for excretion in bile. Given this, the focus of research has shifted to the biological function and composition of HDL, especially in the context of macrophage cholesterol efflux (MCE).[62] MCE is the efflux of FC to plasma lipoproteins, mainly HDL, which was considered as a crucial determinant of atherogenesis and concomitant inflammation in the model of reverse cholesterol transport.[75] Previous studies have reported inconsistent findings for the associations between MCE and ASCVD, with several large studies reporting a positive correlation.[76-79] However, studies have found that FC in HDL is rapidly cleared from circulation within minutes independently of esterification, which casts doubt on this mechanism of cholesterol transport. Moreover, the reversibility of MCE has indicated that high HDL-FC may reversely transfer to macrophages and cause atherosclerosis.[62] Studies also suggest that structural composition of HDL subclasses, each perhaps endowed with particular functionality, rather their cholesterol content and involvement in MCE, may reflect the role of this lipoprotein class in pathophysiological functions.[80]

1.2.6 TG, LPL, and ASCVD

The association between plasma TG levels and the risk of ASCVD has been controversial.[81,82] Although some observational and genetic data support a causal role of TG within TGRLs in the development of ASCVD, this association often becomes nonsignificant after adjustment for HDL-C and non-HDL-C. In certain insulin-resistant conditions (e.g. obesity, metabolic syndrome, and type 2 diabetes), hypertriglyceridemia (defined as a fasting serum TG level greater than 150 mg/dL) is frequently associated

with a low HDL-C, and the prevalence of small and dense cholesterol-depleted LDL particles, which is referred to as atherogenic dyslipidemia.[81] This atherogenic set of interrelated lipoprotein change may jointly result in an increased risk of ASCVD,[83,84] making it difficult to disentangle the relative effect of TG.

Two earlier meta-analyses of prospective cohort studies similarly found a univariate association of TG with CVD.[85,86] However, a meta-analysis from the Emerging Risk Factors Collaboration including about 300,000 individuals without CVD at baseline suggested that the positive association between fasting and nonfasting TG levels and ASCVD (hazard ratio, 1.37; 95% confidence interval, 1.31-1.42; for CHD) was not significant after adjusting for HDL-C and non-HDL-C (0.99, 0.94-1.05).[55] Similarly, additional data from trials investigating the on-treatment TG levels and the risk of ASCVD demonstrated that increased on-treatment TG levels were significantly associated with the risk of CVD events in statin-treated patients,[87,88] although the association was attenuated after adjusting for HDL-C and apoB-to-apoAI ratio in one trial.[88] Taken together, although TG levels appear to provide information about vascular risk, the independence of TG as a causal factor in the development of ASCVD remains controversial based on the evidence from traditional epidemiological studies.

Randomized trials also have not provided consistent evidence that lowering TG levels reduces the risk of cardiovascular events and guidelines do not recommend treatment to lower TG levels.[20] Current available TG-lowering therapies mainly include fibrates, niacin, and omega-3 fatty acids (OM3FAs). Statins also reduce TG levels in individuals with baseline hypertriglyceridemia.[89] As mentioned in ***1.2.5***, in three relevant trials of lipid-lowering therapies, the addition of fibrates and niacin to statin therapy did not result in significant clinical benefit on ASCVD risk, although the plasma TG levels were reduced by 30% to 50%.[60,61,90] However, a subgroup of patients with hypertriglyceridemia (baseline TG > 200 mg/dL) and low HDL-C (< 35 mg/dL) who were on treatment had lower risk of ASCVD events, suggesting the benefit of fibrates and niacin in those with hypertriglyceridemia.[84,90-92] Notably, the baseline TG levels of most individuals in these

trials were lower than 200 mg/dL. Thus, the mean reduction in TG levels in these trials was only about 20 mg/dL to 50 mg/dL, with the corresponding modest reduction in apoB-containing particles that could not reach clinical significance within a short term trial, which appears to explain why these trials evaluating fibrates and niacin failed to consistently suggest a reduced risk of ASCVD. Similar to the trials of fibrates and niacin, studies of OM3FAs have reported inconsistent results and the clinical benefit was significant only in patients with TG levels > 150 mg/dL.[93-95]

Genetic studies investigating the variants in genes involved in TGRLs metabolism, namely *LPL* and the inhibitors of LPL including *APOC3*, *ANGPTL3*, and *ANGPTL4*, provide further insights into the association between TG and the ASCVD. LOF mutations in the *LPL* gene are associated with both higher levels of TG and a higher risk of CVD; while LOF mutations in the *APOC3*, *ANGPTL3*, and *ANGPTL4* genes that encode LPL inhibitors are associated with lower TG levels and a lower risk of CVD.[96-101] These observations suggest that TG may play a role in mitigating the risk of ASCVD and have motivated the development of several therapeutic approaches that potently reduce TG levels.[102,103] MR studies using variants associated with TG levels have suggested a potential causal effect of TG on ASCVD. To overcome the pleiotropy of variants associated with plasma lipids, one multivariable MR study using variants associated with plasma TG, LDL-C, or HDL-C levels found that genetically elevated TG levels significantly correlated with CVD after adjustment for the pleiotropic effect on LDL-C and HDL-C.[104] Other MR studies have demonstrated a similar independent causal relationship of TG with ASCVD and all-cause mortality.[34,35,105,106] Notably, recent MR analyses have further demonstrated that TG-lowering *LPL* variants and LDL-C-lowering *LDLR* variants were associated with similar lower risk of ASCVD per unit change in apoB-containing lipoproteins and the effect of TG and LDL-C determined by these variants were independent and additive, suggesting the potential clinical benefit of lowering TG levels is similar to the benefit of lowering LDL-C per unit change in apoB.[107] Combined together, the results of these MR analyses (as discussed in ***1.2.4*** and ***1.2.6***) suggest that

the causal effect of apoB-containing particles on ASCVD is determined by their concentration in circulation (particle number of apoB) rather than by the cholesterol and TG content within them. Therefore, to reduce the risk of CVD risk by 20% (as can be achieved by lowering LDL-C by 40 mg/dL)[21] it may be expected that TG levels would need to be reduced by 5 times this quantity (the concentration of TGRLs can be estimated by dividing TG concentration by 5 on the mg/dL scale), about 200 mg/dL, to achieve the same reduction in apoB-containing particles.[107] By contrast, the absolute reduction in TG levels in the above mentioned TG-lowering trials was about 20 mg/dL to 50 mg/dL. This may explain the observation that the clinical benefit seemed only to be shown in the subgroup of patients with very high baseline TG. Altogether, the clinical benefit of lowering triglyceride may be proportional to the absolute change in apoB.

1.2.7 Lp(a), LPA, and ASCVD

Lp(a) consists of an LDL-like particle and an additional apolipoprotein(a) [apo(a)] attached to apoB via a disulfide bridge. Plasma Lp(a) levels are mainly determined by the *LPA* gene (~90%).[108] The variation in size of Lp(a) is caused by copy number variations in the *LPA* gene that determines the number of kringle IV repeats, resulting in different isoforms.[108] Like LDL and other apoB-containing particles, Lp(a) has a pathogenic role in development of ASCVD.[109] In addition, apo(a) has homology with plasminogen and could result in a decrease of plasmin synthesis and inhibition of fibrinolysis, leading to thrombosis formation at susceptible sites.[110]

Evidence from observational and genetic studies has strongly suggested an association between plasma concentrations of Lp(a) and increased risk of ASCVD and aortic valve stenosis (AVS).[108,111,112] An individual-level meta-analysis from the Emerging Risk Factors Collaboration including about 126,000 individuals without ASCVD at baseline suggested an independent and modest association of Lp(a) concentration with risk of CHD and stroke after adjustment for lipids and other traditional cardiovascular risk factors.[113] Other meta-analyses of prospective studies consistently showed increased

risk of CHD and stroke at Lp(a) concentrations above 50 mg/dL (about 100 nmol/L).[114,115] A later meta-analysis of 30 studies applied comparable phenotyping and analytic methods to assess the association of apo(a) isoforms with ASCVD risk, and demonstrated that smaller apo(a) isoforms were associated with about a doubling in the risk of CHD and ischaemic stroke, which is relevant from public health perspectives given that approximately 25% to 35% of the population carry small apo(a) isoforms.[116] Moreover, in the recent INTERHEART study of 6086 cases of first MI and 6857 controls stratified by ethnicity and adjusted for age and sex, high Lp(a) concentrations (>50 mg/dL) were associated with an increased risk of MI, independent of established ASCVD risk factors. Although isoform size was inversely associated with Lp(a) concentration and larger isoforms appeared to be associated with a lower risk of MI, this association between isoform and MI was not present after adjustment for Lp(a) concentration.[117] A recent meta-analysis of patients with elevated Lp(a) levels and history of cardiovascular events showed that those with Lp(a) levels above 50 mg/dL on statin therapy were at an increased risk of ASCVD compared those with Lp(a) levels lower than 30 mg/dL, independent of traditional risk factors of CVD.[118] Taken together the primary and secondary prevention studies using the more reliable isoform-insensitive assays supported the view that Lp(a) is an independent risk factor for ASCVD.

Furthermore, evidence from MR studies have suggested that the association between Lp(a) and risk of ASCVD is likely to be causal.[119,120] However, randomized clinical trials of lipid-lowering therapies that reduce Lp(a) concentration by 25% to 35% (such as niacin, CETP inhibitors, and PCSK9 inhibitors) have not provide any evidence that lowering Lp(a) level reduces the risk of ASCVD events.[38,60,61,65,121] Although mediated by the LPLR pathway, statins have a slight Lp(a)-increasing effect rather than a lowering effect.[122] However, the statin trials were not designed to evaluate the Lp(a) lowering effect of these therapies and did not specifically include patients with elevated Lp(a) levels. It has been suggested that patients with higher baseline Lp(a) levels had greater absolute reductions in Lp(a) (> 50 mg/dL) and were likely to derive greater clinical benefit from PCSK9

inhibitors.[123] Studies of other emerging therapies, such as lipid apheresis and antisense therapy, have already reported significant Lp(a) reductions of 35% to 80%, however, limited data of CVD outcomes are available from these studies.[108] In a recent MR study comparing the association of Lp(a)-lowering variants in the *LPA* gene and LDL-C-lowering variants in gene encoding (the targets of currently available therapies with ASCVD) the clinical benefit of lowering Lp(a) was found to be proportional to the absolute reduction in Lp(a) level, apparently independently of changes in LDL-C levels (odds ratio [OR] for a 10 mg/dL genetically determined lower Lp(a), 0.94; 95% CI, 0.93-0.95; OR for a 10 mg/dL genetically determined lower LDL-C, 0.85; 95% CI, 0.82-0.89). Therefore, a large absolute reduction in Lp(a) of approximately 100 mg/dL may be required to achieve equivalent 20% to 25% reduction in ASCVD risk per millimole per liter (38.67 mg/dL) reduction in LDL-C in short-term trial.[124] However, trials of lipid-lowering therapies have tended to result in Lp(a) reductions of only about 3 to 6 mg/dL. Current guidelines recommended that Lp(a) screening should be considered in individuals at intermediate or high risk of CVD, and that Lp(a) > 50 mg/dL (>80th percentile) is considered as an additional risk factor which should be taken into account despite aggressive LDL cholesterol reduction with a statin.[20,112] Finally, randomized controlled trials with selective and potent reductions in plasma Lp(a) levels (i.e., that can lower Lp(a) levels by up to 90%[125]) in primary and secondary prevention settings are needed to confirm or refute the clinical benefit of Lp(a) lowering on ASCVD risk.

1.3 Study setting

The Mexico City Prospective Study (MCPS) is a blood-based prospective cohort study designed to investigate the associations of risk factors with common causes of death including cardiac, cerebrovascular, metabolic and related diseases, as detailed elsewhere.[126,127] Briefly, from 1998 through 2004, households within two districts of Mexico City were systematically visited and about 150,000 adults (100,000 women and 50,000 men) aged 35 years and older were recruited. Demographic characteristics (e.g. age, sex and socioeconomic status), lifestyle factors (e.g. drinking, smoking, and

physical activity), medical comorbidity, medical history, and current medication use were recorded by questionnaire survey. Physical parameters (e.g. height, weight, waist and hip circumferences, and seated blood pressure) were measured and almost all participants provided a non-fasting 10 mL blood sample (from which the plasma and buffy coat were extracted and stored long-term). Participants are followed for cause-specific mortality through electronic probabilistic linkage to the death registry with deaths coded according to the *International statistical classification of diseases and related health problems, 10th revision* (ICD-10).

DNA was extracted from the buffy coat samples and all participants passing basic quality control were genotyped and subsequently imputed to collect genetic characteristics. From the stored baseline plasma sample, the Nightingale Health NMR metabolomics platform was used to quantify more than 225 lipid and other metabolic measures. The combination of detailed metabolomics and genomics characterization provides an unparalleled opportunity to understand the biological mechanisms regulating lipoprotein metabolism and their relevance in the pathogenesis of ASCVD in this large Latin American population. In addition, the extensive NMR measures of lipoprotein particle subclasses and lipid-related traits that could be more informative than conventional lipid measures allows for the identification of new biomarkers of ASCVD by understanding the comparative role of different lipoprotein subclass in cardiovascular events. The use of genetic approaches will be helpful to clarify whether the identified observational associations between lipid traits and ASCVD mortality are causal. Overall, this study is the largest study investigating the association between the genotype and phenotype of NMR lipid-related markers and risk of ASCVD death in a Latin American population.

1.4 Thesis objectives

The main objectives of this thesis are to investigate the association of NMR-measured plasma lipid traits, including the particle size and concentration of lipoprotein subclasses and their associated lipids and apolipoproteins, with the risk of cardiovascular-specific

mortality using both observational regression analyses and MR analyses to strengthen causal inference. Lipid traits used in these analyses will include both traditional lipid panels used in routine clinical practice (LDL-C, HDL-C, and TG) and detailed lipid-related measurements in lipoprotein subclasses categorized by particle size, and the observational analyses will take into account potential confounding and reverse causation, to which MR analyses are less susceptible (providing certain assumptions are met). Observational estimates will be compared with those from MR analyses to make inferences about the likely causal effects of lipid-related traits on cardiovascular mortality risk. Aims of the following chapters include:

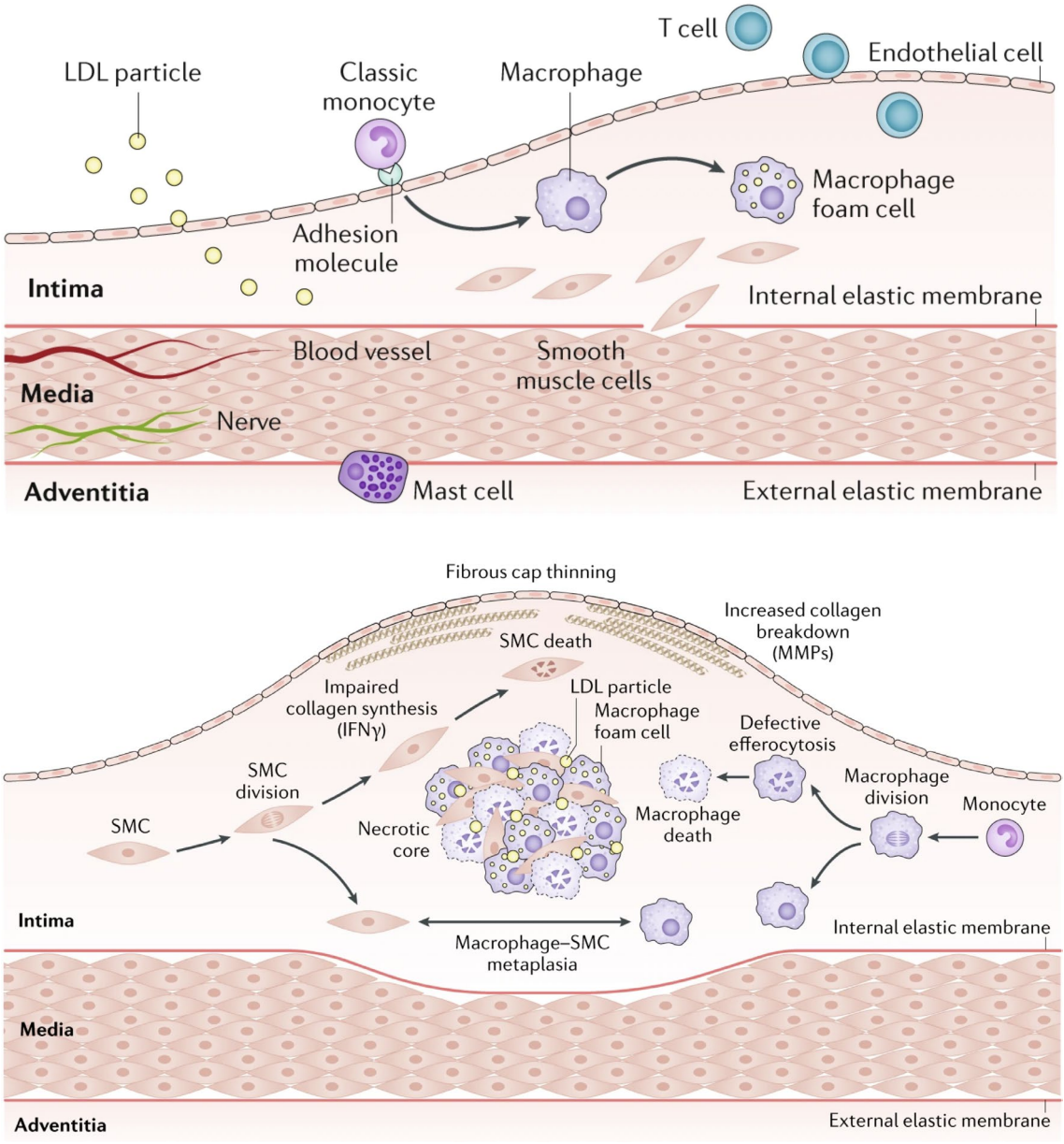
- Summarizing the current epidemiological evidence from observational and Mendelian randomization studies of the association between NMR lipid traits and ASCVD risk, and reviewing the Genome-Wide Association Studies (GWAS) of NMR lipid traits (**Chapter 2**)
- Describing the design of the Mexico City Prospective Study and statistical methods used in this study (**Chapter 3**)
- Describing the inter-correlation of the various NMR lipid traits, the correlation of lipid-related traits measured by NMR spectroscopy and standard clinical chemistry, and the cross-sectional association between NMR lipid traits and baseline characteristics of participants, to compare lipid measures within and across assays and to identify potential confounders (**Chapter 4**)
- Describing the design of GWAS on NMR lipid traits and conducting a GWAS to identify genetic variants associated with the levels of plasma lipids (**Chapter 5**)
- Investigating the prospective associations between NMR lipid traits and the risk of ASCVD after taking into account confounding and reverse causality and assessing whether these association differ by demographic, physical, and lifestyle characteristics (**Chapter 6**)

- Evaluating whether a lifelong genetic predisposition to elevations in NMR lipid traits has a causal effect on the risk of cardiovascular-specific mortality using genetic variants robustly associated with lipid metabolism; Comparing the conventional observational associations between NMR lipid traits and cardiovascular-specific mortality with those from MR analyses to make inferences about the likely causal effects of lipid traits on ASCVD risk (**Chapter 7**)
- Summarizing the key findings and discussing the strengths and limitations of the current study, and the relevance for clinical practice, population health, and future study (**Chapter 8**)

Table 1.1 Classification and characteristics of normally occurring plasma lipoproteins

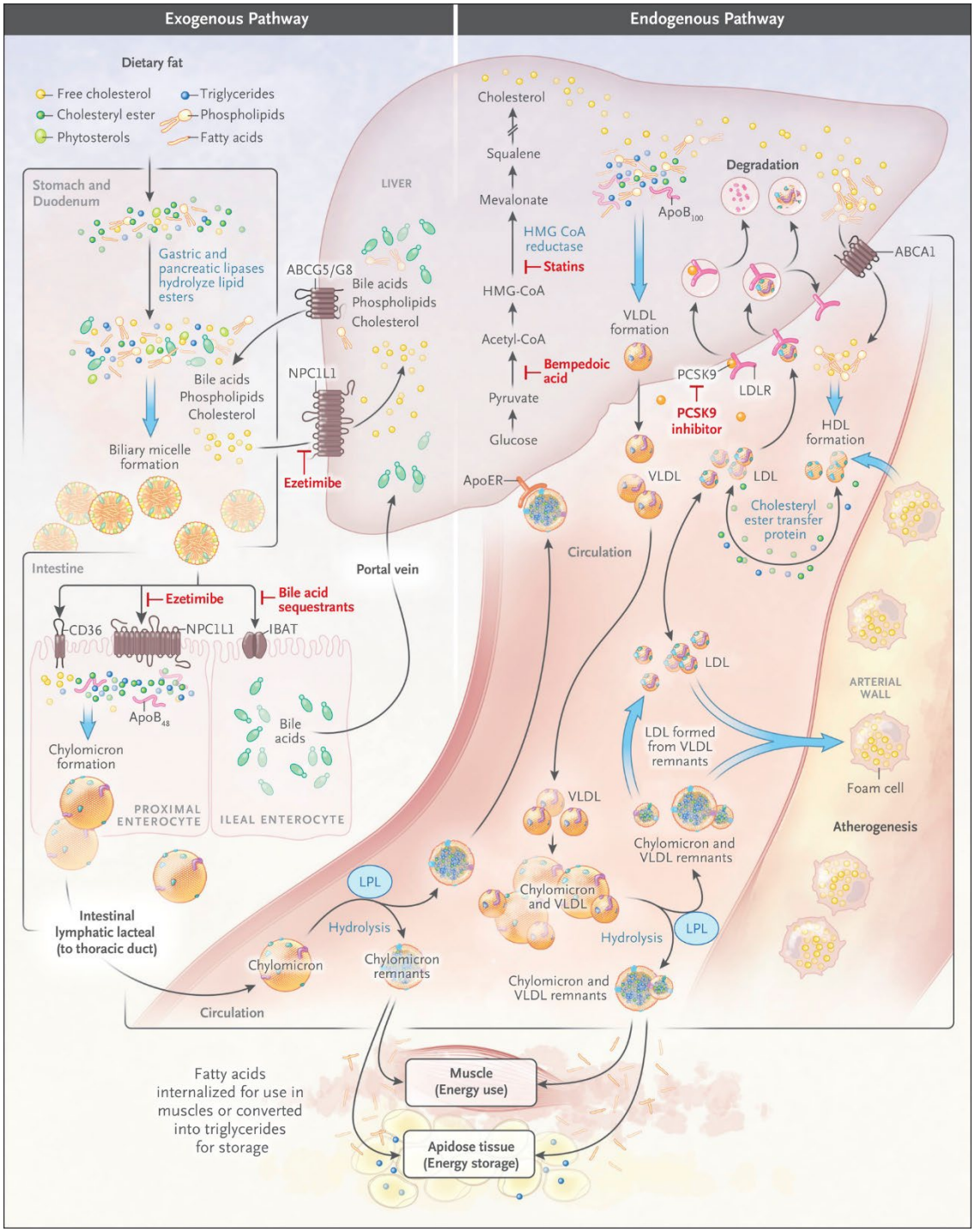
Lipo-protein	Density (g/ml)	Electro-phoretic mobility	Diameter (nm)	Particle composition (weight %)				Major apoli-proteins
				Trigly-ceride	Chole-sterol	Phos-pholipid	Protein	
Chylo-micron	> 0.93	Origin	80-1200	85-95	2-5	3-8	1-2	B-48
VLDL	0.93-1.006	Pre-β	30-80	50	22	19	8	B-100
IDL	1.006–1.019	β	23-35	20	38	23	19	B100
LDL	1.019–1.063	β	18-25	11	47	22	21	B100
HDL_2	1.063–1.125	α	9-12	6	22	30	41	A-I
HDL_3	1.125–1.21	α	5-9	6	15	23	55	A-I

Figure 1.1 Initiation and progression of atherosclerosis



Most schemata of the initiation of atherosclerosis posit a causal role for oxidized LDL particles as ligands for the scavenger receptors that facilitate foam cell formation. Reproduced with permission from Springer Nature, Atherosclerosis. Nat Rev Dis Primers 5, 56 (2019).

Figure 1.2 Lipid metabolism and targets for pharmacotherapy



Reproduced with permission from Erin D. Michos, NEJM. 2019; 381:1557-1567., Copyright Massachusetts Medical Society.

References

1 Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet (London, England)* **396**, 1204-1222, doi:10.1016/s0140-6736(20)30925-9 (2020).

2 Roth, G. A. *et al.* Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019: Update From the GBD 2019 Study. *Journal of the American College of Cardiology* **76**, 2982-3021, doi:10.1016/j.jacc.2020.11.010 (2020).

3 Yusuf, S. *et al.* Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet (London, England)* **364**, 937-952, doi:10.1016/s0140-6736(04)17018-9 (2004).

4 McQueen, M. J. *et al.* Lipids, lipoproteins, and apolipoproteins as risk markers of myocardial infarction in 52 countries (the INTERHEART study): a case-control study. *Lancet (London, England)* **372**, 224-233, doi:10.1016/s0140-6736(08)61076-4 (2008).

5 Borén, J. *et al.* Low-density lipoproteins cause atherosclerotic cardiovascular disease: pathophysiological, genetic, and therapeutic insights: a consensus statement from the European Atherosclerosis Society Consensus Panel. *European heart journal* **41**, 2313-2330, doi:10.1093/eurheartj/ehz962 (2020).

6 Ference, B. A. *et al.* Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. *European heart journal* **38**, 2459-2472, doi:10.1093/eurheartj/ehx144 (2017).

7 Libby, P. *et al.* Atherosclerosis. *Nature reviews. Disease primers* **5**, 56, doi:10.1038/s41572-019-0106-z (2019).

8 Skålén, K. *et al.* Subendothelial retention of atherogenic lipoproteins in early atherosclerosis. *Nature* **417**, 750-754, doi:10.1038/nature00804 (2002).

9 Tabas, I., Williams, K. J. & Borén, J. Subendothelial lipoprotein retention as the initiating process in atherosclerosis: update and therapeutic implications. *Circulation* **116**, 1832-1844, doi:10.1161/circulationaha.106.676890 (2007).

10 Weber, C. & Noels, H. Atherosclerosis: current pathogenesis and therapeutic options. *Nature medicine* **17**, 1410-1422, doi:10.1038/nm.2538 (2011).

11 Moore, K. J. *et al.* Macrophage Trafficking, Inflammatory Resolution, and Genomics in Atherosclerosis: JACC Macrophage in CVD Series (Part 2).

Journal of the American College of Cardiology **72**, 2181-2197, doi:10.1016/j.jacc.2018.08.2147 (2018).

12 Shankman, L. S. *et al.* KLF4-dependent phenotypic modulation of smooth muscle cells has a key role in atherosclerotic plaque pathogenesis. *Nature medicine* **21**, 628-637, doi:10.1038/nm.3866 (2015).

13 Williams, M. C. *et al.* Coronary Artery Plaque Characteristics Associated With Adverse Outcomes in the SCOT-HEART Study. *Journal of the American College of Cardiology* **73**, 291-301, doi:10.1016/j.jacc.2018.10.066 (2019).

14 Bentzon, J. F., Otsuka, F., Virmani, R. & Falk, E. Mechanisms of plaque formation and rupture. *Circulation research* **114**, 1852-1866, doi:10.1161/circresaha.114.302721 (2014).

15 Falk, E., Nakano, M., Bentzon, J. F., Finn, A. V. & Virmani, R. Update on acute coronary syndromes: the pathologists' view. *European heart journal* **34**, 719-728, doi:10.1093/eurheartj/ehs411 (2013).

16 Franck, G. *et al.* Flow Perturbation Mediates Neutrophil Recruitment and Potentiates Endothelial Injury via TLR2 in Mice: Implications for Superficial Erosion. *Circulation research* **121**, 31-42, doi:10.1161/circresaha.117.310694 (2017).

17 Pasterkamp, G., den Ruijter, H. M. & Libby, P. Temporal shifts in clinical presentation and underlying mechanisms of atherosclerotic disease. *Nature reviews. Cardiology* **14**, 21-29, doi:10.1038/nrcardio.2016.166 (2017).

18 Ference, B. A., Kastelein, J. J. P. & Catapano, A. L. Lipids and Lipoproteins in 2020. *Jama* **324**, 595-596, doi:10.1001/jama.2020.5685 (2020).

19 Goldstein, J. L. & Brown, M. S. The LDL receptor. *Arteriosclerosis, thrombosis, and vascular biology* **29**, 431-438, doi:10.1161/atvbaha.108.179564 (2009).

20 Mach, F. *et al.* 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk. *European heart journal* **41**, 111-188, doi:10.1093/eurheartj/ehz455 (2020).

21 Baigent, C. *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 (London, England)* **376**, 1670-1681, doi:10.1016/s0140-6736(10)61350-5 (2010).

22 Silverman, M. G. *et al.* Association Between Lowering LDL-C and Cardiovascular Risk Reduction Among Different Therapeutic Interventions: A Systematic Review and Meta-analysis. *Jama* **316**, 1289-1297, doi:10.1001/jama.2016.13985 (2016).

23 Goldstein, J. L. & Brown, M. S. A century of cholesterol and coronaries: from plaques to genes to statins. *Cell* **161**, 161-172, doi:10.1016/j.cell.2015.01.036 (2015).

24 Nordestgaard, B. G. *et al.* Familial hypercholesterolaemia is

underdiagnosed and undertreated in the general population: guidance for clinicians to prevent coronary heart disease: consensus statement of the European Atherosclerosis Society. *European heart journal* **34**, 3478-3490a, doi:10.1093/eurheartj/eht273 (2013).

25 Cuchel, M. *et al.* Homozygous familial hypercholesterolaemia: new insights and guidance for clinicians to improve detection and clinical management. A position paper from the Consensus Panel on Familial Hypercholesterolaemia of the European Atherosclerosis Society. *European heart journal* **35**, 2146-2157, doi:10.1093/eurheartj/ehu274 (2014).

26 Khera, A. V. *et al.* Diagnostic Yield and Clinical Utility of Sequencing Familial Hypercholesterolemia Genes in Patients With Severe Hypercholesterolemia. *Journal of the American College of Cardiology* **67**, 2578-2589, doi:10.1016/j.jacc.2016.03.520 (2016).

27 Wiegman, A. *et al.* Familial hypercholesterolaemia in children and adolescents: gaining decades of life by optimizing detection and treatment. *European heart journal* **36**, 2425-2437, doi:10.1093/eurheartj/ehv157 (2015).

28 Di Angelantonio, E. *et al.* Lipid-related markers and cardiovascular disease prediction. *Jama* **307**, 2499-2506, doi:10.1001/jama.2012.6571 (2012).

29 Lewington, S. *et al.* 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 (London, England)* **370**, 1829-1839, doi:10.1016/s0140-6736(07)61778-4 (2007).

30 Collins, R. *et al.* Interpretation of the evidence for the efficacy and safety of statin therapy. *Lancet (London, England)* **388**, 2532-2561, doi:10.1016/s0140-6736(16)31357-5 (2016).

31 Lauridsen, B. K., Stender, S., Frikke-Schmidt, R., Nordestgaard, B. G. & Tybjærg-Hansen, A. Genetic variation in the cholesterol transporter NPC1L1, ischaemic vascular disease, and gallstone disease. *European heart journal* **36**, 1601-1608, doi:10.1093/eurheartj/ehv108 (2015).

32 Ference, B. A., Majeed, F., Penumetcha, R., Flack, J. M. & Brook, R. D. Effect of naturally random allocation to lower low-density lipoprotein cholesterol on the risk of coronary heart disease mediated by polymorphisms in NPC1L1, HMGCR, or both: a 2 × 2 factorial Mendelian randomization study. *Journal of the American College of Cardiology* **65**, 1552-1561, doi:10.1016/j.jacc.2015.02.020 (2015).

33 Ference, B. A. *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. *Journal of the American College of Cardiology* **60**, 2631-2639, doi:10.1016/j.jacc.2012.09.017 (2012).

34 Holmes, M. V. *et al.* Mendelian randomization of blood lipids for coronary

heart disease. *European heart journal* **36**, 539-550, doi:10.1093/eurheartj/eht571 (2015).
35 Richardson, T. G. *et al.* Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: A multivariable Mendelian randomisation analysis. *PLoS medicine* **17**, e1003062, doi:10.1371/journal.pmed.1003062 (2020).
36 Welsh, C. *et al.* Comparison of Conventional Lipoprotein Tests and Apolipoproteins in the Prediction of Cardiovascular Disease. *Circulation* **140**, 542-552, doi:10.1161/circulationaha.119.041149 (2019).
37 Sniderman, A. D. *et al.* Apolipoprotein B Particles and Cardiovascular Disease: A Narrative Review. *JAMA cardiology* **4**, 1287-1295, doi:10.1001/jamacardio.2019.3780 (2019).
38 Lincoff, A. M. *et al.* Evacetrapib and Cardiovascular Outcomes in High-Risk Vascular Disease. *The New England journal of medicine* **376**, 1933-1942, doi:10.1056/NEJMoa1609581 (2017).
39 Ference, B. A. *et al.* Association of Genetic Variants Related to CETP Inhibitors and Statins With Lipoprotein Levels and Cardiovascular Risk. *Jama* **318**, 947-956, doi:10.1001/jama.2017.11467 (2017).
40 Ruuth, M. *et al.* Susceptibility of low-density lipoprotein particles to aggregate depends on particle lipidome, is modifiable, and associates with future cardiovascular deaths. *European heart journal* **39**, 2562-2573, doi:10.1093/eurheartj/ehy319 (2018).
41 Krauss, R. M. & Burke, D. J. Identification of multiple subclasses of plasma low density lipoproteins in normal humans. *Journal of lipid research* **23**, 97-104 (1982).
42 Diffenderfer, M. R. & Schaefer, E. J. The composition and metabolism of large and small LDL. *Current opinion in lipidology* **25**, 221-226, doi:10.1097/mol.0000000000000067 (2014).
43 Berneis, K. K. & Krauss, R. M. Metabolic origins and clinical significance of LDL heterogeneity. *Journal of lipid research* **43**, 1363-1379, doi:10.1194/jlr.r200004-jlr200 (2002).
44 Adiels, M. *et al.* Overproduction of large VLDL particles is driven by increased liver fat content in man. *Diabetologia* **49**, 755-765, doi:10.1007/s00125-005-0125-z (2006).
45 Griffin, B. A. *et al.* Rapid isolation of low density lipoprotein (LDL) subfractions from plasma by density gradient ultracentrifugation. *Atherosclerosis* **83**, 59-67, doi:10.1016/0021-9150(90)90131-2 (1990).
46 Benn, M., Nordestgaard, B. G., Jensen, G. B. & Tybjaerg-Hansen, A. Improving prediction of ischemic cardiovascular disease in the general population using apolipoprotein B: the Copenhagen City Heart Study. *Arteriosclerosis, thrombosis, and vascular biology* **27**, 661-670, doi:10.1161/01.ATV.0000255580.73689.8e (2007).
47 Boekholdt, S. M. *et al.* Association of LDL cholesterol, non-HDL cholesterol,

and apolipoprotein B levels with risk of cardiovascular events among patients treated with statins: a meta-analysis. *Jama* **307**, 1302-1309, doi:10.1001/jama.2012.366 (2012).

48 Robinson, J. G., Wang, S. & Jacobson, T. A. Meta-analysis of comparison of effectiveness of lowering apolipoprotein B versus low-density lipoprotein cholesterol and nonhigh-density lipoprotein cholesterol for cardiovascular risk reduction in randomized trials. *The American journal of cardiology* **110**, 1468-1476, doi:10.1016/j.amjcard.2012.07.007 (2012).

49 Thanassoulis, G. *et al.* Relations of change in plasma levels of LDL-C, non-HDL-C and apoB with risk reduction from statin therapy: a meta-analysis of randomized trials. *Journal of the American Heart Association* **3**, e000759, doi:10.1161/jaha.113.000759 (2014).

50 Mora, S. *et al.* Lipoprotein particle profiles by nuclear magnetic resonance compared with standard lipids and apolipoproteins in predicting incident cardiovascular disease in women. *Circulation* **119**, 931-939, doi:10.1161/circulationaha.108.816181 (2009).

51 Lawler, P. R. *et al.* Discordance between Circulating Atherogenic Cholesterol Mass and Lipoprotein Particle Concentration in Relation to Future Coronary Events in Women. *Clinical chemistry* **63**, 870-879, doi:10.1373/clinchem.2016.264515 (2017).

52 Gofman, J. W., Young, W. & Tandy, R. Ischemic heart disease, atherosclerosis, and longevity. *Circulation* **34**, 679-697, doi:10.1161/01.cir.34.4.679 (1966).

53 Gordon, T., Castelli, W. P., Hjortland, M. C., Kannel, W. B. & Dawber, T. R. High density lipoprotein as a protective factor against coronary heart disease. The Framingham Study. *The American journal of medicine* **62**, 707-714, doi:10.1016/0002-9343(77)90874-9 (1977).

54 Castelli, W. P. *et al.* Incidence of coronary heart disease and lipoprotein cholesterol levels. The Framingham Study. *Jama* **256**, 2835-2838 (1986).

55 Di Angelantonio, E. *et al.* Major lipids, apolipoproteins, and risk of vascular disease. *Jama* **302**, 1993-2000, doi:10.1001/jama.2009.1619 (2009).

56 Madsen, C. M., Varbo, A. & Nordestgaard, B. G. Extreme high high-density lipoprotein cholesterol is paradoxically associated with high mortality in men and women: two prospective cohort studies. *European heart journal* **38**, 2478-2486, doi:10.1093/eurheartj/ehx163 (2017).

57 Ko, D. T. *et al.* High-Density Lipoprotein Cholesterol and Cause-Specific Mortality in Individuals Without Previous Cardiovascular Conditions: The CANHEART Study. *Journal of the American College of Cardiology* **68**, 2073-2083, doi:10.1016/j.jacc.2016.08.038 (2016).

58 Wilkins, J. T. *et al.* Coronary heart disease risks associated with high levels of HDL cholesterol. *Journal of the American Heart Association* **3**, e000519, doi:10.1161/jaha.113.000519 (2014).

59 Mazidi, M., Mikhailidis, D. P. & Banach, M. Associations between risk of

overall mortality, cause-specific mortality and level of inflammatory factors with extremely low and high high-density lipoprotein cholesterol levels among American adults. *International journal of cardiology* **276**, 242-247, doi:10.1016/j.ijcard.2018.11.095 (2019).

60 Boden, W. E. *et al.* Niacin in patients with low HDL cholesterol levels receiving intensive statin therapy. *The New England journal of medicine* **365**, 2255-2267, doi:10.1056/NEJMoa1107579 (2011).

61 Landray, M. J. *et al.* Effects of extended-release niacin with laropiprant in high-risk patients. *The New England journal of medicine* **371**, 203-212, doi:10.1056/NEJMoa1300955 (2014).

62 Pownall, H. J., Rosales, C., Gillard, B. K. & Gotto, A. M., Jr. High-density lipoproteins, reverse cholesterol transport and atherogenesis. *Nature reviews. Cardiology* **18**, 712-723, doi:10.1038/s41569-021-00538-z (2021).

63 Ginsberg, H. N. The ACCORD (Action to Control Cardiovascular Risk in Diabetes) Lipid trial: what we learn from subgroup analyses. *Diabetes care* **34 Suppl 2**, S107-108, doi:10.2337/dc11-s203 (2011).

64 Briel, M. *et al.* Association between change in high density lipoprotein cholesterol and cardiovascular disease morbidity and mortality: systematic review and meta-regression analysis. *BMJ (Clinical research ed.)* **338**, b92, doi:10.1136/bmj.b92 (2009).

65 Barter, P. J. *et al.* Effects of torcetrapib in patients at high risk for coronary events. *The New England journal of medicine* **357**, 2109-2122, doi:10.1056/NEJMoa0706628 (2007).

66 Williams, S. A. *et al.* Improving Assessment of Drug Safety Through Proteomics: Early Detection and Mechanistic Characterization of the Unforeseen Harmful Effects of Torcetrapib. *Circulation* **137**, 999-1010, doi:10.1161/circulationaha.117.028213 (2018).

67 Schwartz, G. G. *et al.* Effects of dalcetrapib in patients with a recent acute coronary syndrome. *The New England journal of medicine* **367**, 2089-2099, doi:10.1056/NEJMoa1206797 (2012).

68 Bowman, L. *et al.* Effects of Anacetrapib in Patients with Atherosclerotic Vascular Disease. *The New England journal of medicine* **377**, 1217-1227, doi:10.1056/NEJMoa1706444 (2017).

69 Tall, A. R. & Rader, D. J. Trials and Tribulations of CETP Inhibitors. *Circulation research* **122**, 106-112, doi:10.1161/circresaha.117.311978 (2018).

70 Taheri, H., Filion, K. B., Windle, S. B., Reynier, P. & Eisenberg, M. J. Cholesteryl Ester Transfer Protein Inhibitors and Cardiovascular Outcomes: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Cardiology* **145**, 236-250, doi:10.1159/000505365 (2020).

71 Frikke-Schmidt, R. *et al.* Association of loss-of-function mutations in the

ABCA1 gene with high-density lipoprotein cholesterol levels and risk of ischemic heart disease. *Jama* **299**, 2524-2532, doi:10.1001/jama.299.21.2524 (2008).

72 Voight, B. F. *et al.* Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. *Lancet (London, England)* **380**, 572-580, doi:10.1016/s0140-6736(12)60312-2 (2012).

73 Johannsen, T. H., Frikke-Schmidt, R., Schou, J., Nordestgaard, B. G. & Tybjærg-Hansen, A. Genetic inhibition of CETP, ischemic vascular disease and mortality, and possible adverse effects. *Journal of the American College of Cardiology* **60**, 2041-2048, doi:10.1016/j.jacc.2012.07.045 (2012).

74 Haase, C. L. *et al.* LCAT, HDL cholesterol and ischemic cardiovascular disease: a Mendelian randomization study of HDL cholesterol in 54,500 individuals. *The Journal of clinical endocrinology and metabolism* **97**, E248-256, doi:10.1210/jc.2011-1846 (2012).

75 Rosenson, R. S. *et al.* Cholesterol efflux and atheroprotection: advancing the concept of reverse cholesterol transport. *Circulation* **125**, 1905-1919, doi:10.1161/circulationaha.111.066589 (2012).

76 Khera, A. V. *et al.* Cholesterol efflux capacity, high-density lipoprotein function, and atherosclerosis. *The New England journal of medicine* **364**, 127-135, doi:10.1056/NEJMoa1001689 (2011).

77 Guerin, M. *et al.* Association of Serum Cholesterol Efflux Capacity With Mortality in Patients With ST-Segment Elevation Myocardial Infarction. *Journal of the American College of Cardiology* **72**, 3259-3269, doi:10.1016/j.jacc.2018.09.080 (2018).

78 Soria-Florido, M. T. *et al.* Dysfunctional High-Density Lipoproteins Are Associated With a Greater Incidence of Acute Coronary Syndrome in a Population at High Cardiovascular Risk: A Nested Case-Control Study. *Circulation* **141**, 444-453, doi:10.1161/circulationaha.119.041658 (2020).

79 Li, X. M. *et al.* Paradoxical association of enhanced cholesterol efflux with increased incident cardiovascular risks. *Arteriosclerosis, thrombosis, and vascular biology* **33**, 1696-1705, doi:10.1161/atvbaha.113.301373 (2013).

80 Casula, M., Colpani, O., Xie, S., Catapano, A. L. & Baragetti, A. HDL in Atherosclerotic Cardiovascular Disease: In Search of a Role. *Cells* **10**, doi:10.3390/cells10081869 (2021).

81 Miller, M. *et al.* Triglycerides and cardiovascular disease: a scientific statement from the American Heart Association. *Circulation* **123**, 2292-2333, doi:10.1161/CIR.0b013e3182160726 (2011).

82 Sandesara, P. B., Virani, S. S., Fazio, S. & Shapiro, M. D. The Forgotten Lipids: Triglycerides, Remnant Cholesterol, and Atherosclerotic Cardiovascular Disease Risk. *Endocrine reviews* **40**, 537-557, doi:10.1210/er.2018-00184 (2019).

83 Ballantyne, C. M. *et al.* Influence of low high-density lipoprotein cholesterol

and elevated triglyceride on coronary heart disease events and response to simvastatin therapy in 4S. *Circulation* **104**, 3046-3051, doi:10.1161/hc5001.100624 (2001).

84 Manninen, V. *et al.* Joint effects of serum triglyceride and LDL cholesterol and HDL cholesterol concentrations on coronary heart disease risk in the Helsinki Heart Study. Implications for treatment. *Circulation* **85**, 37-45, doi:10.1161/01.cir.85.1.37 (1992).

85 Sarwar, N. *et al.* Triglycerides and the risk of coronary heart disease: 10,158 incident cases among 262,525 participants in 29 Western prospective studies. *Circulation* **115**, 450-458, doi:10.1161/circulationaha.106.637793 (2007).

86 Hokanson, J. E. & Austin, M. A. Plasma triglyceride level is a risk factor for cardiovascular disease independent of high-density lipoprotein cholesterol level: a meta-analysis of population-based prospective studies. *Journal of cardiovascular risk* **3**, 213-219 (1996).

87 Miller, M. *et al.* Impact of triglyceride levels beyond low-density lipoprotein cholesterol after acute coronary syndrome in the PROVE IT-TIMI 22 trial. *Journal of the American College of Cardiology* **51**, 724-730, doi:10.1016/j.jacc.2007.10.038 (2008).

88 Faergeman, O. *et al.* Plasma triglycerides and cardiovascular events in the Treating to New Targets and Incremental Decrease in End-Points through Aggressive Lipid Lowering trials of statins in patients with coronary artery disease. *The American journal of cardiology* **104**, 459-463, doi:10.1016/j.amjcard.2009.04.008 (2009).

89 Stein, E. A., Lane, M. & Laskarzewski, P. Comparison of statins in hypertriglyceridemia. *The American journal of cardiology* **81**, 66b-69b, doi:10.1016/s0002-9149(98)00041-1 (1998).

90 Ginsberg, H. N. *et al.* Effects of combination lipid therapy in type 2 diabetes mellitus. *The New England journal of medicine* **362**, 1563-1574, doi:10.1056/NEJMoa1001282 (2010).

91 Robins, S. J. *et al.* Relation of gemfibrozil treatment and lipid levels with major coronary events: VA-HIT: a randomized controlled trial. *Jama* **285**, 1585-1591, doi:10.1001/jama.285.12.1585 (2001).

92 Guyton, J. R. *et al.* Relationship of lipoproteins to cardiovascular events: the AIM-HIGH Trial (Atherothrombosis Intervention in Metabolic Syndrome With Low HDL/High Triglycerides and Impact on Global Health Outcomes). *Journal of the American College of Cardiology* **62**, 1580-1584, doi:10.1016/j.jacc.2013.07.023 (2013).

93 Bucher, H. C., Hengstler, P., Schindler, C. & Meier, G. N-3 polyunsaturated fatty acids in coronary heart disease: a meta-analysis of randomized controlled trials. *The American journal of medicine* **112**, 298-304, doi:10.1016/s0002-9343(01)01114-7 (2002).

94 Rizos, E. C., Ntzani, E. E., Bika, E., Kostapanos, M. S. & Elisaf, M. S.

Association between omega-3 fatty acid supplementation and risk of major cardiovascular disease events: a systematic review and meta-analysis. *Jama* **308**, 1024-1033, doi:10.1001/2012.jama.11374 (2012).

95 Yokoyama, M. *et al.* Effects of eicosapentaenoic acid on major coronary events in hypercholesterolaemic patients (JELIS): a randomised open-label, blinded endpoint analysis. *Lancet (London, England)* **369**, 1090-1098, doi:10.1016/s0140-6736(07)60527-3 (2007).

96 Nordestgaard, B. G. *et al.* Heterozygous lipoprotein lipase deficiency: frequency in the general population, effect on plasma lipid levels, and risk of ischemic heart disease. *Circulation* **96**, 1737-1744, doi:10.1161/01.cir.96.6.1737 (1997).

97 Jørgensen, A. B., Frikke-Schmidt, R., Nordestgaard, B. G. & Tybjærg-Hansen, A. Loss-of-function mutations in APOC3 and risk of ischemic vascular disease. *The New England journal of medicine* **371**, 32-41, doi:10.1056/NEJMoa1308027 (2014).

98 Khera, A. V. *et al.* Association of Rare and Common Variation in the Lipoprotein Lipase Gene With Coronary Artery Disease. *Jama* **317**, 937-946, doi:10.1001/jama.2017.0972 (2017).

99 Dewey, F. E. *et al.* Inactivating Variants in ANGPTL4 and Risk of Coronary Artery Disease. *The New England journal of medicine* **374**, 1123-1133, doi:10.1056/NEJMoa1510926 (2016).

100 Stitziel, N. O. *et al.* Coding Variation in ANGPTL4, LPL, and SVEP1 and the Risk of Coronary Disease. *The New England journal of medicine* **374**, 1134-1144, doi:10.1056/NEJMoa1507652 (2016).

101 Wittrup, H. H., Tybjaerg-Hansen, A. & Nordestgaard, B. G. Lipoprotein lipase mutations, plasma lipids and lipoproteins, and risk of ischemic heart disease. A meta-analysis. *Circulation* **99**, 2901-2907, doi:10.1161/01.cir.99.22.2901 (1999).

102 Dewey, F. E. *et al.* Genetic and Pharmacologic Inactivation of ANGPTL3 and Cardiovascular Disease. *The New England journal of medicine* **377**, 211-221, doi:10.1056/NEJMoa1612790 (2017).

103 Gaudet, D. *et al.* Antisense Inhibition of Apolipoprotein C-III in Patients with Hypertriglyceridemia. *The New England journal of medicine* **373**, 438-447, doi:10.1056/NEJMoa1400283 (2015).

104 Do, R. *et al.* Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nature genetics* **45**, 1345-1352, doi:10.1038/ng.2795 (2013).

105 Thomsen, M., Varbo, A., Tybjærg-Hansen, A. & Nordestgaard, B. G. Low nonfasting triglycerides and reduced all-cause mortality: a mendelian randomization study. *Clinical chemistry* **60**, 737-746, doi:10.1373/clinchem.2013.219881 (2014).

106 Sarwar, N. *et al.* Triglyceride-mediated pathways and coronary disease: collaborative analysis of 101 studies. *Lancet (London, England)* **375**,

1634-1639, doi:10.1016/s0140-6736(10)60545-4 (2010).
107 Ference, B. A. *et al.* Association of Triglyceride-Lowering LPL Variants and LDL-C-Lowering LDLR Variants With Risk of Coronary Heart Disease. *Jama* **321**, 364-373, doi:10.1001/jama.2018.20045 (2019).
108 Gencer, B., Kronenberg, F., Stroes, E. S. & Mach, F. Lipoprotein(a): the revenant. *European heart journal* **38**, 1553-1560, doi:10.1093/eurheartj/ehx033 (2017).
109 van der Valk, F. M. *et al.* Oxidized Phospholipids on Lipoprotein(a) Elicit Arterial Wall Inflammation and an Inflammatory Monocyte Response in Humans. *Circulation* **134**, 611-624, doi:10.1161/circulationaha.116.020838 (2016).
110 Boffa, M. B. & Koschinsky, M. L. Lipoprotein (a): truly a direct prothrombotic factor in cardiovascular disease? *Journal of lipid research* **57**, 745-757, doi:10.1194/jlr.R060582 (2016).
111 Tsimikas, S. A Test in Context: Lipoprotein(a): Diagnosis, Prognosis, Controversies, and Emerging Therapies. *Journal of the American College of Cardiology* **69**, 692-711, doi:10.1016/j.jacc.2016.11.042 (2017).
112 Nordestgaard, B. G. *et al.* Lipoprotein(a) as a cardiovascular risk factor: current status. *European heart journal* **31**, 2844-2853, doi:10.1093/eurheartj/ehq386 (2010).
113 Erqou, S. *et al.* Lipoprotein(a) concentration and the risk of coronary heart disease, stroke, and nonvascular mortality. *Jama* **302**, 412-423, doi:10.1001/jama.2009.1063 (2009).
114 Danesh, J., Collins, R. & Peto, R. Lipoprotein(a) and coronary heart disease. Meta-analysis of prospective studies. *Circulation* **102**, 1082-1085, doi:10.1161/01.cir.102.10.1082 (2000).
115 Nave, A. H. *et al.* Lipoprotein (a) as a risk factor for ischemic stroke: a meta-analysis. *Atherosclerosis* **242**, 496-503, doi:10.1016/j.atherosclerosis.2015.08.021 (2015).
116 Erqou, S. *et al.* Apolipoprotein(a) isoforms and the risk of vascular disease: systematic review of 40 studies involving 58,000 participants. *Journal of the American College of Cardiology* **55**, 2160-2167, doi:10.1016/j.jacc.2009.10.080 (2010).
117 Paré, G. *et al.* Lipoprotein(a) Levels and the Risk of Myocardial Infarction Among 7 Ethnic Groups. *Circulation* **139**, 1472-1482, doi:10.1161/circulationaha.118.034311 (2019).
118 Willeit, P. *et al.* Baseline and on-statin treatment lipoprotein(a) levels for prediction of cardiovascular events: individual patient-data meta-analysis of statin outcome trials. *Lancet (London, England)* **392**, 1311-1320, doi:10.1016/s0140-6736(18)31652-0 (2018).
119 Kamstrup, P. R., Tybjaerg-Hansen, A., Steffensen, R. & Nordestgaard, B. G. Genetically elevated lipoprotein(a) and increased risk of myocardial infarction. *Jama* **301**, 2331-2339, doi:10.1001/jama.2009.801 (2009).

120 Clarke, R. *et al.* Genetic variants associated with Lp(a) lipoprotein level and coronary disease. *The New England journal of medicine* **361**, 2518-2528, doi:10.1056/NEJMoa0902604 (2009).
121 Sabatine, M. S. *et al.* Evolocumab and Clinical Outcomes in Patients with Cardiovascular Disease. *The New England journal of medicine* **376**, 1713-1722, doi:10.1056/NEJMoa1615664 (2017).
122 Khera, A. V. *et al.* Lipoprotein(a) concentrations, rosuvastatin therapy, and residual vascular risk: an analysis from the JUPITER Trial (Justification for the Use of Statins in Prevention: an Intervention Trial Evaluating Rosuvastatin). *Circulation* **129**, 635-642, doi:10.1161/circulationaha.113.004406 (2014).
123 Raal, F. J. *et al.* Reduction in lipoprotein(a) with PCSK9 monoclonal antibody evolocumab (AMG 145): a pooled analysis of more than 1,300 patients in 4 phase II trials. *Journal of the American College of Cardiology* **63**, 1278-1288, doi:10.1016/j.jacc.2014.01.006 (2014).
124 Burgess, S. *et al.* Association of LPA Variants With Risk of Coronary Disease and the Implications for Lipoprotein(a)-Lowering Therapies: A Mendelian Randomization Analysis. *JAMA cardiology* **3**, 619-627, doi:10.1001/jamacardio.2018.1470 (2018).
125 Viney, N. J. *et al.* Antisense oligonucleotides targeting apolipoprotein(a) in people with raised lipoprotein(a): two randomised, double-blind, placebo-controlled, dose-ranging trials. *Lancet (London, England)* **388**, 2239-2253, doi:10.1016/s0140-6736(16)31009-1 (2016).
126 Tapia-Conyer, R. *et al.* Cohort profile: the Mexico City Prospective Study. *Int J Epidemiol* **35**, 243-249, doi:10.1093/ije/dyl042 (2006).
127 Alegre-Díaz, J. *et al.* Diabetes and Cause-Specific Mortality in Mexico City. *The New England journal of medicine* **375**, 1961-1971, doi:10.1056/NEJMoa1605368 (2016).

CHAPTER 2: Literature review

Chapter summary

The current chapter systematically reviews previous evidence on the association of *NMR-measured* lipid-related traits and their subfractions with the risk of ASCVD from observational epidemiological studies and MR studies, as well as relevant GWAS on major lipid traits. Findings from observational studies consistently suggested that concentrations of apoB-containing particles, including VLDL-P, IDL-P, and LDL-P, were associated with the risk of a range of ASCVD outcomes. Further MR studies also prioritized apoB as a causal risk factor. However, evidence of the association with particle size of apoB particles was not consistent across studies. Regarding the concentration and size of HDL particles, most observational studies suggested an inverse relationship with risk, with some studies reporting a lack of association. For lipids contained in lipoproteins such as TG and cholesterol, previous genetic and observational studies suggested that these lipids in apoB containing particles were also positively related to risk but their effects may vary by particle sizes. These associations often diminish after adjusting for apoA1 and apoB. Overall, these NMR-based findings broadly align with the understanding of lipids and ASCVD based on chemistry assays, as introduced in Chapter 1. A majority of GWAS based on composite lipid traits measured by traditional chemistry assay in European populations have identified thousands of significant lipid loci, whereas current understanding of the genetic determinants on metabolism of lipid subfractions remain limited. Finally, gaps in current research are highlighted, such as methodological limitations, the need for valid genetic instruments for lipid subfractions, and the lack of evidence from Latin American populations. The MCPS with extensive genetic and NMR data offer a unique opportunity to address these gaps and enhance understanding of lipid metabolism and its relevance to ASCVD.

2.1 Background and objectives

Consistent evidence has established that multiple blood lipids such as LDL-C and other apoB-containing particles are associated with increased risk of ASCVD. However, evidence of the contribution of functionally and chemically diverse lipoprotein subfractions to the development of ASCVD remains limited. Several studies have reported that these lipoprotein subfractions may independently and specifically affect different manifestations of ASCVD and may potentially improve risk assessment over traditional lipids (referred to LDL-C, HDL-C, TC, and TG). The main objective of this systematic review is to summarize the associations between lipoprotein subfractions measured by NMR spectroscopy and ASCVD by evaluating the totality of evidence from epidemiological cohort and case-control studies, and Mendelian randomization studies. Related GWAS of NMR lipid traits and traditional lipids are also included to review current understanding of the genetic regulation of blood lipids and their relationship with ASCVD. Secondary objectives are to critically appraise and discuss the body of evidence, and to identify gaps in the literature and implications for future studies.

2.2 Methods of literature search and systematic review

2.2.1 Search strategy

Consistent with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)[1] and Meta-Analyses of Observational Studies in Epidemiology (MOOSE) guidelines[2] and without language restrictions, a comprehensive search of PubMed and EMBASE online databases from inception through 2022 was performed to identify all relevant cohort, case-control, and MR studies on the associations between lipid-related traits measured by NMR spectroscopy and ASCVD. Additional searches of PubMed and EMBASE databases were conducted to identify GWAS of NMR lipoproteins and subfractions from 2007 (the year the first GWAS was published) through 2022. Representative GWAS of traditional lipids measured by chemical assays were selectively included in the systematic review through a manual search of the GWAS Catalog. Traditional lipid panels were not the focus of this study, and some participants from smaller GWAS on traditional lipids have already been included in larger studies such as those by the Global Lipids Genetics Consortium (GLGC). The reference lists of the identified relevant publications were screened to supplement the bibliographic database searches and identify additional potential data sources that met the eligibility criteria. The search terms were MeSH terms and text words relating to the exposure of interest (nuclear magnetic resonance spectroscopy), outcome (cardiovascular disease), and study characteristics (cohort study, case-control study, MR study, or GWAS). The detailed literature search strategy for PubMed and EMBASE is shown in **Table 2.1** and **Table 2.2**.

2.2.2 Eligibility criteria

Epidemiological studies were included if the following criteria were met: (1) cohort (either prospective or retrospective) OR case-control study design OR MR study; (2) participants: people free from cardiovascular diseases (CVDs) at the time of enrolment in the community-based cohort that was subsequently followed to determine the risk of

CVD incident or cardiovascular death OR patients with CVDs and matched control individuals enrolled in a case-control OR MR study; (3) measurement of lipid related traits: lipids, lipoproteins, and apolipoproteins related traits (concentration and particle size) were measured by nuclear magnetic resonance (NMR) spectroscopy; (4) study reported the risk estimates for CVD incidence or mortality, including the measure of association (e.g., risk ratio [RR], hazard ratio [HR], odds ratio [OR]) and the measure of uncertainty (e.g., 95% CI); 5) duration of follow-up (for observational prospective studies only): the associations between lipid traits and CVD outcomes may potentially be influenced by reverse causality (the reverse effect of diagnosed and undiagnosed CVDs on plasma lipid level), so studies were required to have at least 2 years' mean/median follow up in order to address reverse causality by avoiding overestimation of the association between lipid related traits and the outcomes of interest. GWASs were eligible if they investigated the association between genetic variants and NMR-measured lipoproteins and subfractions and included hypothesis-free estimates.

Cross-sectional studies and case reports, as well as abstracts, reviews, and comments were excluded. Studies base on traditional lipid panels (i.e., low density lipoprotein cholesterol (LDL-C) and other apolipoprotein (apo) B-containing lipoproteins) measured by clinical assays were also excluded. Although traditional clinically used lipid profiles have been used for atherosclerotic cardiovascular disease (ASCVD) risk assessment, previous studies mainly focused on the total concentration of cholesterol or triglyceride of major lipoprotein class, failing to delineate the contribution of functionally diverse lipid subfractions to the risk of cardiovascular diseases and mortality. In this study, studies based on NMR metabolomics that provided quantitative data on lipoprotein subclasses, their lipid concentrations and composition, and apolipoproteins, were focused and included. Studies focusing on fatty acids and other metabolic traits measured by NMR spectroscopy were also excluded. When multiple studies reported similar outcomes based on the same study sample and lipid traits, the most comprehensive study with the largest sample size was included.

2.2.3 Study selection and data extraction

The eligibility of the articles was evaluated by a single reviewer. After removing duplicates, the titles and abstracts of all potentially relevant studies were screened and clearly unrelated articles were excluded. The full texts of the remaining articles were checked and included in the final list if they met eligibility criteria. The data extracted included study characteristics (study name, first author, publication year, country, number of participants, recruitment and follow-up period), participant characteristics (baseline age, percentage of male), main exposure (NMR lipid subclasses), main outcome (NMR lipid traits, association strength with corresponding 95% CIs), adjusted confounding variables, and exclusions of chronic diseases at baseline (for cohort studies). The maximally adjusted risk estimates that were available were extracted.

2.3 Findings of observational study

2.3.1 Study inclusion

The process of identifying relevant case-control or cohort studies investigating the association between NMR lipid traits and risk of ASCVD is shown in **Figure 2.1**. Among the 848 records initially retrieved from PubMed and Embase database, 721 non-duplicated titles were identified, leaving 171 abstracts for further screening. Of these, 64 studies were included in the full-text reading list. Finally, 26 eligible studies including 2 meta-analysis of cohort studies, 14 prospective cohort studies, 9 case-control studies and 1 case-cohort study were included in the systematic review.

The detailed study characteristics are shown in **Table 2.2**. Overall, the studies were published between 2002 and 2021, involving participants from Europe (n = 10), the US (n = 12), Asia (n = 1), and International collaborations (n = 3). The study sample size ranged from 118 to 38,797 participants, with a median sample size of 4387. All cohort studies excluded individuals with prevalent CVD at baseline or were conducted in individuals without CVD, and most of the nested case-control studies were based on cohorts without history of CVD. Few studies excluded CVD cases or deaths that occurred

in the first few years of follow-up. NMR lipid traits were mainly measured by LipoScience (n=19) and Nightingale Health Ltd (n=6), with one study using a platform from Numares AG. Most studies identified CVD outcomes (individual or composite outcomes inclusive of fatal or non-fatal CHD, stroke, and PAD) through hospital records or death certificates, while others combined self-report with medical record confirmation. Adjustment for non-lipid confounders was performed in all studies, though with different types and numbers of variables across studies. Most studies further adjusted for chemically measured lipids to assess the independent effect of NMR lipid-related traits.

2.3.2 Nested case-control and case-cohort studies

Overall, 10 eligible studies with a case-control design published between 2002 and 2020 were identified,[3-12] including 1 case-cohort study based on a clinical trial,[12] 1 case-control study,[4] and 8 nested case-control studies[3,5-11] (5 of which were based on the same EPIC-Norfolk cohort but reported different lipid and lipoprotein particles).[6-10] Participants were mainly from UK (n=5) and US (n=3), with 1 study involving participants from China and 1 study from an International collaboration. The study sample size ranged from 188 to 4662 participants, with a median sample size of 2223. Most of the nested case-control studies excluded participants with the history of CVD at baseline and some further excluded those with cancer and use of statin therapy. All studies adjustment for confounders including chemically measured lipids and other traditional CVD risk factors, though with different types and numbers of variables across studies. The detailed study characteristics and findings are shown in **Table 2**.

In a nested case-control study among middle-aged women with no history of CVD or cancer at baseline from the Women's Health Study, LDL particle (LDL-P) concentration measured by NMR spectroscopy (especially the level of small but not large LDL) was associated with an increased risk of MI, stroke, and fatal CHD after adjustment for chemical lipid measures (TC to HDL-C ratio and TG) and other traditional risk factors (relative risk and 95% CI for participants in the highest versus lowest quartile: 2.90, 1.16-

7.30).[3] However, LDL particle size was no longer significant after adjustment for these factors.[3] Similarly, a case-control study from the Cardiovascular Health study reported that LDL particle number (but not particle size) and VLDL were associated with an increased risk of MI and angina after adjustment for traditional TG and HDL-C in elderly women but not men.[4] In addition, large NMR HDL particle (but not small HDL) and HDL-C were inversely related to ASCVD in both men and women.[4] By contrast, a case-control study of individuals with diabetes suggested that VLDL but not LDL particles were associated with increased ASCVD risk after adjustment for chemically measured lipids while large HDL particles were related to decreased risk.[5]

In five subsequent nested case-control studies based on the EPIC (European Prospective Investigation into Cancer and Nutrition)-Norfolk cohort,[6-10] the relationship of standard chemically measured and NMR lipids and lipoproteins, apolipoproteins with ASCVD risk and risk factors of CVD were assessed. LDL particle concentration but not particle size was related to CHD after adjusting for LDL-C, HDL-C, and TG.[6] HDL particle concentration and HDL particle size were independently associated with decreased risk of ASCVD. However, the association with HDL particle size was abolished after adjustment for TG and apoB, indicating the relationship between HDL size and ASCVD risk was partly explained by factors associated with metabolic syndrome.[8] Notably, after adjusting for apoA-I and apoB, HDL-C and HDL particle size were found to be associated with increased risk at very high values.[7] These discrepant conclusions derived from the same cohort may be due to the difference of HDL subclass classification between studies. NMR LDL/HDL ratio, traditional TC/HDL-C and apoB/apoA-I were similarly associated with increased risk of CHD in both men and women regardless of metabolic syndrome.[9] Moreover, it was found that apoC-III levels were significantly associated with CHD, perhaps due to correlations with VLDL, IDL and small dense LDL particles.[10]

The above mentioned studies mainly focused on the concentration and size of LDL and HDL particles, however the association of lipid content in other lipoprotein subclasses and apolipoproteins with ASCVD were not reported in these studies. In a recent nested

case-control study of the China Kadoorie Biobank, the relationships of 225 metabolic measures by NMR spectroscopy, including lipids, 14 lipoprotein subclasses, and apolipoproteins, with MI, ischemic stroke (IS), and intracerebral hemorrhage (ICH) were comprehensively assessed.[11] After adjustment for traditional CVD risk factors except chemically measured lipids, plasma levels of VLDL, IDL, and LDL particles were positively associated with MI and IS but not ICH. Large but not small HDL particles and cholesterol contained in these subclasses were inversely associated with MI and IS (for 1-SD higher cholesterol in large HDL, OR: 0.79 and 0.88, respectively; cholesterol in small HDL, OR: 0.99 and 1.06, respectively). TG in all lipoproteins, including most HDL particles, were positively associated with MI, with a similar pattern for IS.

In a case-cohort study based on a clinical trial of dalcetrapib (a CETP inhibitor), baseline HDL particles and their subclasses, as well as the change in HDL-P on treatment with dalcetrapib or placebo were not associated with the risk of major adverse cardiovascular events in patients with acute coronary syndrome after adjustment for LDL-C, TG, LDL particles, and VLDL particles.[12]

2.3.3 Cohort studies

Overall, 16 eligible studies with a prospective cohort design published from 2007 to 2021 were identified,[13-28] including 2 meta-analysis of cohort studies[26,27] and 14 prospective cohort studies.[13-25,28] The sample size ranged from 510 to 38,797 participants, with a median sample size of 7256. The follow-up duration ranged from 2 to 15 years. Several studies were based on the same cohort but investigated different lipids and lipoprotein subclasses (e.g. Women's Health Study, Multi-Ethnic Study of Atherosclerosis [MESA], and Justification for the Use of statins in Prevention: an Intervention Trial Evaluating Rosuvastatin study [JUPITER]). Adjustment for traditional non-lipid CVD risk factors was performed in all studies, and most (13 out of 16) further adjusted for chemically measured lipids. Few studies excluded the CVD cases or deaths that occurred in the first few years of follow-up. The detailed study characteristics and findings are shown in **Table 2.2**.

In a prospective cohort study from the Framingham Offspring Study,[13] LDL-P was more strongly related to future CVD in men and women than LDL-C or non-HDL-C. Adding VLDL-P to LDL-P only marginally strengthened the association with CVD compared with LDL-P alone. Notably, LDL-P size and the number of cholesterol molecules per LDL-P progressively decreased as TG and LDL-P levels increased within individuals with same LDL-C levels. This has also been indicated by previous studies showing that variations in TG levels lead to the cholesterol compositional variability of LDL.[29] In addition, this study also suggested that LDL-P became progressively cholesterol-depleted without change in particle size as LDL concentrations decreased within individuals with same TG levels, indicating that LDL-lowering therapies may result in a discordance changes between the levels of LDL-C and LDL-P.

Two subsequent cohort studies based on the Women's Health Study investigated the relationship of both the NMR-measured lipoprotein particles (VLDL, LDL, and HDL subclasses) and chemically measured lipids with the risk of CVD and peripheral artery disease (PAD).[14,23] NMR lipoproteins were significantly associated with incident CVD after adjustment for non-lipid risk factors, with risk estimates comparable but not superior to standard lipid panels or apolipoproteins measured by immunoassay.[14] Specially, LDL-P concentration was similar to apoB in risk prediction and both measurements performed better than LDL-C, however neither measurement was better than total/HDL cholesterol ratio that is readily available from a traditional lipid panel. The components of atherogenic dyslipidemia, including small LDL-P (but not LDL-C), medium and very large VLDL-P, TC to HDL-C ratio, HDL-C, and HDL-P were more strongly associated with incident PAD than incident CVD.[23] However, TG, apoB, non-HDL-C, and TC that were strongly related to CVD were not associated with the risk of PAD. In addition, the addition of HDL-P and LDL-P concentration to TC to HDL-C ratio further identified individuals at risk of PAD.

In three cohort studies based on the Multi-ethnic Study of Atherosclerosis (MESA),[15,17,19] HDL-P was inversely associated with risk of CHD after adjustment for LDL-C, HDL-C, and TG. Small LDL-P was not associated with the CHD risk after adjustment for HDL-C

and TG. LDL-P and LDL-P/VLDL-P ratio were associated with risk of CHD after adjustment for LDL-C and these associations further attenuated after adjustment for other traditional lipids.

In three analyses based on the Justification for the Use of statins in Prevention: an Intervention Trial Evaluating Rosuvastatin (JUPITER) study,[16,18,21] both baseline and on-treatment HDL-P but not HDL-P size showed significant associations with CVD.[16] Regarding HDL-P subclasses, baseline levels of the large and medium HDL particles were inversely associated with the risk of first vascular events.[18] However, on-treatment levels of very large HDL particles were significantly associated with an increased risk of vascular events.

Similarly, NMR measured apoB, apoB-containing lipoprotein particles, particle size of LDL and VLDL, TG, and TC were related to an increased risk of CVD in the FINRISK study,[20] while HDL-P, HDL-C, and HDL-P were inversely associated with CVD risk. HDL-P was inversely related to CVD in a German cohort study.[22] In the Copenhagen General Population Study,[28] VLDL-P, IDL-P, and LDL-P were consistently associated with increased risk of MI.

In a recent meta-analysis of 4 cohort studies,[26] HDL-P was inversely associated with the risk of MI and IS in the overall population after adjustment for cardiometabolic risk factors and HDL-C. However, neither HDL-P or HDL-C concentration was associated with MI in Blacks, indicating ASCVD risk may differ by vascular domain and ethnicity. In another meta-analysis of 7 cohort studies,[27] HDL-P size and cholesterol in HDL-P but not LDL-P were inversely associated with the risk of stroke. TG in LDL-P and very large VLDL-P were associated with increased risk of stroke.

2.4 Summary of observational studies and gap in current evidence

2.4.1 ApoB-containing lipoproteins, LDL/VLDL particle size, and ASCVD

Findings from included observational studies have consistently suggested that

concentration of apoB-containing lipoproteins measured by NMR spectroscopy using two analytic platforms (LipoScience and Nightingale), including VLDL-P, IDL-P, and LDL-P (Lp(a) are not available in NMR metabolomics), were associated with increased risk of CHD,[4,6,11,19,21,28] stroke,[11] and PAD[23] or composite ASCVD outcomes[3,5,13,14,20,24] after adjustment for conventional non-lipid cardiovascular risk factors. These associations were attenuated but remained significant in some studies after further adjusting for chemically measured lipids and apolipoproteins. These results are in line with the current understanding based on evidence of traditional lipid panels that LDL and other apoB-containing lipoproteins are directly implicated in the development of ASCVD, as introduced in **Chapter 1**.

However, evidence of the association between particle size of VLDL and LDL and the risk of ASCVD was not consistent, with some studies suggesting a positive[11,14,20] or inverse[11,14] association and some reporting a lack of association[3,4,6,21,23] after adjustment for non-lipid risk factors in all studies with particle size measurements and additionally for LDL particle concentration in two studies. In addition, evidence of the association between the concentration of VLDL and LDL subclasses and ASCVD was limited and inconsistent. Two studies based on the Women's Health Cohort reporting individual LDL/VLDL subclasses suggested that small but not large LDL was associated with increased risk of ASCVD[3,23] whereas the MESA study reported a lack of association with small LDL particle after adjustment for traditional CVD risk factors and chemically measured lipids,[17] while in the CKB study both large and small LDL particles were associated with MI and stroke after adjustment for traditional non-lipid risk factors only.[11]

2.4.2 Concentration and size of HDL particle and ASCVD

Regarding the concentration and size of HDL particles, most of the included observational studies suggested an inverse relationship of HDL particle concentrations with ASCVD risk after adjustment for non-lipid risk factors,[5,8,11,15,20,22,23,26] with some studies reported a lack of association.[4,12,14,16,27] In two studies, this inverse association

with HDL-P remained significant after additional adjustment for lipid-related traits including LDL particles, HDL-C, and TG.[15,22] However, the association between HDL particle size and ASCVD is relatively inconsistent, with some studies reporting a significant positive[7] or inverse[11,14,20,23,24,27] association but others suggesting a lack of association.[8,16,26] Most studies with the measurement of HDL subclasses concentration suggested that larger but not smaller HDL particles[22] were related to ASCVD risk,[5,11,14,18,23] although some studies reported a lack of association with large HDL-P[8,27] and one study suggested that all HDL-P subclasses were related to the risk of ASCVD.[20]

2.4.3 Lipids contained in lipoprotein particles and ASCVD

Several studies with the measurements of lipids contained in lipoproteins also suggested that TG and cholesterol contained in apoB-containing particles were positively associated with risk of stroke and MI,[11,25,27] while TG in smaller but not larger HDL-P were associated with increased risk of ASCVD.[11,25] Cholesterol in larger but not small HDL-P were inversely associated with MI and stroke.[11] Notably, in one study the association between TG in all 14 lipoprotein subclasses and ASCVD were not significant after adjustment for LDL-C or HDL (or apoA-I and apoB).[25] One study reported that total lipid in apoB-containing particles but not HDL-P, TC, and TG were associated with ASCVD after adjustment for age and sex.[20]

2.4.4 Additional value of NMR lipid traits to traditional lipid panel in ASCVD risk prediction

Whether NMR measured lipoprotein particles improve the prediction of ASCVD independently of conventional lipids and other risk factors remains unclear and inconsistent based on evidence from the included observational studies. Some studies suggested that the addition of NMR-derived lipoprotein measures to traditional lipid measures may improve risk assessment for PAD (adding LDL-P to LDL-C/TC:HDL-C or adding HDL-P to HDL-C),[23] and CHD (adding HDL-P to HDL-C,[5,22] or adding LDL-P/HDL-P to standard lipid panel[19]). However, one study suggested that the addition of LDL-P or apoB to a model that already included TC/HDL-C and non-lipid risk factors provided no

improvement.[14] LDL-P was related more strongly to ASCVD than LDL-C or non-HDL-C in the Framingham Offspring Study,[13] but the Women's Health Study suggested that NMR measures were comparable but not superior to standard lipids,[14] while the additional value of LDL-P to a model including LDL-C was lost after adjustment for HDL-C and TG in the EPIC-Norfolk study.[6]

2.4.5 Gap in current evidence from observational studies and implications for future study

First, although all included studies excluded individuals with the history of ASCVD and some further excluded those with other chronic diseases such as cancer and diabetes, few included cohort studies excluded the ASCVD cases or death that occurred in the first few years of follow-up. Considering the associations between NMR lipid-related traits and cardiovascular outcomes may potentially be influenced by reverse causality (the reverse effect of diagnosed and undiagnosed CVD on lipid profiles; thus plasma lipid may not act as predictors of ASCVD development and survival but may instead be the outcomes of these diseases), future prospective cohort studies based on initially CVD free individuals may need to further exclude new cardiovascular outcomes occurring early in the follow-up duration and have sufficient length of follow-up to minimize the influence of reverse causality.

Second, although all included studies adjusted for conventional non-lipid risk factors of CVD such as age, sex, blood pressure, and smoking, and most studies further adjusted for chemically measured lipids, types and numbers of confounders adjusted varied across studies, making comparison of risk estimates between studies challenging. The associations of lipoprotein particle number with ASCVD after additional adjustment for traditional lipid panel were attenuated, indicating that these relationships may be partly explained by the lipid content within them as measured by chemical assay (although the possibility that lipoprotein particle numbers may identify CVD risk in a subpopulation of individuals with normal cholesterol level cannot be ruled out). In addition, the inverse association between HDL-size and ASCVD observed in analyses without adjustment for

TG was not significant after adjustment for factors associated with metabolic syndrome such as TG and small and dense LDL. However, two studies with the largest sample size[11,27] did not adjust for chemically measured routine lipids. Improved consistency in confounding adjustment and additional inclusion of traditional lipid measures when available in future studies will provide further insight.

Third, the included studies were mainly from Europe and North America and no study was specifically from Latin America, which might affect the generalizability of these findings. Although about 20% of individuals included in MESA had Hispanic ancestry and about 5% of participants in JUPITER were from Latin America,[16,19] the association of NMR lipid traits and ASCVD remains unknown. As the current largest study with measures of NMR metabolomics in the Latin America population, MCPS provides a unique opportunity to assess the association between lipid traits and ASCVD risk with largest sample size and in greater detail than previously possible.

Fourth, although relatively robust methods were used in these studies, results of observational studies may be biased by measurement error, reverse causality, and by residual or unmeasured confounding, and more importantly, did not prove causality. Future studies using methods such as Mendelian randomization are needed to explore whether observed associations between NMR measured lipid traits and ASCVD are causal. With the availability of both NMR metabolomics and genomic data in MCPS, variations in the concentration and size of lipoprotein particle and their lipids contents among individuals can be assessed to better clarify the potential causal effect of specific lipoprotein components to ASCVD risk.

Fifth, although several studies suggested that the association between NMR lipid traits and cardiovascular events varied by vascular domain (MI, ischemic stroke, intracerebral hemorrhage, and PAD), sex, and ethnicity, few included studies conducted stratified analyses to explore potential heterogeneity. Most of observational studies were conducted in single-cohort studies with small or modest sample sizes assessing solely

CHD, stroke, or composite cardiovascular outcomes across various vascular beds. Moreover, because strokes typically make up relatively few events in a single cohort, previous studies commonly combined ischemic and non-ischemic stroke as a composite end point although recent studies have suggested lipid traits do not have consistent effects on stroke subtypes.[11,26] With its large sample size including both male and female adults of Hispanic ancestry and prolonged follow-up (and hence relatively large number of deaths from different cardiovascular events), MCPS provides a unique opportunity to assess these relationships by sex, vascular domain, and other study characteristics.

Finally, plasma lipid and lipoprotein levels are dynamically and metabolically correlated and statistical models used in current observational studies have limited power to investigate potential effect of these interrelationships and to what extent the observed associations are independent of other NMR lipid traits are unclear. Moreover, including all of the highly correlated lipid measurements into one model would likely introduce multi-collinearity problems, potentially violating the criteria of risk factor selection in multivariable MR analyses (the genetic associations of any risk factors included can not be linearly explained by the genetic associations of any other risk factors or the combination of multiple other risk factors). Future studies using novel methods such as multivariable MR approach based on Bayesian model averaging (MR-BMA) that can detect true causal risk factors even when the candidate risk factors are highly correlated (e.g., metabolites measured by high-throughput NMR spectroscopy).[30]

2.5 Findings of Mendelian randomization studies

2.5.1 Study inclusion

The process of identifying relevant MR studies investigating the association between NMR lipid traits and risk of ASCVD is shown in **Figure 2.2**. Among the 611 records initially retrieved from PubMed and Embase database, 574 non-duplicated titles were identified, leaving 52 abstracts for further screening. Of these, 9 studies were included in the full-text reading list. Finally, 3 eligible MR studies were included in the systematic review,[31-

[33] including one three-sample MR study using univariable and multivariable analyses,[33] one three-sample MR study using MR-BMA,[31] and one traditional two-sample MR study using univariable and multivariable analyses.[32]

The detailed study characteristics are shown in **Table 2.3**. The 3 studies were published between 2020 and 2021 and all employed two-sample or three-sample design. These designs overcome the requirement for assessing the exposure and outcome in the same dataset (one-sample MR), allow the sample size to be maximized and increase the statistical power. Additionally, these designs are less prone to the winner's curse bias than one-sample MR. In these two/three-sample MR studies, genetic associations with blood lipids or CVD were extracted from publicly available summary statistics of GWAS based on participants of primarily European-ancestry (70%-100%). Regarding blood lipid traits, all three studies considered both traditional lipid panels (e.g., LDL-C, HDL-C, and TG) and lipoprotein subfractions measured by NMR spectroscopy. The instrumental variables for major lipoprotein-related traits or traditional lipids (LDL-C, HDL-C, TG, apoB, and apoA1) in three MR studies consisted of independent genetic variants associated with these traits at genome-wide significance level ($p<5\times10^{-8}$ and $r^2<0.001$ or $r^2<0.1$) in the GWAS by GLGC. Regarding the instrumental variables for lipoprotein subfractions, two studies used independent genetic variants associated with NMR lipid-related traits at genome-wide significance level ($p<5\times10^{-8}$ and $r^2<0.001$ or $r^2<0.1$) in the GWAS by Kettunen in univariable MR, two studies used genetic variants associated with traditional lipid traits in GLGC GWAS in multivariable MR-BMA or multivariable MR (one of the three studies used two ways of instrument selection). The genetic associations with major lipoprotein-related traits or traditional lipids were estimated in GWAS by GLGC,[32,33] GWAS by MVP and GLGC,[32] GWAS based on UK Biobank[31] or GWAS by GERA[33] (multiple datasets were used in sensitivity or replication analyses in these MR studies). The genetic associations with lipoprotein subfractions were estimated in GWAS by Kettunen.[34] Particularly, one study repeated analysis for major lipoprotein-related traits using genetic associations from the NMR dataset (Kettunen). Regarding cardiovascular

outcomes, one study assessed CHD (GWAS by CARDIoGRAMplusC4D and UK Biobank) and MI (GWAS by UK Biobank), one study assessed CHD (GWAS by CARDIoGRAMplusC4D) and PAD (GWAS by Klarin), one study assessed small vessel stroke (GWAS by MEGASTROKE) and intracerebral haemorrhage (GWAS by ISGC). Genetic associations of instrumental variables with these outcomes were estimated from corresponding GWAS of primarily European-ancestry participants.

2.5.2 Three sample MR studies using traditional univariable and multivariable analyses

Based on recently available genetic data on traditional lipids, NMR-measured lipoprotein subfractions, and CVD, one three-sample Mendelian randomization examined whether lipoprotein subfractions and particle sizes may have a causal effect on CHD beyond traditional lipid risk factors using both univariable and multivariable analyses.[33] As described in **Table 2.3**, two GWAS of traditional lipid panels (TG, TC, LDL-C, and HDL-C), two recent GWAS of the human lipidome measured by NMR spectroscopy, and three GWAS of CHD or MI were used to investigate lipid-related traits and ASCVD. Because lipid subfractions are largely co-determined with traditional lipid risk factors and do not represent independent biological mechanisms, and may lead to multicollinearity in multivariable MR analyses which violate assumptions of valid instruments. Genetic correlation of 77 subfraction traits and particle size with traditional lipid profile was first estimated by LD score regression in lipidome GWAS and traits that have a high genetic correlation (defined as an estimated genetic correlation with TG, LDL-C, HDL-C, apoB, or apoA1 > 0.8) were excluded, leaving 27 relatively independent lipid traits. In univariable MR, one lipidome GWAS was used to select independent genetic instruments that are associated with one or several lipoprotein measures (a set of SNPs that are in linkage equilibrium with each other in a reference panel [distance >10 megebase pairs, r^2 <0.001], and have a p-value < 5×10^{-8} when IVW and weighted median methods vulnerable to weak instrument bias were used in univariable analyses. Multivariable MR analyses for each lipoprotein subfraction or particle size under investigation were then performed to control for potential pleiotropic effects using two set

of exposures (set 1: TC, LDL-C, HDL-C, and lipoprotein measures; set 2: TC, apoB, apoA1). Both GWAS of traditional lipids and lipidome GWAS were used to select instruments that are associated with one of the four exposures ($p<10^{-4}$ when methods that robust to weak instruments providing Instrument Strength Independent of Direct Effect (InSIDE) assumption are met were used in multivariable MR). Finally, the other two GWAS were used to obtain summary genetic association of these selected SNPs with exposures and outcomes of interest, as in a typical two-sample MR design. For example, multivariable MR used the lipidome GWAS (*Davis* et al., 2017)[35] and GWAS of traditional lipids (*GERA*, Hoffmann et al., 2018)[36] datasets for instrument selection in multivariable analyses, the lipidome GWAS (*Kettunen* et al., 2016)[34] and GWAS of traditional lipids (*GLGC*, Willer et al., 2013)[37] datasets for the associations of selected instruments with the exposures, and the CARDIoGRAMplusC4D + UK Biobank dataset for the associations with ASCVD.

The findings[33] of genetic correlation of lipid traits indicated that VLDL subfraction traits and VLDL particle size were closely related to TG, and large HDL subfractions were related to HDL-C and apoA1. Traits with strong correlation were then excluded. Regarding the remaining 27 traits related to the groups of VLDL/LDL/IDL particles or the groups of HDL particles, most traits were strongly correlated with each others within each group. Specifically, most traits in the first group (VLDL/LDL/IDL particles) were positively associated with LDL-C and apoB and those in the second group (HDL particles) were associated with HDL-C and apoA1, which is consistent with epidemiological correlation estimated in observational studies.

The concentration and lipid content of apoB-containing particle subfractions (VLDL, IDL, and LDL) and LDL particle size were uniformly associated with increased risk of MI in univariable analyses,[33] however, these effects were attenuated and close to zero after adjustment for traditional lipids in multivariable analyses. By contrast, the concentration and content of medium and small HDL particles were inversely associated with MI in both univariable and multivariable MR and the effect estimates did not attenuate after

adjusting for traditional lipids. HDL particle size had no effect on MI in univariable MR but were significantly associated with the risk of MI after adjustment. Traditional HDL-C/apoA1 had a weak, non-significant protective effect on CHD, which is consistent with previous MR study on traditional lipids.[38] Overall, these results were generally consistent when the different GWAS datasets were used or when multiple statistical methods (e.g. IVW, weighted median, and RAPS) were used in sensitivity analyses.

2.5.3 Three sample MR studies using MR-BMA analyses

One MR study using a Bayesian model averaging approach (MR-BMA) investigated the association of traditional lipid traits and apoB-containing lipoprotein subfractions with PAD and CHD.[31] MR-BMA is an extension of two-sample multivariable MR that applies a Bayesian variable selection method and aims to identify causal risk factors (rather than the magnitude of causal effect) from a set of potentially highly correlated candidate risk factors (e.g. lipoprotein traits measured by NMR spectroscopy). By including multiple risk factors into one single model, multivariable MR-BMA uses instruments associated with multiple risk factors and thus allows for measured pleiotropy (the effects of genetic variants on other risk factors on competing causal pathways), which is different from the assumption of no pleiotropic effect in standard univariable MR. The causal effect estimates from multivariable MR represents the *direct causal effect* of risk factors (risk factors most proximal to an outcome when there are mediating effects between risk factors) on outcome of interest when other risk factors in the model are held constant and not the *total causal effect* as estimated in univariable MR. Effect estimates from MR-BMA will be conservatively biased towards the null when there is a causal effect but the variance of the estimates will be reduced. Thus, MR-BMA designed for high-dimensional data could provide unbiased estimates but with larger variance.

The instruments used in this study consisted of 145 independent genetic variants associated with at least one of the traditional lipid traits (TG, TC, LDL-C, and HDL-C) at a genome-wide significance level ($p<5\times10^{-8}$ and $r^2<0.001$ in the 1000 Genomes

European Reference Panel) in a GWAS of primarily European-ancestry participants by the GLGC. The exposures of interest in this study included both major lipid-related traits (TG, LDL-C, HDL-C, apoA1, and apoB) and apoB-containing lipoprotein subfractions. Genetic associations with major lipid-related traits were estimated in the UK Biobank study based on about 360,000 European-ancestry participants. Genetic associations with lipoprotein subfractions were estimated in a lipidome GWAS (Kettunen et al., 2016) based on about 25000 European-ancestry participants. Genetic associations with CHD were estimated in GWAS by CARDIoGRAMplusC4D based on about 185,000 participants of primarily European-ancestry (77%) and genetic associations with PAD were estimated in GWAS of about 240,000 multiethnic participants (6% were PAD cases). In the replication analysis, the analysis for major lipid-related traits was repeated using genetic association with NMR dataset. To predict the impact of apoB-lowering on PAD and CHD risk, MR analyses using variants associated with apoB in gene regions that proxy lipid-lowering therapies were also conducted. Independent genetic variants associated with apoB levels at genome-wide significance ($p<5\times10^{-8}$ and $r^2<0.1$) were identified from the GWAS in UK Biobank. Genetic associations with apoB were estimated in NMR GWAS to avoid winner's curse and sample overlap.

In the analysis for major lipid-related traits, apoB was identified as a prioritized risk factor for both PAD (marginal inclusion probability [MIP], defined as the sum of posterior probability over all models where the risk factor present, 0.86; FDR-corrected p=0.003) and CHD (MIP, 0.92; p=0.005) in the primary and replication analyses, which is consistent with the role of apoB in ASCVD previously identified in observational studies (as discussed in section **2.4**). In the subsequent analysis of lipoprotein subfractions, XS.VLDL.P (MIP, 0.91; $p=2.3\times10^{-4}$) and L.LDL.P (MIP, 0.95; p=0.011) were prioritized as the primary apoB-containing risk factors for PAD and CHD, respectively. In the analysis for apoB-lowering effects, genetically proxied apoB lowering was associated with decreased risk of PAD (OR for 1 SD reduction in apoB: 0.87, 0.84-0.91) and CHD (0.66, 0.63-0.69). The effect of apoB-lowering on CHD was greater than PAD (ratio of effects:

3.09, 2.29-4.60). Overall, although apoB was identified as a major lipid-related trait causally related to both CHD and PAD, apoB-containing subfractions had diverse associations with ASCVD and genetically proxied apoB-lowering had a stronger effect on CHD, suggesting potential differences in the role of apoB-containing lipoprotein subfractions in the development of PAD and CHD.

2.5.3 Two sample MR studies using traditional univariable and multivariable analyses

One MR study using a traditional two sample MR design investigated the association of traditional lipid traits and lipoprotein subfractions with small vessel disease (SVD) including small vessel stroke (SVS) and intracerebral haemorrhage (ICH).[32] Genetic instruments associated with traditional lipids were selected from GWAS of the Million Veteran Program (MVP) and GLGC GWAS, and those associated with NMR lipid traits were selected from Kettunen GWAS as in above two MR studies ($p<5\times10^{-8}$ and $r^2<0.1$). Instruments associated with SVS were selected from the GWAS by the MEGASTROKE Consortium while those associated with ICH were selected from the GWAS by the International Stroke Genetics Consortium (ISGC). This study explored the relationship of genetically determined composite HDL-C, LDL-C, and TG with SVD and the relationship of genetically determined lipoprotein subfractions with SVD. Moreover, the effects of genetic predisposition to HDL-C raising, LDL-C lowering, and TG lowering through variants in gene encoding targets of lipid-modifying therapies on SVD were investigated.

Regarding the association between traditional lipids and SVS, HDL-C was associated with a lower risk of SVS in univariable analyses and remained significant in multivariable analyses. In contrast, the positive associations between TG and LDL-C and SVS were attenuated after adjusting for HDL-C or LDL-C/TG (mutually adjusted) in multivariable analyses. Specifically, cholesterol in medium-sized but not large- or extra-large sized HDL-P was associated with lower SVS risk after adjusting for LDL-C and TG. Similar associations between any of the medium-sized HDL-P components (TC, CE, FC,

phospholipids, and total lipids) and SVS risk were observed, suggesting the associations were driven by medium-sized HDL particle as a whole. Regarding other lipoprotein subfractions, cholesterol in all LDL-C, IDL, VLDL particle (except large VLDL) were associated with increased SVS risk after adjustment for TG and HDL-C. TG in small HDL-P but not larger HDL-P were associated with higher risk of SVS. TG in apoB-containing particles (except IDL) was not associated with SVS in either univariable or in multivariable analyses that further adjusted for LDL-C and HDL-C. Genetically proxied HDL-C-raising was associated with lower risk of SVS in analyses using variants associated with HDL-C in the *CETP* locus. While genetic predisposition to LDL-C lowering through *CETP* variants were associated with lower risk of SVS, this effect was entirely reversed after adjusting for the HDL-C altering effects of the variants in multivariable MR. Analyses of genetic variants in LDL-C lowering drug target loci showed no statistically significant results.

Regarding ICH, genetically determined HDL-C, LDL-C, and TG were not associated with ICH risk in either univariable or multivariable analyses. Associations of lipoprotein subfractions with ICH were in an opposite direction as compared with SVS. HDL-C raising variants in the *CETP* locus were associated with a higher risk of ICH.

2.5.4 Gap in current evidence from MR studies and implications for future study

Existing MR studies identified in the systematic review used large genetic datasets to explore potential causal relationships of NMR-derived lipoprotein particle subfractions with a series of key manifestations of ASCVD (e.g. CHD, MI, PAD, SVS, and ICH) after adjustment for traditional lipids. Evidence from MR studies should be interpreted with caution considering its limitations and gaps. First, although multiple tests for the detection of directional pleiotropy and sensitivity analyses provided consistent results in these MR studies, the possibility of bias in MR analyses cannot be ruled out and should be considered in future studies, especially considering that genes that regulate lipoprotein metabolism are rarely specific for one single lipid measure or a single lipoprotein. Second,

both lipid exposures and ASCVD outcomes were studied among participants of primarily European-ancestry, and although 24% to 30% of participants in several multiethnic GWAS were of diverse ancestries, the effect of lipoprotein-related traits may vary across ancestry. Future studies among other populations are therefore warranted to improve the generalizability of current evidence from MR studies. The availability of both genetic and phenotypic data in MCPS will facilitate MR analyses on these associations in a large Latin American population. Third, MR effect estimates reflect the lifetime effect of genetic determinants of blood lipid levels and may not represent the magnitude of effects of short-term lipid-modifying interventions. Again, although no pleiotropic effects were detected in these drug target MR analyses, the estimates of genetic predisposition to lipid-modifying may be influenced by pleiotropy (the possibility that genetic variants located within drug target may influence other cardiometabolic traits). Thus, evidence from drug target MR studies should not be used to directly guide clinical decisions on lipid-modifying therapies. Fourth, estimates for the same lipoprotein subfraction derived from univariable and multivariable MR analyses may vary, reflecting correlated lipid traits within a common pathway. However, not all studies conducted both analyses and one study only performed univariable analyses. Estimates from univariable MR reflect the total effect of an exposure on the outcome but these from multivariable MR identifies the most proximate and direct risk factor to outcome. Notably, in one study using a multivariable framework method (MR-BMA), apoB was prioritized as a risk factor for PAD and CHD, which may give the impression that LDL-C is not a causal risk factor. However, by comparing the estimates from both MR analyses across different studies, this finding does not nullify the causal effect of LDL-C but indicates that the effects of lipid-lowering therapies are likely to be proportional to the change in apoB rather than LDL-C. Thus, future MR studies of detailed lipid-related traits should use both univariable and multivariable analyses to strengthen causal inference. Last, instrument selection in two MR studies was based solely on a GWAS of NMR metabolites in 24,925 European-descent participants published in 2016. Although it remains the current largest GWAS of

circulating lipoproteins and subfractions, the sample size of this GWAS is still relatively modest compared with large disease consortia studies and may not have sufficient statistical power to identify significant variants associated with lipoprotein subfractions. Larger GWASs of NMR lipid traits in diverse populations may facilitate the development of more robust genetic instruments (e.g., the ongoing GWAS of NMR metabolites in MCPS), which may provide further insights into the relationship between lipid-related traits and ASCVD.

2.6 Findings of genome-wide association studies

2.6.1 Study inclusion

The process of identifying relevant GWAS of NMR lipoproteins and subfractions is shown in **Figure 2.3**. Among the 322 records initially retrieved from PubMed and Embase database, 310 non-duplicated titles were identified, leaving 31 abstracts for further screening. Of these, 3 studies were included in the full-text reading list. One additional study from references was also included. Finally, 4 eligible GWAS of NMR lipid traits were included in the systematic review. 4 representative GWAS of traditional lipids by large lipid consortium were included in the review by manual search of GWAS catalog and relevant MR study of lipids and ASCVD. The detailed study characteristics are shown in **Table 2.4**.

2.6.2 Main findings, current gap, and implications for future GWAS

Plasma concentrations of lipid and lipoproteins are heritable risk factors for ASCVD. Genetic studies of plasma lipids can improve the understanding of the genetic factors involved in lipid metabolism and particularly identify potential therapeutic targets for cholesterol management and prevention of CVD. A meta-analysis of 23 studies of 188,578 individuals with genotype data in the GLGC (the GWAS most widely used in MR studies of traditional lipids) previously identified 157 loci associated with levels of HDL-C, LDL-C, TG, and TC, accounting for ~10-12% of total variance.[37] Subsequent GWAS of traditional lipids found additional significant associations, totaling approximately 386

loci. Compared with traditional lipids, lipoproteins and subfractions measured by NMR spectroscopy could be more informative and may serve as good intermediate traits for a GWAS to understand the genetic determinants of full spectrum of circulating lipid traits. With the availability of advanced NMR spectroscopy that can provide hundreds of quantitative metabolic measures from large epidemiological sample collections, a few studies have reported genetic associations for NMR lipid traits either through studies on subsets of lipoprotein subfractions[35,39,40] or GWAS on metabolome[34]. A GWAS meta-analysis of 14 studies including 24,925 individuals with the NMR measures of 123 metabolic traits identified 60 significant loci and found strong associations of *LPA* SNPs with VLDL and TG metabolism.[34] This study was the most widely used GWAS in MR studies of NMR lipid traits. For the majority of the lipid loci, however, their effects on detailed lipoprotein subfractions beyond traditional lipids are unknown.

Substantial gaps and limitations remain in lipid GWAS. First, these GWAS were conducted in individuals of mostly European ancestry, with few studies including other ethnic groups (despite evidence showing that most genetic variation in humans is population-specific and only seen in non-European samples). In the largest GWAS of traditional lipids combining 312,571 Million Veteran Program (MVP) participants and existing data from GLGC, although some variants were shared among the African, European and Hispanic ethnicities, novel coding variants were associated with lipid levels only in the non-European population. Future studies that examine genomic data and NMR lipid traits across diverse ancestries are warranted. Second, these GWAS mainly focused on traditional lipid panels rather than an NMR lipid profile, failing to delineate the contribution of functionally diverse lipoprotein subfractions to the risk of cardiometabolic diseases. As discussed in sections **2.4** and **2.5**, lipoprotein subfractions have been found to be independently and specifically associated with different manifestations of ASCVD and may potentially improve ASCVD risk assessment over traditional lipids. Moreover, previous study suggested that GWAS of lipidomic profiling identified additional variants influencing lipid metabolism that GWAS of traditional lipids

failed to detect. In light of the limited information about the genetics of detailed lipidomic profiles, GWAS emphasizing on lipoprotein subfractions are needed. Third, previous large-scale GWAS for blood lipids are generally meta-analyses of many studies with small sample size, rather than using a single large cohort. Thus, direct conditional analyses for additional independent significant SNPs at each locus have been much smaller.[36] Moreover, the sample size of GWAS of lipoprotein subfractions have only reached several thousand, which is still modest compared with GWAS of traditional lipids and GWAS of chronic diseases and results in a limited number of genetic instruments available for MR analyses. Future studies based on single large cohorts would enhance discovery power and enable novel insights. Forth, estimates of variance explained by detected lipid loci ranged from 11%-23% among traditional lipids, which are substantially below family-based heritability estimates (33%-51%). This suggests that many more lipid-associated variants remain undiscovered. Overall, ongoing GWAS of a single large cohort of about 150,000 Latin American individuals in MCPS using measures of NMR metabolites would help to overcome the challenge of small sample size and help fill a current gap, facilitate the identification of genetic variants influencing the levels of lipoprotein subfractions, and enhance the current understanding of genetic regulation of blood lipids.

2.7 Conclusions

Observational studies consistently show that concentrations of apoB-containing particles are linked to the risk of various ASCVD outcomes. Further MR studies also identify apoB as a causal risk factor. Most studies indicated an inverse relationship between HDL particle concentration and size and risk, though some report no association. For TG and cholesterol, previous studies suggest positive associations with risk, but effects vary by particle size. These associations often diminish after adjusting for apoA1 and apoB. Overall, these NMR-based findings align with current lipid and ASCVD understanding based on chemistry assays. Gaps in current research were highlighted, such as methodological limitations and the lack of evidence from Latin American populations. The

MCPS data offer a unique opportunity to address these gaps and enhance understanding of lipid metabolism and its relevance to ASCVD.

Table 2.1 Search terms used for literature review of observational or MR studies

Search criteria	Search terms (MeSH terms and text words)
Lipid related traits measured by NMR	Nuclear magnetic resonance spectroscopy
Cardiovascular disease	Cardiovascular disease OR heart infarction OR coronary heart disease OR ischemic heart disease OR myocardial infarction OR coronary heart disease OR arterial disease OR vascular disease OR coronary event OR ischemic event OR coronary OR ischemia OR cardiac arrest OR atherosclerotic disease OR atherosclerosis OR atherosclerotic heart disease OR acute coronary syndrome OR coronary arterial disease OR Stroke
Study design	**Search 1:** Prospective OR longitudinal OR cohort OR follow up OR retrospective study OR observational OR case control OR case cohort **Search 2:** Mendelian randomization OR Mendelian randomization OR genetic OR gene OR allele OR genetics OR causal variant OR variant OR single nucleotide polymorphism OR SNP OR genetic association OR loci OR locus OR genome OR instrumental variable analysis
Time period	From inception through 2022
Source	PubMed and EMBASE databases

Table 2.2 Search terms used for literature review of GWAS of NMR lipid related traits

Search criteria	Search terms (MeSH terms and text words)
Lipid related traits measured by NMR	Nuclear magnetic resonance spectroscopy
Study design	**Search 3**: Genome wide association study OR Genome-wide study OR GWAS OR single nucleotide polymorphism OR SNP OR locus OR loci
Time period	From 2007 through 2022
Source	PubMed and EMBASE databases

Table 2.2 Characteristics of included observational studies investigating the association between NMR lipid traits and the risk of ASCVD

Study (Country/Consortium)	Study design	Participants, N (% men)/ CVD events	Baseline age, Y[#]	Recruitment period (Follow-up)	NMR lipid-related traits (Platform)	Exclusions	Adjustments[*]	Risk estimates, HR/OR/RR (95% CI)
Case-control studies								
Blake, 2002[3] (American/ Women's Health Study)	Nested case-control	260 (0%) /65 MI, 53 stroke, and 12 fatal CHD (130)	60 (8.6)	1993-1996 (3 Y)	LDL concentration and particle size (LipoScience)	NA (participants without history of CVD and cancer at baseline)	Chemically measured TC:HDL-C ratio and TG, history of diabetes and hypertension	*Highest vs. lowest quartile:* LDL-P: 2.90 (1.16-7.30) LDL-P size: 1.20 (0.51-2.82)
Kuller, 2002[4] (American/ Cardiovascular Health study)	Case-control	683 (45%) /434 MI and angina	73 (5.5)	1990-1995	Concentrations of 10 subclasses of VLDL, LDL, and HDL, and LDL particle size (LipoScience)	Excluding those with clinical or subclinical diseases in the control group	Chemically measured LDL-C, age, and race	*Highest vs. lowest quartile* by sex: Men: Not significant for NMR VLDL, LDL-P, LDL-C, and HDL Women: Not significant for HDL-P and LDL-

								P size after adjustment; LDL-P: 3.34 (1.50-6.50) Large VLDL-P: 3.26 (1.66-6.40)
Soedamah-Muthu,2003[5] (American)	Nested case-control	118 with Diabetes (48%)/59 MI, angina, and coronary stenosis	35 (6.0)	1950-1980	Concentrations of 9 subclasses of VLDL, LDL, and HDL (LipoScience)	Excluding individuals with TG > 4.5 mmol/L	Chemically measured LDL-C, TG, and HDL-C	*1 SD increment:* Medium HDL: 3.79 (1.17-12.25) Large HDL-P: 0.43 (0.20-0.92) VLDL-P: 2.33 (1.07-5.08) LDL-P and other particles: not significant
El Harchaoui, 2007[6] (UK/EPIC-Norfolk study)	Nested case-control	2888 (63%)/1003 CHD	65 (8.0)	1993-1997 (6 Y)	Concentration and size of LDL-P subclasses (LipoScience)	Excluding individuals with history of heart attack or stroke	Chemically measured LDL-C, TG, and HDL-C	*Highest vs. lowest quartile:* LDL-P: 1.37 (1.04-1.83) LDL-P size adjusting for LDL-P: 1.22 (0.92-1.61)

Van der steeg, 2008[7] (UK/EPIC-Norfolk study)	Nested case-control	2349 (63%)/858 CHD	65 (8.0)	1993-1997 (6 Y)	Particle size of HDL (LipoScience)	Excluding individuals with history of heart attack or stroke	Chemically measured TC, TG, HDL-C, apoA-I, and apoB	After adjustment for apoA-I and apoB, HDL-P size (>10.07 nm vs <8.60 nm): 3.49 (1.37-8.89) (9.85–10.07 nm vs <8.60nm): 2.32 (1.12-4.77) Unadjusted: 0.78 (0.41-1.47)
El Harchaoui, 2009[8] (UK/EPIC-Norfolk study)	Nested case-control	2223 (63%)/822 CHD	65 (8.0)	1993-1997 (6 Y)	Concentration and size of HDL-P (LipoScience)	Excluding individuals with history of heart attack or stroke	Chemically measured TC, TG, HDL-C, apoA-I, and apoB	*Highest vs. lowest quartile:* Adjustment for TG and apoB, HDL-P size: 1.00 (0.71-1.39) HDL-P: 0.50 (0.37–0.66) Large-HDL-P: 0.85 (0.60-1.20)
Arsenault, 2010[9] (UK/EPIC-Norfolk study)	Nested case-control	2529 (63%)/870 CHD	65 (8.0)	1993-1997 (6 Y)	Ratio of LDL-P/HDL-P (LipoScience)	Excluding individuals with history of	Chemically measured TC, TG, HDL-C,	*Highest vs. lowest quartile:* LDL-P/HDL-P: 2.78 (1.86-4.17)

						heart attack or stroke	apoA-I, and apoB	
Van Capelleveen, 2017[10] (UK/EPIC-Norfolk study)	Nested case-control	2223 (63%)/822 CHD	65 (8.0)	1993-1997 (6 Y)	ApoC-III (LipoScience)	Excluding individuals with history of heart attack or stroke	Chemically measured TC, TG, HDL-C, apoA-I, and apoB	*Highest vs. lowest quartile:* ApoC-III: 1.91 (1.48-2.48)
Holmes, 2018[11] (China/China Kadoorie Biobank)	Nested case-control	4662 (48%)/912 MI, 1146 ischemic stroke, 1138 intracerebral hemorrhage	45 (8)	2004-2008 (9 Y)	NMR lipid, 14 lipoprotein subclasses, and apolipoprotein (Nightingale)	Excluding individuals with history of CHD, stroke, cancer, and those using of statin therapy	Traditional non-lipid CVD risk factors	*1 SD increment:* Lipoproteins and lipids showed similar associations with MI and IS, but not with ICH
Salahuddin, 2020[12] (International/dal-Outcomes trial)	Case-cohort	1378 (80%)/476 CHD, stroke, and angina	64 (10)	- (3 Y)	HDL-P subclasses (LipoScience)	Excluding individuals with TG >400 mg/dL	LDL-C, TG, LDL-P, and VLDL-P	*1 SD increment:* HDL-P: 0.98 (0.84-1.15) HDL-C and HDL subclasses: not significant
Cohort studies								
Cromwell, 2007[13]	Cohort	3066 (47%)/531 CVD	51 (10)	1987-1991 (15 Y)	LDL-P and VLDL-P concentration	Excluding individuals with prevalent	Chemically measured	*1 SD increment:* LDL-P: 1.28 (1.17-1.39)

(American, Framingham Offspring cohort)					(LipoScience)	CVD and TG >400 mg/dL at baseline	LDL-C and non-HDL-C	LDL-P + VLDL-P: 1.28 (1.18-1.40)
Mora, 2009[14] (American, Women's Health Study)	Cohort	27673 (0%)/1015 CVD	55 (7)	1992-1995 (11 Y)	LDL-P, VLDL-P, and HDL-P subclasses (LipoScience)	NA (participants without history of CVD and cancer at baseline)	Chemically measured TC, LDL-C, HDL-C, TG, $apoB_{100}$, and apoA-I	*Highest vs. lowest quintile:* LDL-P: 2.51 (1.91-3.30) Large LDL-P: 0.86 (0.72-1.03) HDL-P: 0.91 (0.75-1.12) Large HDL-P: 0.58 (0.48-0.71) VLDL-P: 1.71 (1.38-2.12) LDL-P size: 0.64 (0.52-0.79) HDL-P size: 0.65 (0.51-0.81) VLDL-P size: 1.37 (1.10-1.70)
Aday, 2018[23] (American, Women's Health Study)	Cohort	27888 (0%)/110 PAD	55 (7)	1992-1995 (15 Y)	LDL-P, VLDL-P, and HDL-P	NA (participants without history	Chemically measured TC, LDL-C,	*Highest vs. lowest tertile:*

					subclasses (LipoScience)	of CVD, PAD, and cancer at baseline)	HDL-C, TG, $apoB_{100}$, and apoA-I	LDL-P: 2.03 (1.14-3.59) Large LDL-P: 0.74 (0.44-1.27) Small LDL-P: 2.17 (1.10-4.27) HDL-P: 0.29 (0.16-0.52) Large HDL-P: 0.44 (0.25-0.77) VLDL-P: 1.39 (0.81-2.38) LDL-P size: 0.60 (0.36-1.02) HDL-P size: 0.39 (0.21-0.70) VLDL-P size: 1.22 (0.73-2.04)
Mackey, 2012[15] (American, MESA)	Cohort	5598 (47%)/227 CHD	62 (10)	2000-2002 (6 Y)	HDL-P and LDL-P (LipoScience)	Excluding individuals with prevalent CVD and cancer, TG >400 mg/dL, and	Chemically measured LDL-C, HDL-C, and TG	*Highest vs. lowest quartile:* HDL-P: 0.49 (0.27-0.86) *1 SD increment:* HDL-P: 0.68 (0.54-0.85)

Tsai, 2014[17] (American, MESA)	Cohort	4387 (47%)/234 CHD	62 (10)	2000-2002 (6 Y)	Small LDL-P (LipoScience)	the use of lipid-lowering therapy	Chemically measured HDL-C and TG	*Highest vs. lowest quartile:* Small LDL-P: 1.37 (0.71-2.64)
Steffen, 2015[19] (American, MESA)	Cohort	4679 (47%)/233 CHD	62 (10)	2000-2002 (6 Y)	LDL-P and HDL-P (LipoScience)		Chemically measured TC, LDL-C, HDL-C, and TG	*Highest vs. lowest quartile:* Adjustment for LDL-C: LDL-P: 1.75 (1.11-2.75) LDL-P/HDL-P: 2.30 (1.49-3.57) Adjustment for non-HDL-C: LDL-P: 1.60 (0.98-2.62) LDL-P/HDL-P: 2.25 (1.43-3.55)
Mora, 2013[16] (American, JUPITER)	Cohort	10886 (62%)/ 234 CHD	66 [60-71]	2003-2006 (2 Y)	Concentration and size of HDL-P (LipoScience)	Excluding those with use of lipid-lowering therapy (participants	Chemically measured LDL-C, HDL-C, and TG	Placebo group (Baseline): *Highest vs. lowest tertile:* HDL-P: 0.68 (0.43-1.08)

						without history of CVD and diabetes at baseline)		HDL-P size: 1.06 (0.65-1.73) *1 SD increment:* HDL-P: 0.81 (0.67-0.97) HDL-P size: 0.94 (0.77-1.15)
Akinkuolie, 2015[18] (American, JUPITER)	Cohort	12487(62%)/308 CHD	66 [60-71]	2003-2006 (2 Y)	HDL-P subclasses (LipoScience)		Traditional non-lipid CVD risk factors	Baseline: Large HDL-P: 0.85 (0.76-0.95) Medium HDL-P:0.89 (0.80-0.98) On-treatment: Very large HDL-P: 1.65 (1.02-2.66)
Lawler, 2017[21] (American, JUPITER)	Cohort	11984 (64%)/298 CHD	66 [60-71]	2003-2006 (2 Y)	Concentration and size of LDL-P and VLDL-P (LipoScience)		Chemically measured HDL-C	Baseline: LDL-P: 1.19 (1.02-1.38) LDL-P size: 0.87 (0.75-1.00) VLDL-P: 1.21 (1.04-1.41)

								VLDL-P size: 1.10 (0.95-1.27)
Würtz, 2015[20] (Finland, FINRISK 1997 study)	Cohort	7256 (49%)/800 CVD	48 (13)	1997 (15 Y)	NMR lipid, 14 lipoprotein subclasses, and apolipoprotein (Nightingale)	Excluding individuals with prevalent CVD	Chemically measured TC and HDL-C	*1 SD increment:* Adjustment for age and sex: ApoB: 1.26 (1.16-1.36) ApoA-I: 0.85 (0.79-0.91) TG: 1.33 (1.22-1.43) TC: 1.11 (1.02-1.20) HDL-C: 0.77 (0.72-0.83) HDL-P size: 0.86 (0.79-0.92) LDL-P size: 1.09 (1.01-1.17) VLDL-P size: 1.18 (1.09-1.26) Total lipid in 14 lipoprotein subclasses:

								Significant for apoB-containing particles; Very large HDL-P: 0.94 (0.88-1.01) Large HDL-P: 0.73 (0.68-0.79) Medium HDL-P: 0.86 (0.80-0.92) Small HDL-P: 0.92 (0.86-0.99)
Silbernagel, 2017[22] (Germany)	Cohort	510 (69%)/47 CVD	63 (10)	1997-2000 (10)	Concentration and size of HDL-P (Numares AG)	Excluding individuals with chronic diseases	Chemically measured LDL-C and TG	*Highest vs. lowest tertile:* HDL-P: 0.37 (0.18-0.77) Small HDL-P: 0.56 (0.27-1.16)
Akbaraly, 2018[24] (UK, Whitehall II study)	Cohort	4824 (72%)/697 CVD	60 (6)	1997-1999 (16)	VLDL and HDL particle size, and total lipid in lipoprotein subclasses (Nightingale)	Excluding individuals with CVD, cancer, and other reported chronic diseases	-	*1 SD increment:* VLDL-P: 1.12 (1.03-1.21) HDL-P size: 0.86 (0.79-0.94) Significant for total lipids in

								VLDL-P, IDL-P, and LDL-P
Joshi, 2020[25] (UK, UCLEB) Consortium)	Cohort	11560 (66%)/1613 CVD	59 (9)	1998-2000 (7)	Triglyceride in 14 lipoprotein subclasses, apoA-I, and apoB (Nightingale)	NA (participants without ASCVD at baseline)	Chemically measured LDL-C and HDL-C	*1 SD increment:* For TG in: Large VLDL-P: 1.12 (1.03-1.21) IDL-P: 1.21 (1.12-1.30) Small HDL-P: 1.19 (1.09-1.29) Large HDL-P: 0.98 (0.88-1.08) Not significant for TG in 14 subclasses after adjustment for LDL-C and HDL-C (or apoA-I and apoB)
Singh, 2020[26] (Meta-analysis of 4 cohorts: MESA, DHS, ARIC, and PREVEND)	Cohort (meta-analysis)	15784 (46%)/515 MI and 321 IS	57 (13)	2000-2002/ 1997-1998 (8-12)	Concentration and size of HDL-P (LipoScience)	NA (participants without ASCVD at baseline)	Chemically measured LDL-C, HDL-C, and TG	*Highest vs. lowest quartil*e: HDL-P: ASCVD: 0.64 (0.53-0.78)

								Stroke and MI: 0.69 (0.54-0.88) MI: 0.64 (0.47-0.87) Stroke: 0.75 (0.51-1.12) HDL-P size (ASCVD): 0.91 (0.77-1.09)
Vojinovic, 2020[27] (Meta-analysis of 7 cohorts: Rotterdam, Whitehall II, FINRISK97, DILGOM, PROSPER, EGCUT, and FHS)	Cohort (meta-analysis)	38797 (37%-75%) /1791 stroke	47-77	- (2-15)	14 lipoprotein subclasses and their lipids (LipoScience in FHS and Nightingale in others)	Excluding participants with a history of stroke at baseline	-	*1 SD increment:* HDL-P size: 0.92 (0.87-0.98) HDL-P-C:0.92 (0.87-0.97) Large LDL-P-TG: 1.09 (1.03-1.14) Medium LDL-P-TG: 1.09 (1.03-1.14) Small LDL-P-TG: 1.07 (1.01-1.12) Very large VLDL-P-TG: 1.05 (1.00-1.10)

								Not significant for other lipoprotein subclasses
Johansen, 2021[28] (Denmark, Copenhagen General Population Study)	Cohort	29039 (48%)/2309 MI	62 [51-72]	2003-2015 (10)	ApoB-containing lipoprotein subclasses and their lipids (Nightingale)	NA (participants without MI at baseline)	Chemically measured HDL-C and TG	*1 SD increment:* Significant for all VLDL-P subclasses; IDL-P: 1.16 (1.11-1.21) L-LDL-P: 1.16 (1.11-1.21) M-LDL-P: 1.16 (1.11-1.22) S-LDL-P: 1.16 (1.11-1.21)

Abbreviations: Y, years; LDL-P, LDL particle; MI, Myocardial Infarction; EPIC, European Prospective Investigation into Cancer and Nutrition; IDEAL, Incremental Decrease in End Points through Aggressive Lipid Lowering; IS, ischemic stroke; ICH, intracerebral hemorrhage; NA, not applicable; MESA, Multi-Ethnic Study of Atherosclerosis; JUPITER, Justification for the Use of statins in Prevention: an Intervention Trial Evaluating Rosuvastatin; PAD, peripheral artery disease; UCLEB, University College London–Edinburgh–Bristol; DHS, Dallas Heart Study; ARIC, Atherosclerosis Risk in Communities Study; PREVEND, Prevention of Renal and Vascular End-stage Disease; DILGOM, Dietary, Lifestyle and Genetic Determinants of Obesity and Metabolic Syndrome; PROSPER, Prospective Study of Pravastatin in the Elderly at Risk; EGCUT, Estonian biobank; FHS, Framingham Heart Study;

[#]: Mean (SD)/Median[IQR]/range

[*]: Only lipid-related covariates were listed because almost all of studies adjusted for non-lipid risk factors of CVD

Table 2.3 Characteristics of included Mendelian randomization studies investigating the association between NMR lipid traits and the risk of ASCVD

Study	Study design	Study population (GWAS consortium/first author of the GWAS)	Sample size	Ancestry	NMR lipid-related traits (Platform)	Instruments	Adjustments	Risk estimates, HR/OR/RR (95% CI)
Zhao, 2021[33]	Three sample (univariable and multivariable MR)	Traditional lipid: GERA GLGC NMR lipid subfraction: Davis Kettunen CVD outcomes: CARDIoGRAMplusC4D (CAD) CARDIoGRAMplusC4D+UK Biobank (CAD) UK Biobank (MI)	 94674 188578 8372 24925 185000 - 360420	 Multiethnic European Finnish European Mostly European Mostly European European	Kettunen GWAS: 14 lipoprotein subclasses and there lipid content, and particle size (Nightingale)	For univariable MR: 1: Using GWAS for traditional lipid traits to select SNPs for lipid subfraction traits (e.g. HDL-C was used to select SNPs for HDL subclasses and particle size) 2: Using GWAS for NMR lipid traits to select SNPs for lipid	In multivariable MR: Set 1: TG, HDL-C, and LDL-C Set 2: TG, apoA1, and apoB	Significant in both univariable and multivariable MR (RAPS): *1 SD increment:* M-HDL-P: 0.62 (0.45-0.86) M-HDL-L: 0.62 (0.45-0.84) Significant in multivariable MR: *1 SD increment:* HDL-D: 1.34 (1.10-1.64) S-HDL-L: 0.80 (0.67-0.95) S-HDL-P: 0.79 (0.67-0.93)

						subfraction traits For multivariable MR: Using both GWAS to select SNPs associated ($P < 10^{-4}$) with at least one of the TG/LDL-C/HDL-C/subfraction		
Levin, 2021[31]	Three sample (multivariable MR-BMA)	Traditional lipid: GLGC UK Biobank NMR lipid subfraction: Kettunen CVD outcomes: CAD: CARDIoGRAMplusC4D PAD: Klarin	188578 361194 24925 185000 243060	European European European Mostly European (77%) Multiethnic	Kettunen GWAS: 14 lipoprotein subclasses and their lipid content, and particle size (Nightingale)	145 independent genetic variants associated with one of traditional lipid traits (TG, TC, LDL-C, and HDL-C) at a genome-wide	Mutually adjusted in MR-BMA: For traditional lipids: apoB, apoA1, LDL-C, HDL-C, TG. For NMR lipid traits: 10	For major lipid traits: apoB for PAD and CHD For lipoprotein subfractions: XS.VLDL.P for PAD and L.LDL.P for CHD

						significance level ($p<5\times10^{-8}$ and $r^2<0.001$)	apoB particle subfractions	*1 SD reduction* in apoB: PAD: 0.87 (0.84-0.91) CHD: 0.66 (0.63-0.69)
Georgakis, 2020[32]	Two sample (univariable and multivariable MR)	Traditional lipid: Klarin (MVP&GLGC)	617303	Multiethnic	Kettunen GWAS: 14 lipoprotein subclasses and their lipid content, and particle size (Nightingale)	For traditional lipid panel: genetic variants associated with one of traditional lipid traits (253 for TG, 219 for LDL-C, and 312 HDL-C) at a genome-wide significance level ($p<5\times10^{-8}$) For lipoprotein subfractions:	For traditional lipids: HDL-C, LDL-C, and TG; For lipoprotein subfractions: adjusted for other traditional lipids (e.g. for cholesterol in different HDL-P adjusted for LDL-C and TG)	*1 SD increment:* SVS: Traditional lipids: Univariable MR (IVW): HDL-C: 0.85 (0.78-0.92) LDL-C: 1.09 (1.01-1.17) TG: 1.14 (1.04-1.24) Multivariable MR (IVW): HDL-C: 0.89 (0.83-0.96) LDL-C: 1.08 (0.99-1.17)
		GLGC	188577	Multiethnic				
		NMR lipid subfraction: Kettunen	24925	European				
		Cerebral small vessel disease: Small vessel stroke: MEGASTROKE	298777	Multiethnic				
		ICH: ISGC	3026	European				

					genetic variants associated with specific subfraction at a genome-wide significance level ($p<5\times10^{-8}$ $r^2<0.1$)	TG: 1.08 (0.95-1.22) Lipoprotein subfractions: M-HDL-C but not L/XL-HDL-C was significant in both analyses; Significant for S/M/L-LDL-C, IDL-C, and S/M-VLDL-C but not L-VLDL-C in both analyses; Only S-HDL-TG was significant in both analyses; *1 SD increment* in HDL-C by genetic variants in *CETP* locus: 0.82 (0.75-0.89)

Abbreviations: ICH, Intracerebral haemorrhage; SVS, Small vessel stroke;

Table 2.4 Characteristics of included GWAS of NMR lipid traits and representative GWAS of traditional lipids

Study	Lipid traits	GWAS consortium (year) /first author	Ethnicity and ancestry of study population	Sample size	Adjustments	Variance explained	Findings (URL to summary genetics)
GWAS of NMR lipids							
Chasman, 2009[39]	17 NMR lipid-related traits and 5 traditional lipids (LDL-C, HDL-C, TG, apoA1, $apoB_{100}$)	WGHS	European women	17296 (WGHS)	Age, BMI, menopausal status, smoking, hormone replacement therapy, and 10 genetic principal components	2.1%-17.2%	43 loci (10 novel) -
Petersen, 2011[40]	15 NMR lipoprotein subfractions	KORA	European	1791 (KORA)	Age, age^2, sex	1.5%-4.5%	8 loci -
Kettunen, 2016[34]	123 NMR metabolites including lipid-related traits	Kettunen	European	24925 (14 cohorts)	Age, sex, and the first 10 genetic principal components	0.2%-12.5%	62 loci (8 novel) http://www.computationalmedicine.fi/data#NMR_GWAS

Davis, 2017[35]	72 NMR lipid-related traits	Davis	Finnish men	8372 (METSIM study)	Age, age^2, lipid-lowering therapy, and smoking	0.3%-1.1%	47 loci (5 novel) http://csg.sph.umich.edu/boehnke/public/metsim-2017-lipoproteins/
GWAS of traditional lipids							
Willer, 2013[37]	Traditional lipid panel (LDL-C, HDL-C, TG, and TC)	GLGC (2013)	Multiethnic (European, African, South Asian, East Asian)	188578 (60 studies)	Age, age^2, genetic principal components	In the Framingham offspring cohort: 1.6%-2.6%	157 loci (62 novel) http://csg.sph.umich.edu/abecasis/public/lipids2013/
Liu, 2017[41]	Traditional lipid panel (LDL-C, HDL-C, TG, and TC)	GLGC (2017)	Multiethnic (European, African, South Asian, East Asian)	300,000 (73 studies)	Age, age, genetic principal components	-	444 SNPs (75 novel) in 250 loci
Hoffmann, 2018[36]	Traditional lipid panel (LDL-C, HDL-C, TG, and TC)	GERA	Multiethnic (non-Hispanic white, Latino, African American, South Asian, East Asian)	Discovery: 94674 Replication: 94595 (GLGC) & 460088 (UKB)	Age^3, BMI^3, sex, top 10 genetic principal components for non-Hispanic whites and 6 for other groups	11%-23%	171 loci (60 novel) ftp://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/

Klarin, 2018[42]	Traditional lipid panel (LDL-C, HDL-C, TG, and TC)	MVP	Multiethnic (White [European], Black [African], Hispanic)	Discovery: 297000 Replication: 319000 (GLGC)	Age, age^2, sex, top 10 genetic principal components of ancestry	8.8%-12.3%	386 loci (118 novel) -
Richardson, 2020[43]	Traditional lipid panel (LDL-C, HDL-C, TG, and TC) and apolipoprotein (apoB and apoA-I)	UKBB	European	459000 (UKBB)	Age, sex, and genotyping chip	-	220-534 significant SNPs (56%-93% novel) -
Graham, 2021[44]	Traditional lipid panel (LDL-C, HDL-C, nonHDL-C, TG, and TC) and apolipoprotein (apoB and apoA-I)	GLGC (2021)	Multiethnic (admixed African or African, East Asian, European, Hispanic, South Asian)	1.65 million individuals	Age, sex, genetic principal components	9%-12%	1,765 significant lead SNPs (237 novel)

Abbreviations: METSIM, METabolic Syndrome In Men; WGHS, Women's Genome Health Study; KORA, The Cooperative Health Research in the Region of Augsburg; MVP, Million Veteran Program; UKBB, UK Biobank; SNP, single nucleotide polymorphism;

References

1 Page, M. J. *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ (Clinical research ed.)* **372**, n71, doi:10.1136/bmj.n71 (2021).

2 Stroup, D. F. *et al.* Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *Jama* **283**, 2008-2012, doi:10.1001/jama.283.15.2008 (2000).

3 Blake, G. J., Otvos, J. D., Rifai, N. & Ridker, P. M. Low-density lipoprotein particle concentration and size as determined by nuclear magnetic resonance spectroscopy as predictors of cardiovascular disease in women. *Circulation* **106**, 1930-1937, doi:10.1161/01.cir.0000033222.75187.b9 (2002).

4 Kuller, L. *et al.* Nuclear magnetic resonance spectroscopy of lipoproteins and risk of coronary heart disease in the cardiovascular health study. *Arteriosclerosis, thrombosis, and vascular biology* **22**, 1175-1180, doi:10.1161/01.atv.0000022015.97341.3a (2002).

5 Soedamah-Muthu, S. S., Chang, Y. F., Otvos, J., Evans, R. W. & Orchard, T. J. Lipoprotein subclass measurements by nuclear magnetic resonance spectroscopy improve the prediction of coronary artery disease in Type 1 diabetes. A prospective report from the Pittsburgh Epidemiology of Diabetes Complications Study. *Diabetologia* **46**, 674-682, doi:10.1007/s00125-003-1094-8 (2003).

6 El Harchaoui, K. *et al.* Value of low-density lipoprotein particle number and size as predictors of coronary artery disease in apparently healthy men and women: the EPIC-Norfolk Prospective Population Study. *Journal of the American College of Cardiology* **49**, 547-553, doi:10.1016/j.jacc.2006.09.043 (2007).

7 van der Steeg, W. A. *et al.* High-density lipoprotein cholesterol, high-density lipoprotein particle size, and apolipoprotein A-I: significance for cardiovascular risk: the IDEAL and EPIC-Norfolk studies. *Journal of the American College of Cardiology* **51**, 634-642, doi:10.1016/j.jacc.2007.09.060 (2008).

8 El Harchaoui, K. *et al.* High-density lipoprotein particle size and concentration and coronary risk. *Annals of internal medicine* **150**, 84-93, doi:10.7326/0003-4819-150-2-200901200-00006 (2009).

9 Arsenault, B. J. *et al.* Lipid assessment, metabolic syndrome and coronary heart disease risk. *European journal of clinical investigation* **40**, 1081-1093, doi:10.1111/j.1365-2362.2010.02357.x (2010).

10 van Capelleveen, J. C. *et al.* Apolipoprotein C-III Levels and Incident Coronary Artery Disease Risk: The EPIC-Norfolk Prospective Population

Study. *Arteriosclerosis, thrombosis, and vascular biology* **37**, 1206-1212, doi:10.1161/atvbaha.117.309007 (2017).
11 Holmes, M. V. *et al.* Lipids, Lipoproteins, and Metabolites and Risk of Myocardial Infarction and Stroke. *Journal of the American College of Cardiology* **71**, 620-632, doi:10.1016/j.jacc.2017.12.006 (2018).
12 Salahuddin, T. *et al.* Association of high-density lipoprotein particle concentration with cardiovascular risk following acute coronary syndrome: A case-cohort analysis of the dal-Outcomes trial. *American heart journal* **221**, 60-66, doi:10.1016/j.ahj.2019.12.003 (2020).
13 Cromwell, W. C. *et al.* LDL particle number and risk of future cardiovascular disease in the Framingham Offspring Study-Implications for LDL management. *Journal of Clinical Lipidology* **1(6)**, 583-592 (2007).
14 Mora, S. *et al.* Lipoprotein particle profiles by nuclear magnetic resonance compared with standard lipids and apolipoproteins in predicting incident cardiovascular disease in women. *Circulation* **119**, 931-939, doi:10.1161/circulationaha.108.816181 (2009).
15 Mackey, R. H. *et al.* High-density lipoprotein cholesterol and particle concentrations, carotid atherosclerosis, and coronary events: MESA (multi-ethnic study of atherosclerosis). *Journal of the American College of Cardiology* **60**, 508-516, doi:10.1016/j.jacc.2012.03.060 (2012).
16 Mora, S., Glynn, R. J. & Ridker, P. M. High-density lipoprotein cholesterol, size, particle number, and residual vascular risk after potent statin therapy. *Circulation* **128**, 1189-1197, doi:10.1161/circulationaha.113.002671 (2013).
17 Tsai, M. Y. *et al.* New automated assay of small dense low-density lipoprotein cholesterol identifies risk of coronary heart disease: the Multi-ethnic Study of Atherosclerosis. *Arteriosclerosis, thrombosis, and vascular biology* **34**, 196-201, doi:10.1161/atvbaha.113.302401 (2014).
18 Akinkuolie, A. O., Lawler, P. R., Glynn, R. J., Ridker, P. M. & Mora, S. HDL particle subclasses, rosuvastatin therapy, and risk of vascular events: An analysis from the jupiter trial. *Circulation. Conference: American Heart Association's* **132** (2015).
19 Steffen, B. T. *et al.* Use of lipoprotein particle measures for assessing coronary heart disease risk post-American Heart Association/American College of Cardiology guidelines: the Multi-Ethnic Study of Atherosclerosis. *Arteriosclerosis, thrombosis, and vascular biology* **35**, 448-454, doi:10.1161/atvbaha.114.304349 (2015).
20 Würtz, P. *et al.* Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts. *Circulation* **131**, 774-785, doi:10.1161/circulationaha.114.013116 (2015).
21 Lawler, P. R. *et al.* Atherogenic Lipoprotein Determinants of Cardiovascular Disease and Residual Risk Among Individuals With Low Low-Density Lipoprotein Cholesterol. *Journal of the American Heart Association* **6**, doi:10.1161/jaha.117.005549 (2017).

22 Silbernagel, G. *et al.* High-Density Lipoprotein Subclasses, Coronary Artery Disease, and Cardiovascular Mortality. *Clinical chemistry* **63**, 1886-1896, doi:10.1373/clinchem.2017.275636 (2017).
23 Aday, A. W. *et al.* Lipoprotein Particle Profiles, Standard Lipids, and Peripheral Artery Disease Incidence. *Circulation* **138**, 2330-2341, doi:10.1161/circulationaha.118.035432 (2018).
24 Akbaraly, T. *et al.* Association of circulating metabolites with healthy diet and risk of cardiovascular disease: analysis of two cohort studies. *Scientific reports* **8(1)**, 8620 (2018).
25 Joshi, R. *et al.* Triglyceride-containing lipoprotein sub-fractions and risk of coronary heart disease and stroke: A prospective analysis in 11,560 adults. *European journal of preventive cardiology* **27**, 1617-1626, doi:10.1177/2047487319899621 (2020).
26 Singh, K. *et al.* Associations Between High-Density Lipoprotein Particles and Ischemic Events by Vascular Domain, Sex, and Ethnicity: A Pooled Cohort Analysis. *Circulation* **142**, 657-669, doi:10.1161/circulationaha.120.045713 (2020).
27 Vojinovic, D. *et al.* Association of circulating metabolites in plasma or serum and risk of stroke: Meta-analysis from seven prospective cohorts. *Neurology* **96**, e1110-1123, doi:10.1212/wnl.0000000000011236 (2020).
28 Johansen, M. *et al.* Per-Particle Triglyceride-Rich Lipoproteins Imply Higher Myocardial Infarction Risk Than Low-Density Lipoproteins: Copenhagen General Population Study. *Arteriosclerosis, thrombosis, and vascular biology* **41**, 2063-2075, doi:10.1161/atvbaha.120.315639 (2021).
29 Ference, B. A. *et al.* Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. *European heart journal* **38**, 2459-2472, doi:10.1093/eurheartj/ehx144 (2017).
30 Zuber, V., Colijn, J. M., Klaver, C. & Burgess, S. Selecting likely causal risk factors from high-throughput experiments using multivariable Mendelian randomization. *Nature communications* **11**, 29, doi:10.1038/s41467-019-13870-3 (2020).
31 Levin, M. G. *et al.* Prioritizing the Role of Major Lipoproteins and Subfractions as Risk Factors for Peripheral Artery Disease. *Circulation* **144**, 353-364, doi:10.1161/circulationaha.121.053797 (2021).
32 Georgakis, M. K. *et al.* Genetic determinants of blood lipids and cerebral small vessel disease: role of high-density lipoprotein cholesterol. *Brain* **143**, 597-610, doi:10.1093/brain/awz413 (2020).
33 Zhao, Q. *et al.* A Mendelian randomization study of the role of lipoprotein subfractions in coronary artery disease. *eLife* **10**, doi:10.7554/eLife.58361 (2021).
34 Kettunen, J. *et al.* Genome-wide study for circulating metabolites identifies

62 loci and reveals novel systemic effects of LPA. *Nature communications* **7**, 11122, doi:10.1038/ncomms11122 (2016).

35 Davis, J. P. *et al.* Common, low-frequency, and rare genetic variants associated with lipoprotein subclasses and triglyceride measures in Finnish men from the METSIM study. *PLoS genetics* **13**, e1007079, doi:10.1371/journal.pgen.1007079 (2017).

36 Hoffmann, T. J. *et al.* A large electronic-health-record-based genome-wide study of serum lipids. *Nature genetics* **50**, 401-413, doi:10.1038/s41588-018-0064-5 (2018).

37 Willer, C. J. *et al.* Discovery and refinement of loci associated with lipid levels. *Nature genetics* **45**, 1274-1283, doi:10.1038/ng.2797 (2013).

38 Holmes, M. V. *et al.* Mendelian randomization of blood lipids for coronary heart disease. *European heart journal* **36**, 539-550, doi:10.1093/eurheartj/eht571 (2015).

39 Chasman, D. I. *et al.* Forty-three loci associated with plasma lipoprotein size, concentration, and cholesterol content in genome-wide analysis. *PLoS genetics* **5**, e1000730, doi:10.1371/journal.pgen.1000730 (2009).

40 Petersen, A. K. *et al.* Genetic associations with lipoprotein subfractions provide information on their biological nature. *Human molecular genetics* **21**, 1433-1443, doi:10.1093/hmg/ddr580 (2012).

41 Liu, D. J. *et al.* Exome-wide association study of plasma lipids in >300,000 individuals. *Nature genetics* **49**, 1758-1766, doi:10.1038/ng.3977 (2017).

42 Klarin, D. *et al.* Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program. *Nature genetics* **50**, 1514-1523, doi:10.1038/s41588-018-0222-9 (2018).

43 Richardson, T. G. *et al.* Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: A multivariable Mendelian randomisation analysis. *PLoS medicine* **17**, e1003062, doi:10.1371/journal.pmed.1003062 (2020).

44 Graham, S. E. *et al.* The power of genetic diversity in genome-wide association studies of lipids. *Nature* **600**, 675-679, doi:10.1038/s41586-021-04064-3 (2021).

CHAPTER 3: Overall study design and methods

Chapter summary

The current chapter outlines the study design, recruitment, baseline assessment, and follow-up for mortality outcomes in MCPS. A brief introduction to the nuclear magnetic resonance metabolomics platform used to measure lipidome is presented. Finally, the general statistical and epidemiological methods, including selection criteria of participants and covariables, and baseline characteristics are described. Detailed statistical methods for specific analyses such as GWAS, observational, and MR are detailed in the subsequent relevant chapters.

3.1 Study design, procedures, and follow-up

3.1.1 Overall study design and recruitment of participants

The Mexico City Prospective Study (MCPS) is a population-based prospective cohort study designed to investigate the associations of traditional and emerging risk factors with the major causes of death in Mexico.[1,2] Following the study protocol established by the collaboration between researchers at the National Autonomous University of Mexico (Mexico City) and the Clinical Trial Service Unit & Epidemiological Studies Unit (University of Oxford), two neighboring geographical districts of Mexico City (Coyoacán and Iztapalapa) were selected due to the contrasting socio-economic characteristics and relatively stable populations consisting of local residents and migrants (**Figure 3.1**). Between 1995 and 1997, all households in the selected districts were recorded by census-style door-to-door interviews and basic information of each household member (age, sex, and address) was recorded. Then, between 1998 and 2004, 112,333 households were systematically visited and, from these, 106,059 (94%) yielded 159,755 eligible participants aged 35 years of older who provided consent to take part. Women accounted for a larger proportion (66%) of participants than men because they were more likely to be at home during working hours when investigators mostly visited households.

3.1.2 Baseline assessment procedures

After providing informed consent, detailed information of each participant was collected by trained nurses using a standardized questionnaire administered at the participant's home. The questionnaire collected sociodemographic and lifestyle factors, use of medications and medical history (**Table 3.1**). Physical measurements (seated blood pressure and anthropometry) were measured and nearly all participants provided a 10 ml non-fasting blood sample, with time since last meal recorded.

The collected information was entered directly into a handheld data recorder that was programmed to query highly implausible values and the data were automatically

uploaded to the study's electronic database. The database was then reviewed regularly to identify further potential data errors. Two types of recorders were used in the baseline interviews: *Microwand 32ES* barcode optical reading device (Hand Held Products) between 1998 and 2001, and *SPT 500* personal digital assistants with Palm operating system (Symbol) between 2001 and 2004.

3.1.3 Physical measurements

Anthropometric parameters such as height, weight, and waist and hip circumference were measured by trained nurses using standard equipment that was calibrated every two weeks, with participants wearing light clothing. Standing height without shoes was measured (to the nearest 0.1 cm) using wooden triangle and a non-stretchable measuring tape on a flat surface, with the head kept in Frankfort-Plane position. Weight was measured using portable analogue scales (to the nearest kg) in 1998-2003 and portable electronic scales (to the nearest 0.1 kg) in 2003-2004. Body mass index was calculated as the weight in kilograms divided by the height in square meters. Waist circumference was measured (to the nearest 0.1 cm) using a non-stretchable measuring tape at the midpoint between the iliac crest and the costal margin. Hip circumference was measured using a non-stretchable measuring tape around the femoral epicondyles. Sitting blood pressure was measured (to the nearest 1 mmHg) three times at 3-minute intervals using a standard mercury sphygmomanometer with a suitably sized cuff.

3.1.4 Blood sample

Almost all participants (155,487, 97.3% of those with a completed questionnaire) provided a non-fasting blood sample, with time since last meal recorded. The 10 mL blood sample was collected into a single EDTA tube labelled with a bar-coded identifier that was linked to each participant's record. After collection, the blood sample was placed in an insulated box containing chilled packs (4-10°C) before transportation to a central laboratory in Mexico City at the end of the end of working day for overnight storage (4°C). The next working day, the blood sample was extracted and centrifuged. The plasma was

separated and stored in two identical 1.8 mL bar-coded cryovials and the buffy coat was separated and stored in one cryovial of the same type. The set of three cryovials were placed separately into three storage boxes and stored in -80°C freezers before shipment on dry ice to Oxford (UK) for long-term storage in tanks containing liquid nitrogen vapor (-150°C). The mean delay (at 4-10°C) between sample collection and separation was about 24 hours.[2]

3.1.5 Follow-up and adjudication

In Mexico, almost all adult deaths are medically certified and coded according to the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10), with only a small proportion of deaths attributed to unknown causes.[2,3] This information, together with personal identifying information, is systematically recorded in a Ministry of Health database. Participants are followed for cause-specific mortality through linkage to the Mexican Death Registry (*Sistema Epidemiológico y Estadístico de las Defunciones, SEED*) using a Soundex probabilistic matching algorithm based on personal identifying information (name, age, sex, and place of death). Field validation of 7,000 matched deaths confirmed the reliability of the matching algorithm in more than 95% of the cases. By December 2020, 30,635 medically certified deaths had been reliably identified and linked to participants in MCPS. Linked mortality information includes date and place of death, ICD-10 coded underlying and contributory causes of death, time duration between onset of disease and death, and medical certification at the time of death. Study clinicians then reviewed the death records and, when necessary, recoded the underlying cause of death. For example, diabetes was accepted as the underlying cause only for deaths due to acute diabetic crises, and other such deaths were recoded to alternative causes such as vascular, renal, or infection, based on the records in parts 1 and 2 of the certificate. Deaths attributed primarily to a risk factor were also recoded to an epidemiologically appropriate alternative underlying disease. For example, deaths attributed to hypertension or hypercholesterolemia were recoded to a vascular disease.

3.1.6 Funding and ethics approval

This work was supported by the Mexican Health Ministry; the National Council of Science and Technology for Mexico; Wellcome [058299/Z/99]; Cancer Research UK; the British Heart Foundation [RE/13/1/30181]; and the UK Medical Research Council [MC_UU_00017/2, MR/Z504543/1]. Formal approval to share MCPS data with commercial institutions was sought and obtained from the Medical Ethics Committee of the National Autonomous University of Mexico (FMED/CEI/MHU/001/2020). Ethics approval was obtained from the National Council of Science and Technology in Mexico, the Mexican Ministry of Health, and the University of Oxford (135/2014). All participants recruited into this study provided written informed consent to participate.

3.2 Measurement of NMR lipids, lipoproteins, and apolipoproteins

3.2.1 Proton NMR spectroscopic method

In this study, Nightingale Health Ltd high-throughput metabolomics platform are used to quantify lipid-related measurements in Finland and UK.[4] The NMR method used the characteristic signals emitted by each lipoprotein subclass to directly measure the concentration of that subclass particle. The signals were broadcast by methyl groups on the lipids contained in the particle that were determined by the particle diameter (proportional to the lipid mass concentration) instead of the variations in lipid composition (e.g. the relative proportion of TG and CE in the lipoprotein, because each of them contribute 2 methyl groups).[5,6] The methodology used to generate and process NMR metabolomics data generally consist of three steps: (1) acquisition of proton NMR spectra of the plasma specimens (with a 500/600 MHz spectrometer in this study); (2) computational decomposition of lipid methyl signal envelope in these spectra to yield the contributing signal amplitudes of the 14 lipoprotein subclasses; (3) conversion of these signal amplitudes to lipoprotein subclass concentrations using experimentally determined factors conversion that relate the signal amplitudes of purified subclass standards to their chemically measured cholesterol and TG concentrations.

3.2.2 Sample management and quality control

In 2018-2020, plasma sample of 153,888 participants were sub-aliquoted into one vial of 110 μL and two vials of 200 μL and were refrozen for storage or transportation. At the time of analysis, around 120,000 participants had measured data of NMR metabolites. Most of samples were analysed at Nightingale Health Ltd (Kuopio, Finland) using a 500 Mhz spectrometer. Following the same protocol, several batches were analysed at the Clinical Trial Service Unit's (CTSU) Wolfson laboratory (Oxford, United Kingdom) using a 600 Mhz spectrometer. Samples went through two thawing cycles including one for reformatting and one for analysis.

For each batch, samples were analysed in racks with 96 3-mm NMR tubes (140 μL 1:1 sample to buffer ratio), and the first and the last tube were set as high- and low-level quality control (QC) samples, respectively. The QC samples were aliquots from two pools of human EDTA plasma and were processed following the same steps. Although QC limits are pre-determined, means and standard deviations of eight NMR metabolic measures (HDL-C, LDL-C, TC, TG, ApoA1, ApoB, and Creatinine) were tracked using Levey-Jennings plots as an additional QC to verify the acceptability of NMR measurements. In addition, comparison between measurements by Nightingale Health Ltd and those by CTSU Wolfson Laboratory (based on multiple set of samples analysed in both sites) were performed, with good inter-laboratory agreement observed.

3.2.3 NMR lipid-related traits

In total, the Nightingale Health Ltd metabolomics platform quantified 228 lipid and other metabolic measurements (including 131 direct measures and 97 indirect measures derived from post-estimation algorithms), either as absolute concentrations of each metabolic measure or as ratios. In this study investigating the association between blood lipid levels and the risk of vascular mortality, an array of 104 NMR measurements related to lipid traits including both direct and derived measures were used as main exposures. The array of NMR lipid measures includes the absolute concentrations (particle number)

of 14 lipoprotein subclasses (6 VLDL, IDL, 3 LDL, and 4 HDL), the lipids contained in each lipoprotein subclass (cholesterol (total, esterified, and free), triglycerides, phospholipids, and total lipids), mean lipoprotein particle size (VLDL, LDL, and HDL), concentrations of ApoB-100 and ApoA1, and the ApoB/ApoA1 ratio (**Table 3.2**). The ratios between lipoprotein lipids and measures of fatty acids were not included. This NMR lipid array included most of the direct measures and excluded most of the derived measures such as ratios. The results in the tables and figures were sorted by the lipid metabolic groups described above to facilitate comparison between groups.

All the subfraction lipid traits are named with three components separated by hyphens: the first component indicates the particle size (For VLDL: XS, S, M, L, XL, XXL; For LDL: S, M, L; For HDL: S, M, L, XL); the second component indicates the lipoprotein fraction according to their density (VLDL, IDL, LDL, HDL); the third component indicates the lipid traits (C for total cholesterol, CE for cholesterol esters, FC for free cholesterol, TG for triglycerides, PL for phospholipids, L for total lipids, P for particle concentration, and D for particle size). For example, L-LDL-L refers to the concentration of total lipids contained in the large LDL particles. Besides lipoprotein subfractions, concentration of the composite lipid measures (TG, TC, LDL-C, HDL-C) and apolipoprotein (ApoA1, ApoB) were also measured. Notably, most of the direct lipid measures in the array are not independent and several lipid measures are considered multiple times in the NMR biomarkers platform. For example, the total lipid in any lipoprotein particle subclass is the sum of the free and esterified cholesterol (FC and CE), the triglycerides (TG), and the phospholipids (PL) contained in that subclass (e.g. L-LDL-L=L-LDL-FC + L-LDL-CE + L-LDL-TG + L-LDL-PL). The total cholesterol in large LDL particle (L-LDL-C) was also considered within the total cholesterol in LDL particle (LDL-C), within the total cholesterol (TC).

3.2.4 Statistical quality control

The NMR metabolomics platform conducted automated statistical quality control steps

prior to data release. Further statistical QC steps applied for NMR lipid measures including handling of zero values and missing values, and data transformations were also performed prior to analysis. Zero values indicate that the actual measurements may be lower than the detection threshold of the platform and were replaced with the lowest non-zero value of the relevant NMR measure where applicable. Missing values and the underlying cause of missingness will be reported when the platform was not able to reliably quantify a relevant NMR measure. The potential reasons include the value was rejected by automatic quality control (NA), the value could not be quantified due to sample irregularity (TAG), and undefined reasons (NDEF). Overall, 5% of participants with at least one missing value of NMR measure were excluded from analyses. The distribution of most NMR lipid measures exhibit some degree of right-skew (**Figure 3.2**). To ensure the NMR lipid measures were approximately normally distributed, all lipid measures were log-transformed and then standardized by subtracting the mean and then dividing the standard deviation after zero values and missing values in the dataset has been handled.

3.3 General statistical methods

3.3.1 Selection of variables and variable derivation

3.3.1.1 Categorical variables at baseline

Selected demographic and lifestyle characteristics at baseline assessment were categorized into groups: age category (10-year age groups); district of residence (Coyoacán and Iztapalapa); attained educational level (University or college, high school, elementary school, others); smoking status (never, former, less than daily, less than 10 cigarettes per day, 10 or more cigarettes per day); and alcohol consumption (never, former, up to 3 times per month, up to 3 times per week, more than 3 times per week).

3.3.1.2 Additional variables at baseline

At baseline, participants reported the history of doctor-diagnosed chronic diseases

including diabetes mellitus, coronary heart disease, stroke, chronic kidney disease, emphysema, liver cirrhosis, and cancer of any type. Prior diabetes mellitus was further defined using the criteria including a self-reported history of medical diagnosis, the use of glucose-lowering medications (insulin, sulphonylureas, biguanides, or other), or a glycosylated haemoglobin of 6.5% or higher at recruitment. Participants also reported regular medication use in the past 12 months before recruitment and the information was categorised by study clinicians into pharmacological groups based on the therapeutic application of individual drugs (e.g. antihypertensives, glucose-lowering, lipid-lowering medications, etc). Within each pharmacological groups, the information was categorised into subgroups based on the pharmacological action (e.g. beta-blockers, biguanides, statins, etc). Lipid-lowering medications included statins, fibrates, resins, and niacin.

Blood pressure was categorized into meaningful clinical categories: SBP (<125, ≥125 to <135, ≥135 to <145, ≥145 to <155, ≥155 to <165, and ≥165 mm Hg); DBP (<65, ≥65 to <75, ≥75 to <85, ≥85 to <95, ≥95 to <105, and ≥105 mm Hg). Glycosylated haemoglobin (HbA1c) was categorized into <7, 7-9, and >9 percent units (%). Fasting duration was categorized into <8, 8-12, and >12 hours. Location where NMR measures were analyzed was categorized as Oxford or Finland.

3.3.2 Selection of participants

Of the 159,755 participants recruited into the study, participants who were aged ≥85 (n=2,466), or were using lipid-lowering medications (n=890), or had missing or extreme values of selected covariates at baseline (n=8,225). Threshold for extreme values were defined based on a priori knowledge. Those without NMR lipid measures at the time of analysis were further excluded (n=25,631). Considering the potential effect of reverse causality in observational analyses, participants with self-reported chronic diseases including cardiovascular disease, cancer, or chronic renal, hepatic, or pulmonary disease at baseline were also excluded (n=8135) as well as those with uncertain mortality matches. Flowchart of participant selection are shown in **Figure 3.3**.

3.4 Baseline characteristics

Table 3.3 describes the baseline socio-demographic and lifestyle characteristics of participants included in the main analyses. The study population consisted of a total of 107,504 individuals aged 35 to 84 years without history of chronic diseases (CVD, cancer, or chronic renal, hepatic, or pulmonary disease) who were not using lipid-lowering medications and had complete measurement of NMR lipid-related traits and other relevant covariates at baseline. The mean age of included participants was 51 (SD 12) years, 33% (n=35,518) were male, and 44% (n=47,683) were residents of Coyoacán district. According to the self-reported baseline survey, 16% (n=16,897) of participants had a University or College degree, 7% (n=7,017) of participants smoked 10 or more cigarettes per day and 5% (n=5,792) of participants consumed alcohol more than 3 times per week.

Table 3.4 describes the baseline biological and medical characteristics of participants included in the main analyses. Mean BMI was 29 (SD 5) kg/m^2, mean SBP was 127 (SD 17) mmHg and mean DBP was 83 (SD 10) mmHg, and HbA1c was 6.1 (SD 1.7)%. Mean fasting duration was 3.9 (SD 4.3) hours and only 12% of participants reported a fasting duration longer than 8 hours. 13% (n=13,662) of participants meet the definition of prior diabetes (self-reported diagnosis, use of glucose-lowering medications, or HbA1c ≥ 6.5% at baseline).

Men and women had similar baseline characteristics except attained educational level, smoking status, alcohol consumption, and physical measures (**Table 3.3** and **Table 3.4**). Compared with women, men were more likely to have a College or University degree, smoke and consume alcohol regularly, be taller and heavier, and had a lower mean BMI.

Reference

1. Alegre-Díaz J, Herrington W, López-Cervantes M, et al. Diabetes and Cause-Specific Mortality in Mexico City. The New England journal of medicine 2016;375(20):1961-1971. (In eng). DOI: 10.1056/NEJMoa1605368.
2. Tapia-Conyer R, Kuri-Morales P, Alegre-Díaz J, et al. Cohort profile: the Mexico City Prospective Study. Int J Epidemiol 2006;35(2):243-9. (In eng). DOI: 10.1093/ije/dyl042.
3. Hernández B, Ramírez-Villalobos D, Romero M, Gómez S, Atkinson C, Lozano R. Assessing quality of medical death certification: Concordance between gold standard diagnosis and underlying cause of death in selected Mexican hospitals. Population health metrics 2011;9:38. (In eng). DOI: 10.1186/1478-7954-9-38.
4. Soininen P, Kangas AJ, Würtz P, Suna T, Ala-Korpela M. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. Circulation Cardiovascular genetics 2015;8(1):192-206. (In eng). DOI: 10.1161/circgenetics.114.000216.
5. Otvos JD, Jeyarajah EJ, Bennett DW. Quantification of plasma lipoproteins by proton nuclear magnetic resonance spectroscopy. Clinical chemistry 1991;37(3):377-86. (In eng).
6. Otvos JD, Jeyarajah EJ, Bennett DW, Krauss RM. Development of a proton nuclear magnetic resonance spectroscopic method for determining plasma lipoprotein concentrations and subspecies distributions from a single, rapid measurement. Clinical chemistry 1992;38(9):1632-8. (In eng).

Table 3.1 Information collected at baseline by questionnaire survey

Socio-demographic
Date of birth
Sex
District of residence
Attained education level
Income (household, personal)
Occupation
Lifestyle
Smoking (passive, active status and amount)
Alcohol use (status, type, frequency)
Physical activity
Dietary pattern (fruits/vegetables intake, fried foods)
Sleep duration
Time since last meal
Health
Healthcare insurance (provider)
Doctor-diagnosed disease (type, time)
Medication regularly used in the last year
Reproductive history

Table 3.2 Description of NMR lipid measures used in the main analyses

ID	Trait	Name	Label	Quantification Method
1	Particle concentration of	Chylomicrons and extremely large VLDL particles (mmol/l)	XXL-VLDL-P	Derived
2		Very large VLDL particles (mmol/l)	XL-VLDL-P	Derived
3		Large VLDL particles (mmol/l)	L-VLDL-P	Derived
4		Medium VLDL particles (mmol/l)	M-VLDL-P	Derived
5		Small VLDL particles (mmol/l)	S-VLDL-P	Derived
6		Very small VLDL particles (mmol/l)	XS-VLDL-P	Derived
7		IDL particles (mmol/l)	IDL-P	Derived
8		Large LDL particles (mmol/l)	L-LDL-P	Derived
9		Medium LDL particles (mmol/l)	M-LDL-P	Derived
10		Small LDL particles (mmol/l)	S-LDL-P	Derived
11		Very large HDL particles (mmol/l)	XL-HDL-P	Derived
12		Large HDL particles (mmol/l)	L-HDL-P	Derived
13		Medium HDL particles (mmol/l)	M-HDL-P	Derived
14		Small HDL particles (mmol/l)	S-HDL-P	Derived
15	Total cholesterol in	Chylomicrons and extremely large VLDL particles (mmol/l)	XXL-VLDL-C	Derived
16		Very large VLDL particles (mmol/l)	XL-VLDL-C	Derived
17		Large VLDL particles (mmol/l)	L-VLDL-C	Derived
18		Medium VLDL particles (mmol/l)	M-VLDL-C	Derived
19		Small VLDL particles (mmol/l)	S-VLDL-C	Derived
20		Very small VLDL particles (mmol/l)	XS-VLDL-C	Derived
21		IDL particles (mmol/l)	IDL-C	Derived
22		Large LDL particles (mmol/l)	L-LDL-C	Derived
23		Medium LDL particles (mmol/l)	M-LDL-C	Derived
24		Small LDL particles (mmol/l)	S-LDL-C	Derived
25		Very large HDL particles (mmol/l)	XL-HDL-C	Derived
26		Large HDL particles (mmol/l)	L-HDL-C	Derived
27		Medium HDL particles (mmol/l)	M-HDL-C	Derived
28		Small HDL particles (mmol/l)	S-HDL-C	Derived
29	Free cholesterol in	Chylomicrons and extremely large VLDL particles (mmol/l)	XXL-VLDL-FC	Direct
30		Very large VLDL particles (mmol/l)	XL-VLDL-FC	Direct
31		Large VLDL particles (mmol/l)	L-VLDL-FC	Direct
32		Medium VLDL particles (mmol/l)	M-VLDL-FC	Direct
33		Small VLDL particles (mmol/l)	S-VLDL-FC	Direct
34		Very small VLDL particles (mmol/l)	XS-VLDL-FC	Direct
35		IDL particles (mmol/l)	IDL-FC	Direct
36		Large LDL particles (mmol/l)	L-LDL-FC	Direct
37		Medium LDL particles (mmol/l)	M-LDL-FC	Direct
38		Small LDL particles (mmol/l)	S-LDL-FC	Direct
39		Very large HDL particles (mmol/l)	XL-HDL-FC	Direct
40		Large HDL particles (mmol/l)	L-HDL-FC	Direct
41		Medium HDL particles (mmol/l)	M-HDL-FC	Direct
42		Small HDL particles (mmol/l)	S-HDL-FC	Direct
43	Esterified cholesterol in	Chylomicrons and extremely large VLDL particles (mmol/l)	XXL-VLDL-CE	Direct
44		Very large VLDL particles (mmol/l)	XL-VLDL-CE	Direct
45		Large VLDL particles (mmol/l)	L-VLDL-CE	Direct
46		Medium VLDL particles (mmol/l)	M-VLDL-CE	Direct
47		Small VLDL particles (mmol/l)	S-VLDL-CE	Direct
48		Very small VLDL particles (mmol/l)	XS-VLDL-CE	Direct

Table 3.2 Description of NMR lipid measures used in the main analyses

ID	Trait	Name	Label	Quantification Method
49		IDL particles (mmol/l)	IDL-CE	Direct
50		Large LDL particles (mmol/l)	L-LDL-CE	Direct
51		Medium LDL particles (mmol/l)	M-LDL-CE	Direct
52		Small LDL particles (mmol/l)	S-LDL-CE	Direct
53		Very large HDL particles (mmol/l)	XL-HDL-CE	Direct
54		Large HDL particles (mmol/l)	L-HDL-CE	Direct
55		Medium HDL particles (mmol/l)	M-HDL-CE	Direct
56		Small HDL particles (mmol/l)	S-HDL-CE	Direct
57	Triglycerides in	Chylomicrons and extremely large VLDL particles (mmol/l)	XXL-VLDL-TG	Direct
58		Very large VLDL particles (mmol/l)	XL-VLDL-TG	Direct
59		Large VLDL particles (mmol/l)	L-VLDL-TG	Direct
60		Medium VLDL particles (mmol/l)	M-VLDL-TG	Direct
61		Small VLDL particles (mmol/l)	S-VLDL-TG	Direct
62		Very small VLDL particles (mmol/l)	XS-VLDL-TG	Direct
63		IDL particles (mmol/l)	IDL-TG	Direct
64		Large LDL particles (mmol/l)	L-LDL-TG	Direct
65		Medium LDL particles (mmol/l)	M-LDL-TG	Direct
66		Small LDL particles (mmol/l)	S-LDL-TG	Direct
67		Very large HDL particles (mmol/l)	XL-HDL-TG	Direct
68		Large HDL particles (mmol/l)	L-HDL-TG	Direct
69		Medium HDL particles (mmol/l)	M-HDL-TG	Direct
70		Small HDL particles (mmol/l)	S-HDL-TG	Direct
71	Phospholipids in	Chylomicrons and extremely large VLDL particles (mmol/l)	XXL-VLDL-PL	Direct
72		Very large VLDL particles (mmol/l)	XL-VLDL-PL	Direct
73		Large VLDL particles (mmol/l)	L-VLDL-PL	Direct
74		Medium VLDL particles (mmol/l)	M-VLDL-PL	Direct
75		Small VLDL particles (mmol/l)	S-VLDL-PL	Direct
76		Very small VLDL particles (mmol/l)	XS-VLDL-PL	Direct
77		IDL particles (mmol/l)	IDL-PL	Direct
78		Large LDL particles (mmol/l)	L-LDL-PL	Direct
79		Medium LDL particles (mmol/l)	M-LDL-PL	Direct
80		Small LDL particles (mmol/l)	S-LDL-PL	Direct
81		Very large HDL particles (mmol/l)	XL-HDL-PL	Direct
82		Large HDL particles (mmol/l)	L-HDL-PL	Direct
83		Medium HDL particles (mmol/l)	M-HDL-PL	Direct
84		Small HDL particles (mmol/l)	S-HDL-PL	Direct
85	Total lipids in	Chylomicrons and extremely large VLDL particles (mmol/l)	XXL-VLDL-L	Derived
86		Very large VLDL particles (mmol/l)	XL-VLDL-L	Derived
87		Large VLDL particles (mmol/l)	L-VLDL-L	Derived
88		Medium VLDL particles (mmol/l)	M-VLDL-L	Derived
89		Small VLDL particles (mmol/l)	S-VLDL-L	Derived
90		Very small VLDL particles (mmol/l)	XS-VLDL-L	Derived
91		IDL particles (mmol/l)	IDL-L	Derived
92		Large LDL particles (mmol/l)	L-LDL-L	Derived
93		Medium LDL particles (mmol/l)	M-LDL-L	Derived
94		Small LDL particles (mmol/l)	S-LDL-L	Derived
95		Very large HDL particles (mmol/l)	XL-HDL-L	Derived
96		Large HDL particles (mmol/l)	L-HDL-L	Derived
97		Medium HDL particles (mmol/l)	M-HDL-L	Derived

98		Small HDL particles (mmol/l)	S-HDL-L	Derived
99	Mean size of	VLDL particles (nm)	VLDL-D	Direct
100		LDL particles (nm)	LDL-D	Direct
101		HDL particles (nm)	HDL-D	Direct
102		Apolipoprotein A-I (g/l)	ApoA1	Direct
103		Apolipoprotein B (g/l)	ApoB	Direct
104		Ratio of apolipoprotein B to apolipoprotein A-I (ratio)	ApoB-ApoA1	Derived
105		Total LDL cholesterol (mmol/l)	LDL-C	Derived
106		Total HDL cholesterol (mmol/l)	HDL-C	Derived
107		Total Triglycerides (mmol/l)	TG	Derived

Based on the Nightingale Health Ltd metabolomics platform, NMR lipid measures were estimated directly from characteristic spectra obtained using NMR spectroscopy, or from post-estimation algorithms or calculations (e.g. ratios or sums for total concentration).

Table 3.3 Selected socio-demographic and lifestyle characteristics of participants by sex

Characteristic	Men (n=35518)	Women (n=71986)	Overall (n=107504)
Age, mean (SD), year	52 (12)	51 (12)	51 (12)
Age group, year			
35-44	12199 (34%)	26712 (37%)	38911 (36%)
45-54	9776 (28%)	20482 (28%)	30258 (28%)
55-64	6844 (19%)	12999 (18%)	19843 (18%)
65-74	4562 (13%)	8202 (11%)	12764 (12%)
74-84	2137 (6%)	3591 (5%)	5728 (5%)
District of residence			
Coyoacán	16885 (48%)	30798 (43%)	47683 (44%)
Iztapalapa	18633 (52%)	41188 (57%)	59821 (56%)
Educational level			
University or college	8490 (24%)	8407 (12%)	16897 (16%)
High school	9209 (26%)	17390 (24%)	26599 (25%)
Elementary school	14703 (41%)	36020 (50%)	50723 (47%)
Others	3116 (9%)	10169 (14%)	13285 (12%)
Smoking status			
Never	7179 (20%)	45045 (63%)	52224 (49%)
Former	10431 (29%)	10297 (14%)	20728 (19%)
Less than daily	5798 (16%)	6540 (9%)	12338 (11%)
Less than 10 cigarettes per day	7459 (21%)	7738 (11%)	15197 (14%)
10+ cigarettes per day	4651 (13%)	2366 (3%)	7017 (7%)
Alcohol consumption			
Never	2079 (6%)	18520 (26%)	20599 (19%)
Former	3140 (9%)	3253 (5%)	6393 (6%)
Up to 3 times per month	13420 (38%)	11583 (16%)	25003 (23%)
Up to 3 times per month	12680 (36%)	37037 (51%)	49717 (46%)
>3 times per week	4199 (12%)	1593 (2%)	5792 (5%)

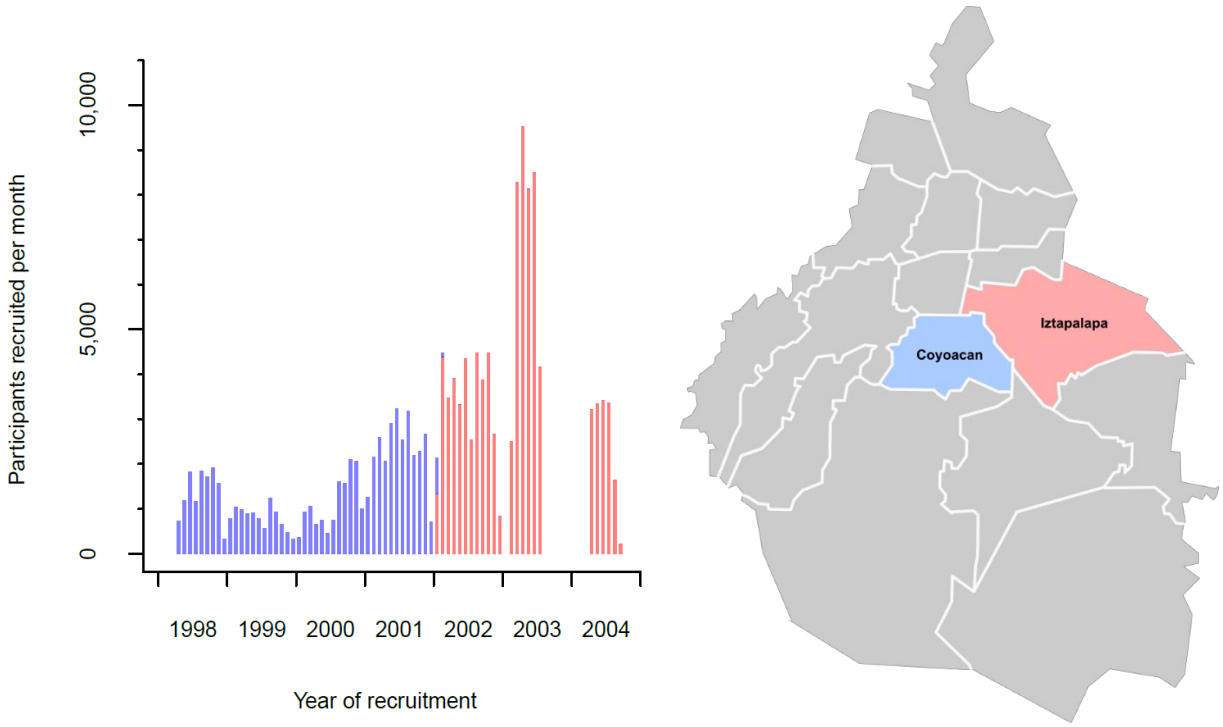
Abbreviation: BMI, body mass index (calculated as weight in kilograms by height in meters squared). SD, standard deviation.
Table excludes individuals with prior chronic disease (other than diabetes) at baseline (e.g. cardiovascular disease, cancer, or chronic renal, hepatic, or pulmonary disease), missing or extreme values of relevant covariates, uncertain mortality marches, or missing values of NMR lipid measures.

Table 3.4 Selected biological and medical characteristics of participants by sex

Characteristic	Men (n=35518)	Women (n=71986)	Overall (n=107504)
Blood pressure, mean (SD), mmHg			
Systolic	128 (16)	127 (17)	127 (17)
Diastolic	84 (10)	83 (10)	83 (10)
HbA_{1c}, %			
<7	30949 (87%)	62592 (87%)	93541 (87%)
7-9	1746 (5%)	3629 (5%)	5375 (5%)
≥9	2823 (8%)	5765 (8%)	8588 (8%)
Mean (SD)	6.1 (1.7)	6.1 (1.7)	6.1 (1.7)
Fasting duration, hours			
<8	30634 (86%)	63543 (88%)	94177 (88%)
8-12	1095 (3%)	1401 (2%)	2496 (2%)
≥12	3789 (11%)	7042 (10%)	10831 (10%)
Mean (SD)	4.1 (4.4)	3.8 (4.2)	3.9 (4.3)
Diabetes*	4495 (13%)	9167 (13%)	13662 (13%)
Physical measures			
Weight, kg	76 (13)	68 (13)	70 (13)
Height, cm	165 (7)	151 (6)	156 (9)
BMI, kg/m^2	28 (4)	30 (5)	29 (5)

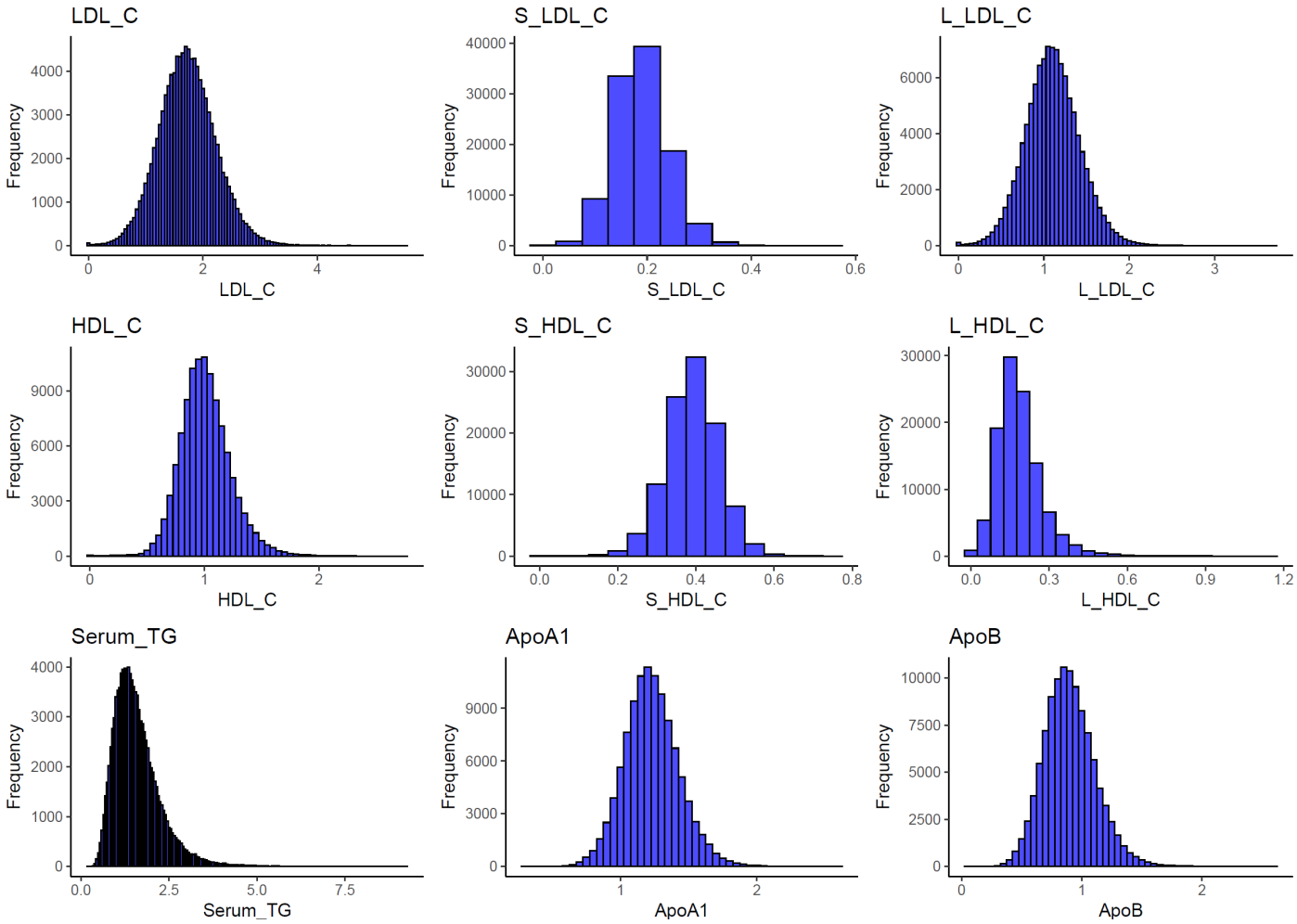
Abbreviation: BMI, body mass index (calculated as weight in kilograms by height in meters squared). SD, standard deviation.
Table excludes individuals with prior chronic disease (other than diabetes) at baseline (e.g. cardiovascular disease, cancer, or chronic renal, hepatic, or pulmonary disease), missing or extreme values of relevant covariates, uncertain mortality matches, or missing values of NMR lipid measures.
*Prior diabetes was defined using the criteria including a self-reported history of medical diagnosis, the use of glucose-lowering medications (insulin, sulphonylureas, biguanides, or other), or a glycosylated haemoglobin of 6.5% or higher at recruitment.

Figure 3.1 Mexico City Prospective Study recruitment sites and dates, 1998-2004



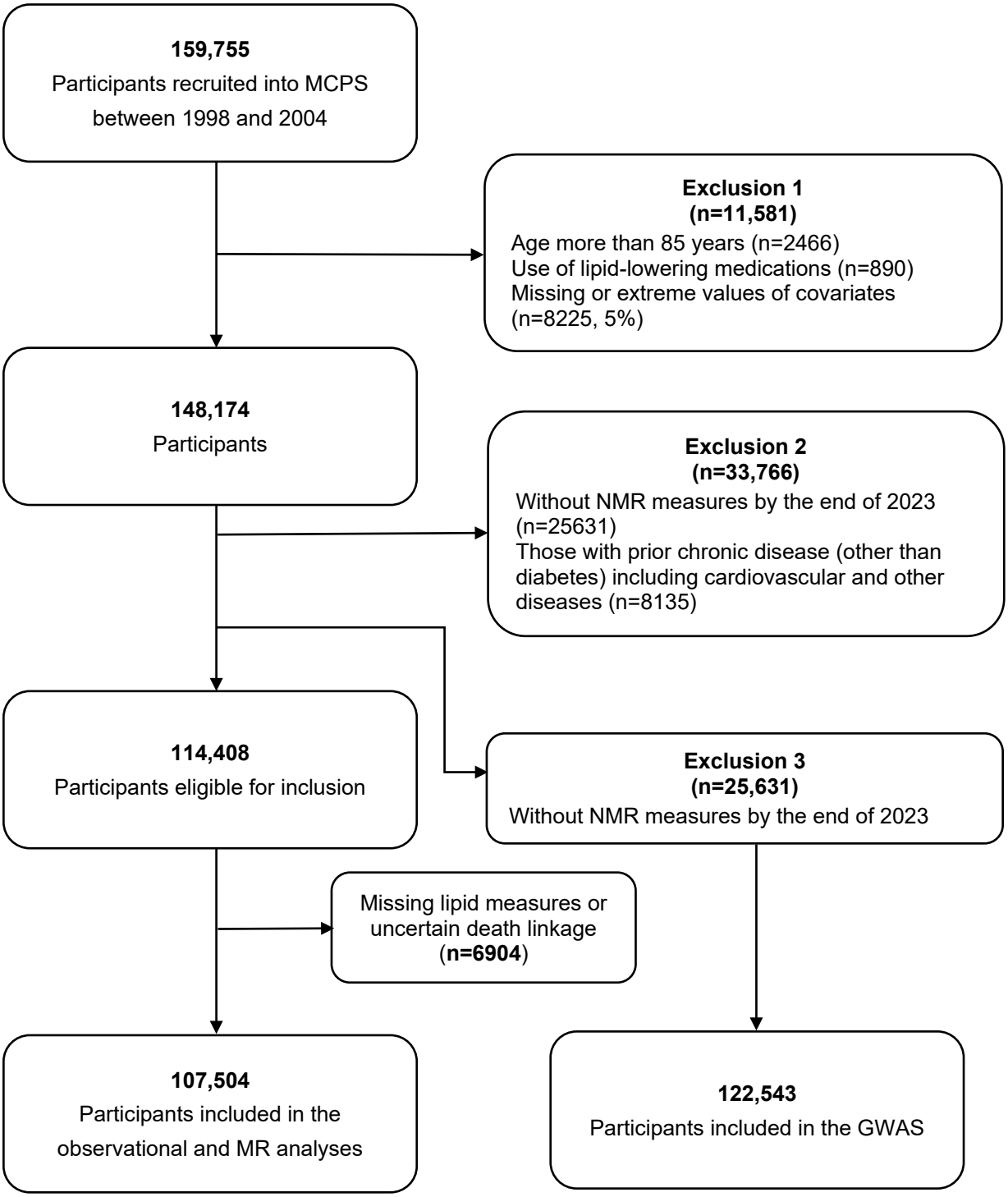
Recruitment patterns represent the monthly numbers of participants enrolled into the Mexico City Prospective Study from two neighbouring districts in Mexico City (Coyoacán and Iztapalapa). Recruitment stopped between mid-2003 and early 2004 due to a meteorological disaster in the country.

Figure 3.2 Distribution of major NMR lipid measures before data transformation



The units were g/L for ApoA1 and ApoB, and mmol/L for other lipid traits.

Figure 3.3 Flowchart of participant selection



CHAPTER 4: Cross-sectional associations of NMR lipid traits with demographic and lifestyle factors

Chapter summary

Although NMR-based metabolomics provides detailed measurements of the lipidome profile and advances the prediction of cardiometabolic diseases, their correlation patterns with each other, with the clinical chemistry lipid panel, and with sociodemographic and lifestyle factors remains limited, especially in Latin American populations. Based on 107,504 MCPS participants with NMR data available and without chronic disease or use of lipid-lowering drugs at baseline, the current chapter systematically assesses the correlation patterns within the NMR lipid panel and their cross-sectional associations with clinical chemistry lipids, as well as with multiple participant and sample-related characteristics at baseline. Complex correlation patterns within and across lipoprotein categories were observed, largely in line with their biochemical structures and interactions in lipid metabolism. These complex correlation patterns reflect the nature of lipoprotein metabolism and make exclusive interpretations on the effects of individual components of lipid traits challenging, justifying the use of both univariable and multivariable methods to quantify the independent and joint association of lipid traits with ASCVD risk in observational and MR analyses. A high correlation between NMR spectroscopy and clinical biochemistry confirmed the robustness of baseline lipid measures in MCPS. NMR lipid traits were also associated with a range of baseline characteristics, such as age, sex, region, education, and smoking, which are key risk factors for ASCVD and should be considered as confounding factors in subsequent observational analyses to limited potential bias.

4.1 Background

Lipid metabolism is a continuum and the composition of lipid particles in the circulation is physiologically correlated. Given the complexities and inter-relatedness of lipid traits, a comprehensive NMR approach might be more preferable and informative than use of traditional chemical assays to quantify the concentration and size of lipoprotein subfractions and their lipid contents that may have contrasting relevance to the ASCVD risk. Previous observational studies of NMR lipid traits primarily in European populations (**Chapter 2**) supported the evidence based on traditional lipid panels that LDL and other apoB-containing lipoproteins are directly implicated in the development of ASCVD, as introduced in **Chapter 1**. However, questions remain on the correlation between NMR-measured lipid profiles, and their association with traditional lipids measured by chemical assays and confounding factors including demographic, lifestyle, and sample-processing characteristics. Before investigating the prospective associations between lipid traits and ASCVD risk, it is important to clarify the associations between potential confounders and lipid traits at baseline. Assessing the intercorrelation of lipid traits may also inform the use of pleiotropy-robust methods in the MR analyses to avoid biased results.

Several studies of lipids and ASCVD also reported a modest to high correlation between lipid traits measured by NMR spectroscopy and clinical chemistry.[1-3] Lipid-related biomarkers in the NMR metabolomics array correlate with each other according to their functional, biochemical, and metabolic properties.[1,3,4] A UK study using the same NMR platform as in the MCPS suggested that 14 lipoprotein subfractions (VLDL, IDL, LDL, and HDL) were positively correlated with one another (range of correlation coefficients, 0.33 to 0.91) and with clinical chemistry-measured LDL-C, and were negatively correlated with HDL-C (-0.42 to -0.33).[1] In the US Women's health Study, concentration of total LDL particles was strongly correlated with LDL-C, apoB, and non-HDL-C (0.66 to 0.78), and was correlated with small (0.63) but not medium (0.03) LDL particles. In the China Kadoorie Biobank study, major lipid-related traits (TC, LDL-C, HDL-C, and TG) measured by clinical chemistry were strongly correlated with those measured by NMR

spectroscopy (0.77 to 0.90). However, the correlation patterns of NMR lipid traits in Latin American population are largely unexplored.

Blood lipid levels are affected by genes, lifestyle, illness, diet, hormone levels, and other environmental factors.[5] Studies of traditional lipid profiles have suggested that lipoproteins and lipids also vary by age and sex.[6,7] Premenopausal women tend to have a less atherogenic lipid profile (higher HDL-C, lower LDL-C, VLDL-C, and TG, and larger average size of LDL) and lower risk of CVDs than age-matched men. These differences tend to be only apparent after puberty and before reproductive senescence. The mechanisms for this cardioprotective effect of female sex may include the effects of sex hormones such as estrogens and sex-chromosomes in tissues. Studies of age-related trends in lipid levels suggested that TC, LDL-C, and TG levels increase up to middle age and then decrease,[8] and have reported linear or curvilinear effects of age in both sexes.[9,10]

Similar to age and sex, lifestyle factors such as smoking and alcohol drinking may affect the blood lipid profile. Smoking is a major causal risk factor for ASCVD and may exert its atherogenic effects through altering blood lipid levels. A meta-analysis suggested a dose-response effect of smoking on blood cholesterol levels, with higher TC, TG, LDL-C, and VLDL-C, and lower HDL-C and apoA1 in smokers than in nonsmokers.[11] Alcohol drinking showed a dose-response relationship with high HDL-C and apoA1 in another meta-analysis,[12] which indirectly support the protective effect of moderate alcohol consumption on CHD observed in several observational studies.[13] Moreover, obesity is associated with alterations in lipid metabolism, insulin resistance, and metabolic syndrome, and increased risk of ASCVD. About 60-70% of individuals with obesity had dyslipidemia including increased TG, VLDL, small dense LDL, apoB, and non-HDL-C levels, and decreased HDL-C and apoA1.[14] Chronic diseases such as CVDs, diabetes, and kidney diseases may also affect blood lipid levels. However, participants with chronic diseases at baseline were excluded in the subsequent observational regression analyses to avoid

reverse causality, the relationships of chronic diseases with lipid profile are not assessed in this thesis.

Overall, although previous studies largely based on traditional lipid panels suggest that blood lipid levels may vary by demographic and lifestyle factors, the variation of NMR lipid profile by these factors remains unclear, especially in Latin American population. Given the complexity of lipid metabolism, further studies with large sample sizes and detailed measures of lipid traits are needed to elucidate the differences in the NMR lipid profile between confounding factors and the underlying mechanisms. The objectives of this chapter are to: a) Characterize the correlation between NMR-measured lipid profile; b) Assess the association between NMR-measured lipid profile and traditional lipids measured by chemical assay; and c) Explore the cross-sectional relationship of NMR-measured lipid profile with a set of demographic, lifestyle, and sample-processing characteristics at baseline.

4.2 Methods

4.2.1 Study design and study population

The overall study design and recruitment of participants is described in **Chapter 3**. Briefly, participants aged 35 to 84 years at baseline who had no prior history of self-reported chronic diseases except diabetes (cardiovascular disease, cancer, or chronic renal, hepatic, or pulmonary disease), were not using lipid-lowering treatment, with complete data on NMR metabolomics and covariates, and with reliable mortality matches were included. For the cross-sectional analyses of the associations between adiposity and NMR lipid traits, participants with diabetes were also excluded. Baseline characteristics of participants included in this chapter are shown in **Chapter 3**.

4.2.2 NMR lipid-related traits

Two sets of lipid-related phenotypes were used in the analyses. All lipid-related traits were measured by high-throughput NMR metabolomics platform (the Nightingale Health

Ltd.), as detailed in **Chapter 3**. The first array of *lipoprotein subfractions* includes 107 NMR lipid-related measurements out of 228 lipid and other metabolic measurements, representing a spectrum of lipoprotein particle subclasses with heterogeneous composition and size, as well with different metabolic and functional characteristics. The second array of *composite lipid traits* includes major blood lipids (TG, LDL-C, and HDL-C) and apolipoproteins (ApoB and ApoA1), which are consistent with the traditional lipid panels used in clinical practice and research.

4.2.3 Cross-sectional analyses

To assess the correlation pattern of 107 NMR lipid-related measurements, Pearson's correlation coefficients were calculated for each pair of log-transformed NMR lipid traits. Similarly, the correlation pattern of 5 NMR composite lipid traits were assessed using Pearson's correlation coefficients.

The composite lipid traits including TG, LDL-C, HDL-C, ApoB, and ApoA1 were measured by clinical chemistry at the CTSU Wolfson laboratory (Oxford, United Kingdom) in a random sample of 1000 participants. NMR-measured and clinical chemistry-measured composite lipid traits were both available for 521 samples. These were then compared using scatterplots and Pearson's correlation coefficients.

To assess the variation of composite lipid traits by baseline characteristics, differences (in SD units) from the overall study population mean of log-transformed composite lipid traits across levels of demographic (age, sex, BMI, region, and education level), lifestyle (smoking and alcohol consumption), and sample-processing characteristics (fasting duration) were calculated. The mean valuables were adjusted simultaneously for other relevant baseline characteristics, when appropriate. To explore the shape of association with age, the trajectories of lipid levels with age were fitted by locally estimated scatterplot smoothing (LOESS) regression stratified by sex.

4.3 Results

4.3.1 Correlations of NMR lipid traits at baseline

Correlation patterns of 107 NMR-measured lipid traits, including lipoproteins, lipids in lipoprotein subfractions, particle sizes, and apolipoproteins, are shown in **Figure 4.1**. Overall, lipid subfractions formed metabolic clusters consisting of strongly and positively correlated lipid traits, according to their relevant biochemical and metabolic characteristics, such as VLDL, IDL, LDL, and HDL as well as their lipid contents. There were moderate to high correlations across the clusters according to the properties of lipid metabolism, such the highly positive correlations between ApoB-containing lipoprotein particles (VLDL, IDL, and LDL) and ApoB, and between HDL related traits and ApoA1. The concentration of TG in VLDL and VLDL levels were negatively with LDL diameters and cholesterol in HDL subfractions, which was consistent with the patterns in atherogenic dyslipidemia (hypertriglyceridemia, low HDL-C, and high small dense cholesterol-depleted LDL particles) in certain insulin-resistant conditions (e.g., obesity, metabolic syndrome, and type 2 diabetes). Heterogeneity for lipoprotein particles with varying sizes was also observed. For HDL particles, smaller particles were positively and larger particles negatively related to ApoB-containing particles.

Correlation patterns of composite lipid traits, including LDL-C (S/L-LDL-C), HDL-C (S/L-HDL-C), TG, ApoB, and ApoA1, overall and by sex are shown in **Figure 4.2**. The overall correlations were generally consistent with that of detailed NMR lipid spectrum, with strong positive correlations between ApoB-containing particles and negative correlations of TG and VLDL with LDL (r=-0.38 for TG and LDL-D) and HDL (r=-0.24 for VLDL-D and LDL-D) size. TG was more strongly related to S-LDL-C (0.38) than L-LDL-C (0.19). Notably, the negative correlation between TG and LDL size was more pronounced in men than in women (-0.47 vs. -0.31).

4.3.2 Correlations of NMR- and clinical chemistry-measured composite lipid traits

The correlations of composite lipid traits (LDL-C, HDL-C, TG, ApoB, and ApoA1) measured by NMR versus chemical assay are shown in **Figure 4.3**. There were moderate to high correlations for LDL-C (r=0.81), HDL-C (r=0.72), TG (r=0.79), ApoA1 (r=0.67), and ApoB (r=0.84).

4.3.3 Correlations of NMR-measured composite lipid traits and baseline characteristics

4.3.3.1 Demographic characteristics

Age and sex

The trajectories of major lipid traits with age by sex are shown in **Figure 4.4**. Non-linear or linear associations of lipid traits with age were observed and broadly consistent in men and women. Overall, most traits exhibited a linear increase either across all age groups or up to the ages of 55-65, after which their levels either peaked and then declined or stabilized in older age. For HDL-related traits and ApoA1, men consistently exhibit higher levels than women across all age groups. In contrast, for ApoB-containing particles and TG, men had higher levels than women before the age of 45-50, but lower levels thereafter. The trends for lipoprotein subfractions were broadly similar with the composite lipid traits (e.g., S-LDL-C or L-LDL-C vs LDL-C; S-HDL-C or L-HDL-C vs HDL-C).

Mean differences from population mean value of NMR lipid traits by sex are shown in **Table 4.1**. Lipid traits showed variation by sex that broadly aligned to their relevant biochemical and metabolic clusters. ApoB, LDL-C, LDL-D, HDL-C, and HDL-D were generally higher in women than in men. VLDL related traits and TG were generally higher in men than in women. The largest differences were observed for HDL-D, with a 0.32 population SD lower in men than in women. Mean differences from population mean value of NMR lipid traits by age groups are shown in **Table 4.2**. Overall, most lipid subfractions showed a curvilinear relationship across age groups, with individuals at 55-64 years old had the highest levels.

District of residence and educational level

Mean differences from population mean value of lipid traits by district of residence (Coyoacán and Iztapalapa) or different levels of education are shown in **Table 4.2**. Modest differences were observed for most lipid traits across regions (0.1-0.2 SD). The greatest differences were observed for VLDL and HDL particles, average lipoprotein sizes (VLDL-D and HDL-D), and ApoA1. Most lipid traits varied across levels of educational attainment. VLDL and HDL related traits, and TG were lower in individuals with higher education level, whereas IDL and LDL related traits were higher in individuals with higher education level. The largest difference was observed for HDL-D, with about 0.4 SD differences between the highest and lowest levels of education.

4.3.3.2 Lifestyle characteristics

Smoking

Mean differences from population mean value of lipid traits by different levels of smoking are shown in **Table 4.3**. Increasing smoking levels (from never smoking to smoking more than 10 cigarettes per day) were correlated with higher levels of VLDL, LDL, TG, ApoB, and ApoB to ApoA1 ratio, and with lower levels of HDL-C, HDL-D, ApoA1, and LDL-D. The largest difference was observed for HDL-D, with about 0.4 SD differences between the highest and lowest levels of smoking.

Alcohol consumption

Mean differences from population mean value of lipid traits by different levels of alcohol consumption are shown in **Table 4.3**. There were modest differences for most lipid traits between alcohol consumption levels. For example, HDL particles and ApoA1 tended to be higher in those with higher alcohol consumption.

BMI

Mean differences from population mean value of lipid traits by obesity are shown in **Table 4.3**. There were strong differences for most lipid traits between individuals with and

without obesity, with those with obesity had higher VLDL (e.g., mean population SD 0.4 for XXL-VLDL-P), and TG levels across lipoproteins (e.g., mean population SD 0.25 for overall TG).

4.3.3.3 Sample-processing characteristics

Fasting duration

Mean differences from population mean value of lipid traits by fasting duration are shown in **Table 4.3**, with strong differences for most lipid traits between fasting duration groups, as expected. For example, lipid traits related to most VLDL subclasses (particularly chylomicrons), TG, XL- and L-HDL, VLDL-D, and HDL-D were lower in individuals with longer fasting duration (12+h and 8-12h versus <8h). Cholesterol in VLDL, IDL, and LDL, ApoB, ApoB to ApoA1 ratio and LDL-D were higher in those with longer fasting duration.

4.4 Discussion

We assessed the cross-sectional association of NMR-measured lipid traits of lipoprotein subfractions (size, lipid content, and concentration) and apolipoproteins with socio-demographic, lifestyle, and sample-processing characteristics in the currently largest Latin American population.

4.4.1 Correlations of NMR lipid traits at baseline

Correlation patterns of NMR lipid biomarkers within biochemical and metabolic groups were generally consistent with those observed in previous studies. Chemically similar lipid subfractions within the same lipoprotein category, such as the cholesterol in VLDL with varying particle sizes, tend to be positively correlated as they represent the continuum of lipid metabolism. Lipids, lipoproteins, and apolipoproteins were bound to be correlated with each other due to the chemical structure. For example, ApoB was positively related to all ApoB containing particles and their lipid contents, such as VLDL, IDL, and LDL, as each particle contains exactly one ApoB molecule, which formed during the lipolysis and of triglyceride-rich lipoproteins by the enzyme LPL, HL, and CETP.

Chemically differentiated lipid traits, such as LDL particle size, HDL-C, and TG, were also correlated due to shared lipid metabolism pathways and lipid metabolic enzymes. For example, extremely high TG levels in certain conditions such as obesity and type 2 diabetes are associated with low levels of HDL-C and smaller, denser LDL particles through the action of CETP and HL, referred to as atherogenic dyslipidemia (see **Chapter 1**).[15,16]

These complex correlation patterns reflect the nature of lipoprotein metabolism and make exclusive interpretations on the effects of individual components of lipid traits challenging,[17] justifying the use of both univariable and multivariable methods to quantify the independent and joint association of lipid traits with ASCVD risk in observational and MR analyses.

4.4.2 Correlations of NMR-measured and clinical chemistry-measured composite lipid traits

The high correlations between NMR-measured and clinical chemistry-measured lipid traits are largely aligned with those from previous studies from China[2] and UK[18,19]. Of note, the mean value of NMR-measured LDL-C (1.73 mmol/L) was lower than that of clinical chemistry-measured LDL-C (2.81 mmol/L) in MCPS and means of LDL-C in external cohorts (e.g., 2.20 mmol/L in CKB). However, these differences in absolute concentration of LDL-C would not influence the relative risk of ASCVD associated with between-person *differences* in the measures. Overall, the consistency between analytical methods (NMR and chemistry assay), as observed previously, [2,18,19] suggests the robustness of baseline lipid measurements in the current study.

4.4.3 Correlations of NMR-measured composite lipid traits and baseline characteristics

Most lipid traits were correlated with multiple demographic, lifestyle, and sample-processing baseline characteristics that are also risk factors for multiple chronic diseases, which may partly mediate or confound the association between lipid traits and the risk of ASCVD.

4.4.3.1 Age and sex

We observed HDL-C and ApoA1 are consistently higher in women than in men across all age groups. LDL-C and TG are higher in men before ages 45-50 years but higher in women afterward. These findings are consistent with a previous study in UK, suggesting the VLDL are higher in men whereas LDL and HDL are higher in women.[20] Our findings further showed the differences by sex may vary according to age. These sex-based differences with age may partly be explained by the variation in sexual hormones, with such differences in lipids are stronger in adults than pre-pubescent individuals. As we did not account for other covariables, these differences are also likely to be related to the variation in lifestyle factors (e.g., smoking, alcohol consumption, sleep duration, physical activity, and diet), physical measures (e.g., BMI and blood pressure), and prevalent chronic diseases (e.g., diabetes).

4.4.3.2 District of residence and educational level

Most lipid traits varied by study region and educational level, reflecting the differences related to socio-economic conditions. Participants from more socio-economically deprived area (Iztapalapa) had a more atherogenic lipid profile, such as higher VLDL and ApoB. Similarly, VLDL and TG were lower in individuals with higher education levels. However, there was an opposite direction for other atherogenic lipid traits such as LDL-C and IDL-C, which were higher in individuals with higher education levels. Socio-economic status is one of the key risk factors for cardiometabolic diseases and death, with the underlying mechanisms including limited access to healthcare, unhealthy lifestyle choices, increased stress levels, and reduced opportunities for physical activity, which may partially explain the observed variations in lipid levels. A previous study showed the sex differences in the association between socio-economic status and lipid levels and the related variations may be largely attributed to potentially modifiable factors such as obesity, physical activity and dietary intake.[21] LDL-C and TC levels were not related to educational class indices among men, but a positive relationship was observed

among women, after adjusting for age, BMI, and alcohol use.[21] Further investigations that stratified by sex and account for additional covariables are warranted.

4.4.3.3 BMI and obesity

We observed that obesity was strongly associated with hypertriglyceridemia, driven by increasing levels of VLDL particles and total TG, as well as decreased HDL-C and smaller LDL size, whereas LDL-related traits (except TG) were largely unaffected. These findings are consistent with the pattern of atherogenic dyslipidemia in certain insulin-resistant conditions (e.g., obesity, metabolic syndrome and type 2 diabetes)[15,16] and with previous studies in geographically diverse populations.[22,23] Considering the extremely high level of obesity and TG in the study population, BMI and obesity should be accounted for in the analyses of association with ASCVD.

4.4.3.4 Smoking and alcohol consumption

We observed that smoking was related to atherogenic lipid profiles, such as higher ApoB to ApoA1 ratio and TG, and lower HDL-C, which is consistent with previous studies.[11,24,25] The potential explanations included oxidative stress, inflammation, and the adverse effects of smoking on lipid metabolism (e.g., decreased activity of LPL in heavy smoker).[24] Smoking is also associated with increased BMI, central obesity, and other lifestyle factors that may affect plasma lipid levels.[26]

We observed that alcohol consumption was related to modest differences in ApoA1 and HDL-C and no obvious changes in TG and other lipid traits which is consistent with previous studies.[12] Previous studies have suggested that such changes underlie the observed protective effect of moderate alcohol consumption on ASCVD in observational studies,[12] although such adverse association may be biased by reverse causality and confounding and contradict with the findings from genetic studies where reduction of alcohol consumption is beneficial for cardiovascular health.[27] Other explanations include the association of alcohol drinking and multiple adverse health conditions that may also

directly affect lipid metabolism, including metabolic syndrome, chronic liver disease, hepatocellular carcinoma.[28]

4.4.3.5 Fasting duration

We observed modest differences in NMR lipid measurements by fasting duration, a sample processing related factor in NMR metabolomics. The slightly increased VLDL levels, especially the chylomicrons and very large VLDLs, and TG in non-fasting samples, reflect the elevated levels of intestinal-derived lipoproteins after meal. In line with previous studies, levels of cholesterol and HDL-C differed little across the groups of fasting duration, suggesting that fasting for clinical routine lipid panels is largely unnecessary.[29] Fasting only accounted for less than 2% of the variation in most NMR metabolites.[30] Non-fasting lipid panels may better reflect an individual's typical lipid levels throughout the day, providing a more accurate assessment of cardiovascular risk.

4.4.4 Strength and limitations

This chapter has systematically evaluated the correlation patterns within the NMR lipid panel and their association with the clinical chemistry lipid panel, as well as with various lifestyle and socioeconomic factors, which were previously assessed in mainly European populations. The large sample size also enabled the detailed assessment of correlation patterns within NMR metabolites and across multiple baseline characteristics. Several key limitations are also noted. As the aim of the current chapter is *not* to assess the independent association of a specific factor with multiple lipid characteristics, the cross-sectional analyses were univariable and did not take into account other covariates. Variations in plasma lipids are likely to be jointly influenced by prevalent diseases, lifestyle factors and physical measures. Further investigations with bespoke study designs should further assess the independent and combined effects of participant characteristics and NMR methods. We excluded participants with chronic diseases at baseline and therefore did not assess their effect on the lipid profile, as this is beyond the scope of this chapter. Finally, a potential future direction is to investigate the impact

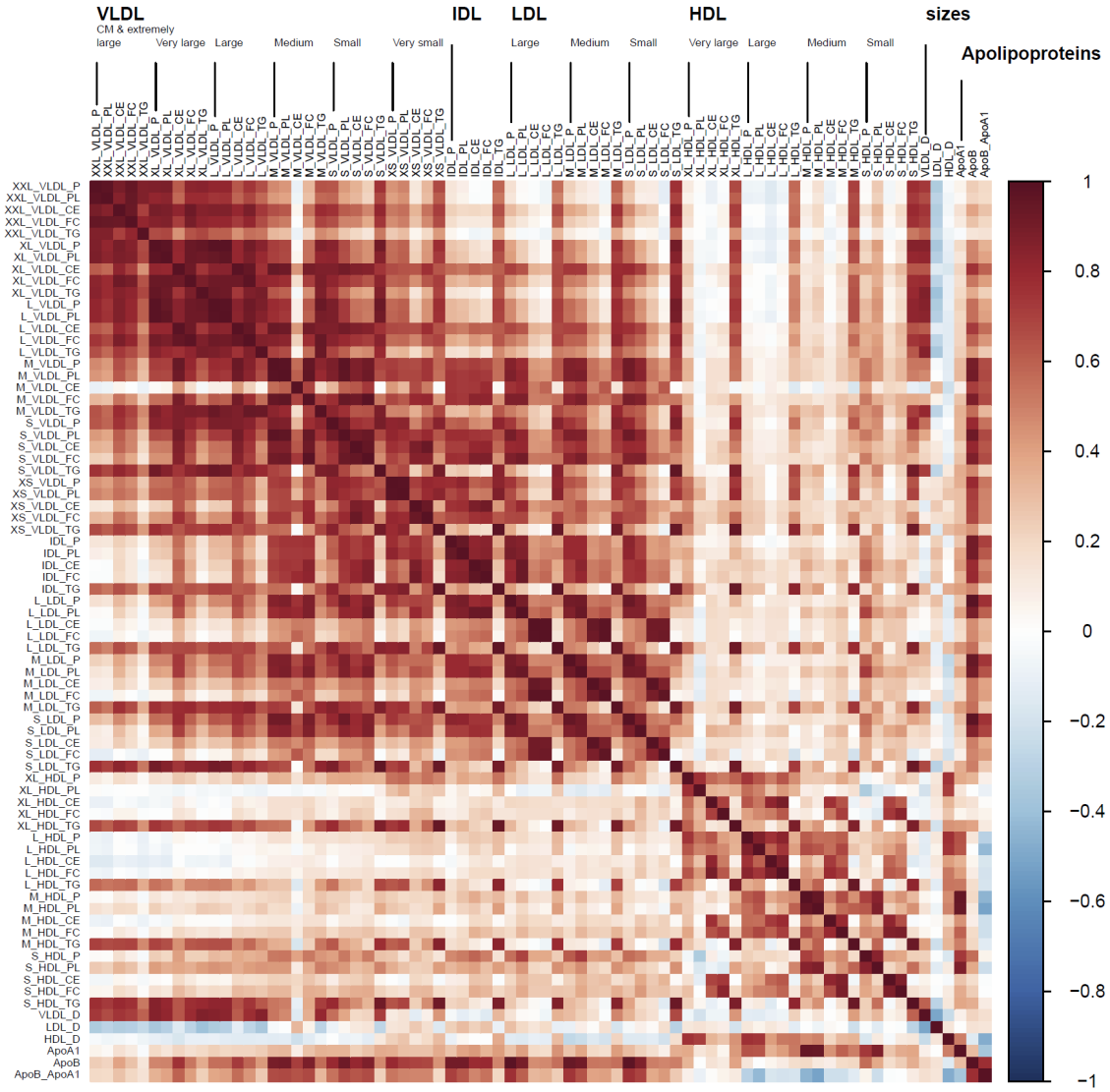
of fasting duration more directly by examining repeated measurements taken at varying fasting intervals in the same individuals. Overall, the association with multiple baseline characteristics should be considered preliminary, with the aim of informing subsequent analyses of the association with ASCVD, rather than their specific individual effect, which deviated from the main aims of the current thesis.

4.5 Conclusion

In summary, complex correlation patterns within and across the same lipoprotein category within NMR lipid panels were observed, which are largely in line with their biochemical structures and interaction in lipid metabolism. High correlation between NMR spectroscopy and clinical biochemistry confirmed the robustness of baseline lipid measures in MCPS. NMR lipid traits were also associated with a range of baseline characteristics, such as age, sex, region, education, and smoking, which are key risk factors for ASCVD and should be considered as confounding factors in subsequent observational analyses to limited potential bias.

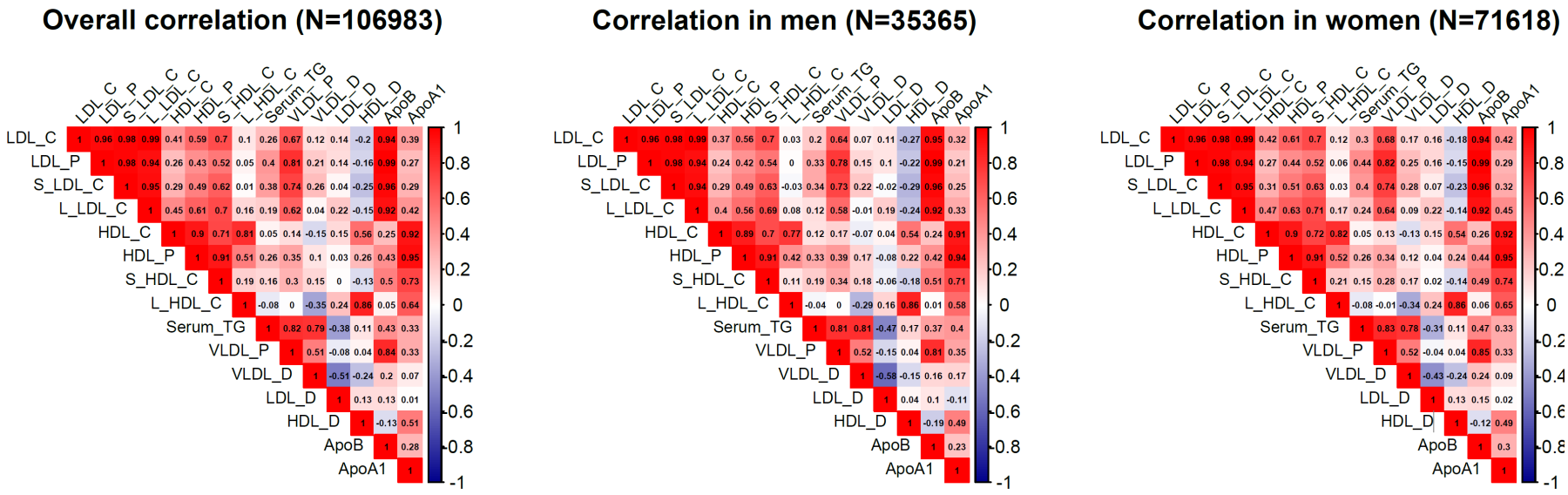
Figure 4.1. Correlation patterns of NMR lipid traits

Lipoproteins, lipids by lipoprotein subclasses, particle sizes, and apolipoproteins



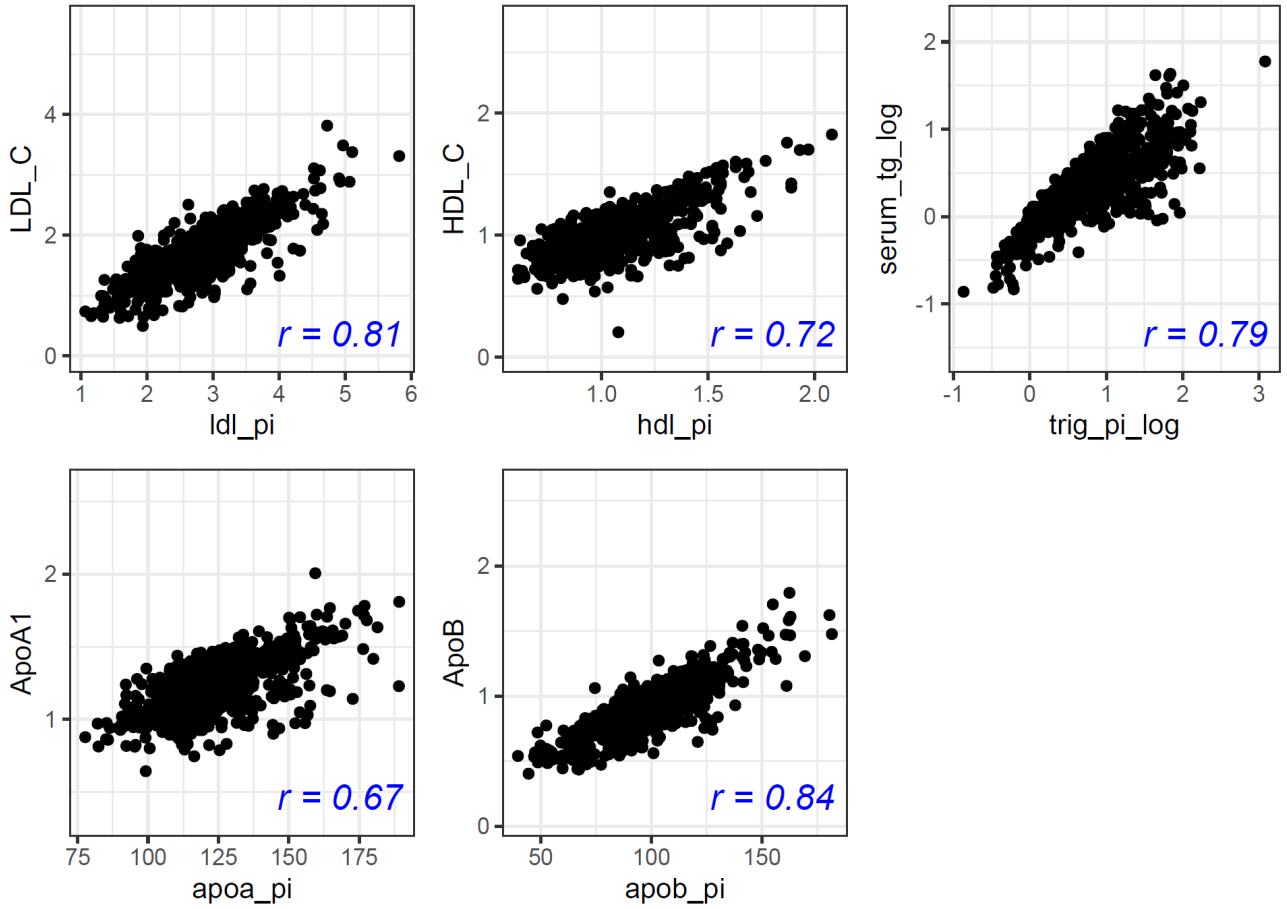
AcAce=acetoacetate; Ace=acetate; Ala=alanine; Alb=albumin; Apo−A1=apolipoprotein A1; Apo−B=apolipoprotein B; bOHBut=3−hydroxybutyrate; Cit=citrate; Crea=creatinine; DHA=docosahexaenoic acid; FA=fatty acids; FAw3=omega−3 fatty acids; FAw6=omega−6 fatty acids; Glc=glucose; Gln=glutamine; Glyc−A=glycoprotein acetyls; HDL=high density lipoproteins; HDL−D=high density lipoprotein particle diameter; His=histidine; IDL=intermediate density lipoproteins; Ile=isoleucine; L=large; LA=linoleic acid; Lac=lactate; LDL=low density lipoproteins; LDL−D=low density lipoprotein particle diameter; Leu=leucine; LP=lipoprotein; M=medium; MUFA=monounsaturated fatty acids; PC=phosphatidylcholines; Phe=phenylalanine; PUFA=polyunsaturated fatty acids; S=small; SFA=saturated fatty acids; SM=sphingomyelins; TotFA=total fatty acids; TotCho=total cholines; Tyr=tyrosine; Val=valine; VLDL=very low density lipoproteins; VLDL−D=very low density lipoprotein particle diameter; XL=very large; XS=very small; XXL=extremely large.

Figure 4.2. Correlation patterns of composite lipid traits



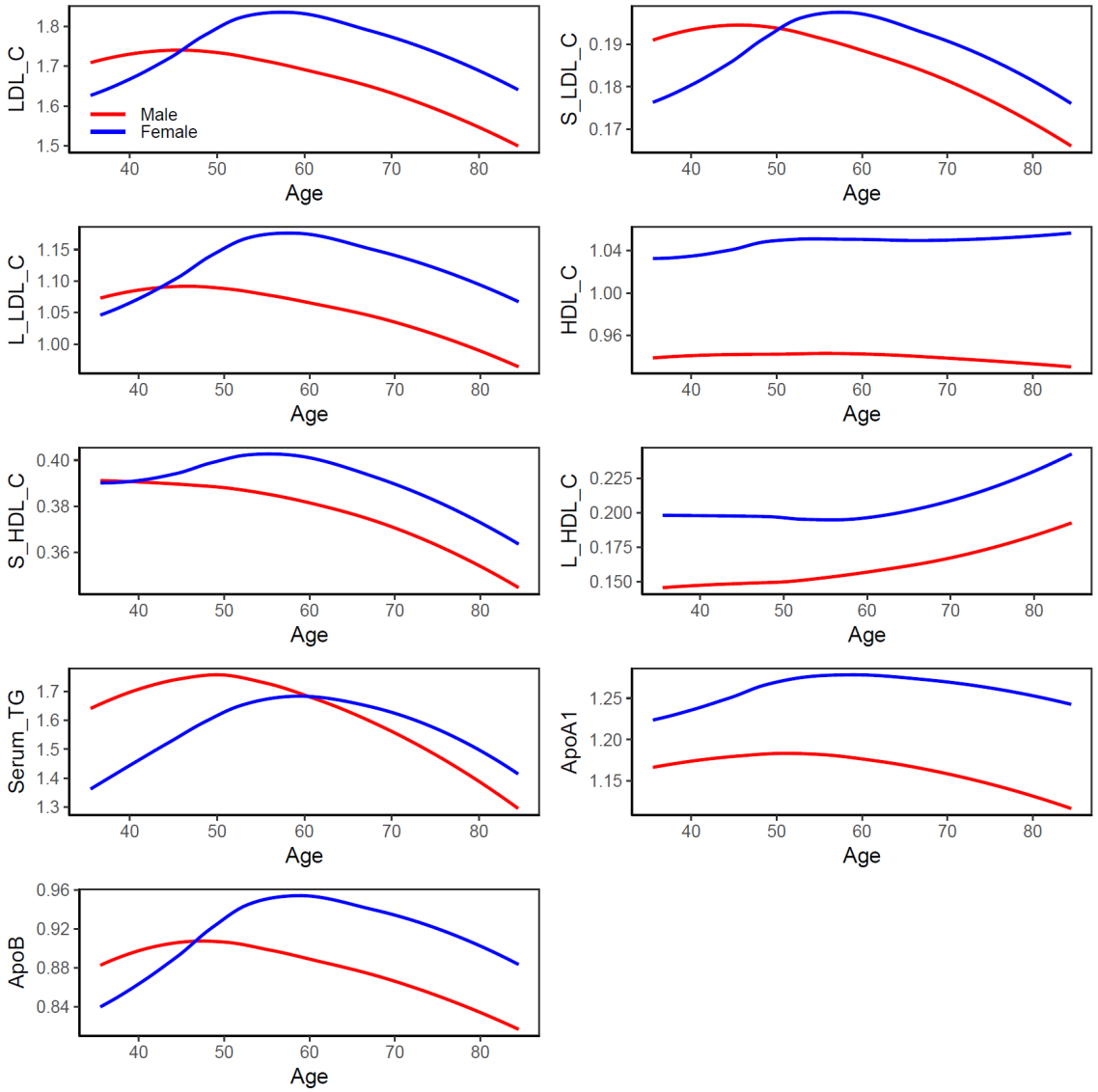
Analyses excluded those with chronic diseases at baseline (except diabetes). VLDL=very low density lipoproteins; LDL=low density lipoproteins; HDL=high density lipoproteins; D, diameter; L=large; S=small.

Figure 4.3. Comparison of lipid traits measured by NMR versus clinical chemical assay



Analyses included 521 participants with both NMR and clinical chemical lipid measures (LDL-C, HDL-C, TG, ApoA1, and ApoB). The x-axis value was lipid trait measured by clinical chemistry assay and the y-axis value was NMR lipid measures.

Figure 4.4. Trajectory of major lipid traits with age by sex



Analyses excluded those with chronic diseases at baseline (except diabetes). Trajectories were fitted by Loess regression. LDL=low density lipoproteins; HDL=high density lipoproteins; D, diameter; L=large; S=small.

Table 4.1. NMR lipid traits by sex

ID	Biomarker	Men Difference (SD units) from mean log-NMR biomarkers	Women Difference (SD units) from mean log-NMR biomarkers	All Mean(SD)
N		35518	71986	.
1	Concentration of chylomicrons and extremely large VLDL particles (XXL-VLDL-P) (mmol/l)	0.12	-0.06	2435.3 (1768.6)
2	Concentration of very large VLDL particles (XL-VLDL-P) (mmol/l)	0.15	-0.08	4666.4 (2495.9)
3	Concentration of large VLDL particles (L-VLDL-P) (mmol/l)	0.15	-0.08	13080 (6084.8)
4	Concentration of medium VLDL particles (M-VLDL-P) (mmol/l)	0.03	-0.01	38644 (13069)
5	Concentration of small VLDL particles (S-VLDL-P) (mmol/l)	0.05	-0.02	46638 (13998)
6	Concentration of very small VLDL particles (XS-VLDL-P) (mmol/l)	-0.16	0.08	61135 (15178)
7	Concentration of IDL particles (IDL-P) (mmol/l)	-0.17	0.08	317471 (76567)
8	Concentration of large LDL particles (L-LDL-P) (mmol/l)	-0.07	0.04	776055 (189351)
9	Concentration of medium LDL particles (M-LDL-P) (mmol/l)	0.05	-0.03	313767 (89514)
10	Concentration of small LDL particles (S-LDL-P) (mmol/l)	0.00	0.00	181794 (40511)
11	Concentration of very large HDL particles (XL-HDL-P) (mmol/l)	-0.21	0.10	213217 (60161)
12	Concentration of large HDL particles (L-HDL-P) (mmol/l)	-0.31	0.15	919947 (430807)
13	Concentration of medium HDL particles (M-HDL-P) (mmol/l)	-0.26	0.13	2.94E6 (761327)
14	Concentration of small HDL particles (S-HDL-P) (mmol/l)	-0.09	0.04	8.65E6 (1.43E6)
15	Cholesterol in chylomicrons and extremely large VLDL (XXL-VLDL-C) (mmol/l)	0.10	-0.05	0.08 (0.05)
16	Cholesterol in very large VLDL (XL-VLDL-C) (mmol/l)	0.10	-0.05	0.07 (0.03)
17	Cholesterol in large VLDL (L-VLDL-C) (mmol/l)	0.10	-0.05	0.1 (0.0)
18	Cholesterol in medium VLDL (M-VLDL-C) (mmol/l)	-0.09	0.04	0.2 (0.1)
19	Cholesterol in small VLDL (S-VLDL-C) (mmol/l)	-0.05	0.02	0.2 (0.1)
20	Cholesterol in very small VLDL (XS-VLDL-C) (mmol/l)	-0.19	0.09	0.2 (0.0)
21	Cholesterol in IDL (IDL-C) (mmol/l)	-0.18	0.09	0.8 (0.2)
22	Cholesterol in large LDL (L-LDL-C) (mmol/l)	-0.05	0.03	1.1 (0.3)
23	Cholesterol in medium LDL (M-LDL-C) (mmol/l)	0.00	0.00	0.4 (0.1)
24	Cholesterol in small LDL (S-LDL-C) (mmol/l)	0.00	0.00	0.2 (0.1)
25	Cholesterol in very large HDL (XL-HDL-C) (mmol/l)	-0.11	0.06	0.07 (0.02)
26	Cholesterol in large HDL (L-HDL-C) (mmol/l)	-0.19	0.09	0.2 (0.1)
27	Cholesterol in medium HDL (M-HDL-C) (mmol/l)	-0.14	0.07	0.4 (0.1)
28	Cholesterol in small HDL (S-HDL-C) (mmol/l)	-0.03	0.01	0.4 (0.1)
29	Free cholesterol in chylomicrons and extremely large VLDL (XXL-VLDL-FC) (mmol/l)	0.10	-0.05	0.03 (0.02)
30	Free cholesterol in very large VLDL (XL-VLDL-FC) (mmol/l)	0.12	-0.06	0.03 (0.02)
31	Free cholesterol in large VLDL (L-VLDL-FC) (mmol/l)	0.14	-0.07	0.06 (0.03)
32	Free cholesterol in medium VLDL (M-VLDL-FC) (mmol/l)	-0.03	0.01	0.09 (0.03)
33	Free cholesterol in small VLDL (S-VLDL-FC) (mmol/l)	-0.05	0.02	0.06 (0.02)
34	Free cholesterol in very small VLDL (XS-VLDL-FC) (mmol/l)	-0.15	0.07	0.06 (0.02)
35	Free cholesterol in IDL (IDL-FC) (mmol/l)	-0.19	0.09	0.2 (0.1)
36	Free cholesterol in large LDL (L-LDL-FC) (mmol/l)	-0.07	0.03	0.3 (0.1)
37	Free cholesterol in medium LDL (M-LDL-FC) (mmol/l)	-0.04	0.02	0.1 (0.0)
38	Free cholesterol in small LDL (S-LDL-FC) (mmol/l)	-0.02	0.01	0.05 (0.01)
39	Free cholesterol in very large HDL (XL-HDL-FC) (mmol/l)	-0.06	0.03	0.02 (0.00)
40	Free cholesterol in large HDL (L-HDL-FC) (mmol/l)	-0.19	0.09	0.05 (0.02)
41	Free cholesterol in medium HDL (M-HDL-FC) (mmol/l)	-0.14	0.07	0.06 (0.02)
42	Free cholesterol in small HDL (S-HDL-FC) (mmol/l)	-0.04	0.02	0.1 (0.0)
43	Cholesteryl esters in chylomicrons and extremely large VLDL (XXL-VLDL-CE) (mmol/l)	0.10	-0.05	0.04 (0.03)
44	Cholesteryl esters in very large VLDL (XL-VLDL-CE) (mmol/l)	0.07	-0.04	0.04 (0.01)

ID	Biomarker	Men	Women	All
		Difference (SD units) from mean log-NMR biomarkers	Difference (SD units) from mean log-NMR biomarkers	Mean(SD)
45	Cholesteryl esters in large VLDL (L-VLDL-CE) (mmol/l)	0.07	-0.03	0.06 (0.02)
46	Cholesteryl esters in medium VLDL (M-VLDL-CE) (mmol/l)	-0.13	0.07	0.09 (0.04)
47	Cholesteryl esters in small VLDL (S-VLDL-CE) (mmol/l)	-0.04	0.02	0.1 (0.0)
48	Cholesteryl esters in very small VLDL (XS-VLDL-CE) (mmol/l)	-0.21	0.10	0.1 (0.0)
49	Cholesteryl esters in IDL (IDL-CE) (mmol/l)	-0.17	0.08	0.6 (0.2)
50	Cholesteryl esters in large LDL (L-LDL-CE) (mmol/l)	-0.05	0.02	0.8 (0.2)
51	Cholesteryl esters in medium LDL (M-LDL-CE) (mmol/l)	0.01	0.00	0.3 (0.1)
52	Cholesteryl esters in small LDL (S-LDL-CE) (mmol/l)	0.00	0.00	0.1 (0.0)
53	Cholesteryl esters in very large HDL (XL-HDL-CE) (mmol/l)	-0.10	0.05	0.05 (0.01)
54	Cholesteryl esters in large HDL (L-HDL-CE) (mmol/l)	-0.16	0.08	0.1 (0.1)
55	Cholesteryl esters in medium HDL (M-HDL-CE) (mmol/l)	-0.10	0.05	0.3 (0.1)
56	Cholesteryl esters in small HDL (S-HDL-CE) (mmol/l)	-0.02	0.01	0.3 (0.1)
57	Triglycerides in chylomicrons and extremely large VLDL (XXL-VLDL-TG) (mmol/l)	0.09	-0.05	0.2 (0.2)
58	Triglycerides in very large VLDL (XL-VLDL-TG) (mmol/l)	0.17	-0.08	0.1 (0.1)
59	Triglycerides in large VLDL (L-VLDL-TG) (mmol/l)	0.18	-0.09	0.2 (0.1)
60	Triglycerides in medium VLDL (M-VLDL-TG) (mmol/l)	0.15	-0.07	0.3 (0.1)
61	Triglycerides in small VLDL (S-VLDL-TG) (mmol/l)	0.12	-0.06	0.2 (0.1)
62	Triglycerides in very small VLDL (XS-VLDL-TG) (mmol/l)	0.00	0.00	0.09 (0.03)
63	Triglycerides in IDL (IDL-TG) (mmol/l)	-0.03	0.02	0.1 (0.0)
64	Triglycerides in large LDL (L-LDL-TG) (mmol/l)	0.00	0.00	0.1 (0.0)
65	Triglycerides in medium LDL (M-LDL-TG) (mmol/l)	0.06	-0.03	0.04 (0.01)
66	Triglycerides in small LDL (S-LDL-TG) (mmol/l)	0.11	-0.05	0.02 (0.01)
67	Triglycerides in very large HDL (XL-HDL-TG) (mmol/l)	-0.01	0.00	0.009 (0.004)
68	Triglycerides in large HDL (L-HDL-TG) (mmol/l)	-0.14	0.07	0.03 (0.02)
69	Triglycerides in medium HDL (M-HDL-TG) (mmol/l)	-0.08	0.04	0.06 (0.02)
70	Triglycerides in small HDL (S-HDL-TG) (mmol/l)	0.06	-0.03	0.06 (0.02)
71	Phospholipids in chylomicrons and extremely large VLDL (XXL-VLDL-PL) (mmol/l)	0.11	-0.05	0.05 (0.04)
72	Phospholipids in very large VLDL (XL-VLDL-PL) (mmol/l)	0.14	-0.07	0.05 (0.03)
73	Phospholipids in large VLDL (L-VLDL-PL) (mmol/l)	0.15	-0.07	0.09 (0.04)
74	Phospholipids in medium VLDL (M-VLDL-PL) (mmol/l)	0.00	0.00	0.1 (0.0)
75	Phospholipids in small VLDL (S-VLDL-PL) (mmol/l)	0.00	0.00	0.1 (0.0)
76	Phospholipids in very small VLDL (XS-VLDL-PL) (mmol/l)	-0.15	0.07	0.1 (0.0)
77	Phospholipids in IDL (IDL-PL) (mmol/l)	-0.21	0.10	0.3 (0.1)
78	Phospholipids in large LDL (L-LDL-PL) (mmol/l)	-0.10	0.05	0.4 (0.1)
79	Phospholipids in medium LDL (M-LDL-PL) (mmol/l)	0.00	0.00	0.2 (0.0)
80	Phospholipids in small LDL (S-LDL-PL) (mmol/l)	0.03	-0.02	0.09 (0.02)
81	Phospholipids in very large HDL (XL-HDL-PL) (mmol/l)	-0.23	0.12	0.07 (0.03)
82	Phospholipids in large HDL (L-HDL-PL) (mmol/l)	-0.29	0.14	0.2 (0.1)
83	Phospholipids in medium HDL (M-HDL-PL) (mmol/l)	-0.25	0.13	0.4 (0.1)
84	Phospholipids in small HDL (S-HDL-PL) (mmol/l)	-0.14	0.07	0.6 (0.1)
85	Total lipids in chylomicrons and extremely large VLDL (XXL-VLDL-L) (mmol/l)	0.12	-0.06	0.3 (0.2)
86	Total lipids in very large VLDL (XL-VLDL-L) (mmol/l)	0.16	-0.08	0.3 (0.1)
87	Total lipids in large VLDL (L-VLDL-L) (mmol/l)	0.17	-0.08	0.4 (0.2)
88	Total lipids in medium VLDL (M-VLDL-L) (mmol/l)	0.06	-0.03	0.6 (0.2)
89	Total lipids in small VLDL (S-VLDL-L) (mmol/l)	0.04	-0.02	0.5 (0.1)
90	Total lipids in very small VLDL (XS-VLDL-L) (mmol/l)	-0.16	0.08	0.4 (0.1)
91	Total lipids in IDL (IDL-L) (mmol/l)	-0.20	0.10	1.2 (0.3)

ID	Biomarker	Men	Women	All
		Difference (SD units) from mean log-NMR biomarkers	Difference (SD units) from mean log-NMR biomarkers	Mean(SD)
92	Total lipids in large LDL (L-LDL-L) (mmol/l)	-0.10	0.05	1.6 (0.4)
93	Total lipids in medium LDL (M-LDL-L) (mmol/l)	0.02	0.00	0.6 (0.2)
94	Total lipids in small LDL (S-LDL-L) (mmol/l)	0.03	-0.02	0.3 (0.1)
95	Total lipids in very large HDL (XL-HDL-L) (mmol/l)	-0.25	0.12	0.1 (0.0)
96	Total lipids in large HDL (L-HDL-L) (mmol/l)	-0.37	0.18	0.4 (0.2)
97	Total lipids in medium HDL (M-HDL-L) (mmol/l)	-0.27	0.13	0.8 (0.2)
98	Total lipids in small HDL (S-HDL-L) (mmol/l)	-0.12	0.06	1.0 (0.2)
99	Average diameter for VLDL particles (VLDL size) (nm)	0.27	-0.13	39.0 (1.1)
100	Average diameter for LDL particles (LDL size) (nm)	-0.25	0.12	23.9 (0.1)
101	Average diameter for HDL particles (HDL size) (nm)	-0.32	0.16	9.5 (0.2)
102	Apolipoprotein A1 (ApoA1) (g/l)	-0.28	0.14	1.2 (0.2)
103	Apolipoprotein B (ApoB) (g/l)	-0.05	0.02	0.9 (0.2)
104	Ratio of apolipoprotein B to apolipoprotein A1 (ApoB/ApoA1) (ratio)	0.14	-0.07	0.7 (0.2)
105	LDL cholesterol (LDL-C) (mmol/l)	-0.03	0.02	1.7 (0.5)
106	HDL cholesterol (HDL-C) (mmol/l)	-0.27	0.14	1.0 (0.2)
107	Total triglycerides (Total-TG) (mmol/l)	0.13	-0.06	1.6 (0.7)

To compare the NMR lipid traits between men and women, the table shows the difference between the standardized mean of log-transformed NMR lipid traits of the relevant comparison group and the standardized mean of log-transformed NMR lipid traits of the study population.

Table 4.2. NMR lipid traits by age, region, and education level

		Difference from study population mean, SD units											
		Age, years						District of residence			Educational level		
	Biomarker	35-44	45-54	55-64	65-74	75-84		Coyoacan	Iztapalapa		University	High school	Elementary	Other
N		38911	30258	19843	12764	5728	.	47683	59821	.	16897	26599	50723	13285
.	XXL_VLDL_P	-0.11	0.08	0.12	0.03	-0.14	.	-0.08	0.07	.	-0.09	-0.06	0.04	0.06
.	XL_VLDL_P	-0.12	0.10	0.14	0.00	-0.23	.	0.00	0.00	.	-0.04	-0.04	0.03	0.00
.	L_VLDL_P	-0.11	0.10	0.14	0.00	-0.23	.	0.06	-0.05	.	-0.02	-0.03	0.03	-0.01
.	M_VLDL_P	-0.11	0.09	0.12	0.00	-0.16	.	0.06	-0.05	.	0.04	0.00	0.01	-0.09
.	S_VLDL_P	-0.18	0.10	0.17	0.08	-0.09	.	0.04	-0.03	.	-0.03	-0.05	0.04	0.00
.	XS_VLDL_P	-0.24	0.09	0.19	0.17	0.11	.	-0.04	0.03	.	-0.07	-0.09	0.05	0.08
.	IDL_P	-0.13	0.08	0.12	0.06	-0.03	.	-0.04	0.03	.	0.05	0.00	0.00	-0.10
.	L_LDL_P	-0.11	0.08	0.11	0.02	-0.07	.	0.03	-0.02	.	0.06	0.00	0.00	-0.09
.	M_LDL_P	-0.09	0.09	0.11	-0.02	-0.18	.	0.03	-0.03	.	0.10	0.02	0.00	-0.15
.	S_LDL_P	-0.12	0.09	0.12	0.01	-0.13	.	0.03	-0.02	.	0.07	0.00	0.00	-0.11
.	XL_HDL_P	-0.13	0.00	0.06	0.17	0.32	.	0.09	-0.08	.	-0.12	-0.09	0.04	0.18
.	L_HDL_P	-0.06	-0.01	0.02	0.10	0.19	.	0.10	-0.08	.	-0.11	-0.05	0.03	0.13
.	M_HDL_P	-0.04	0.04	0.04	-0.01	-0.11	.	0.02	-0.01	.	-0.02	0.01	0.00	-0.02
.	S_HDL_P	0.00	0.08	0.05	-0.12	-0.35	.	0.00	0.00	.	0.11	0.07	-0.02	-0.21
.	XXL_VLDL_C	-0.11	0.09	0.12	0.02	-0.18	.	-0.06	0.05	.	-0.08	-0.05	0.04	0.04
.	XL_VLDL_C	-0.11	0.10	0.13	0.00	-0.20	.	0.00	0.00	.	-0.01	-0.03	0.03	-0.04
.	L_VLDL_C	-0.11	0.10	0.13	0.00	-0.19	.	0.02	-0.02	.	0.00	-0.02	0.02	-0.04
.	M_VLDL_C	-0.07	0.06	0.07	0.01	-0.06	.	0.01	0.00	.	0.07	0.01	0.00	-0.11
.	S_VLDL_C	-0.14	0.08	0.13	0.06	-0.04	.	-0.01	0.01	.	0.02	-0.02	0.02	-0.06
.	XS_VLDL_C	-0.16	0.06	0.12	0.12	0.10	.	-0.04	0.04	.	-0.02	-0.05	0.03	0.00
.	IDL_C	-0.09	0.06	0.08	0.03	-0.03	.	0.00	0.00	.	0.06	0.01	0.00	-0.11
.	L_LDL_C	-0.03	0.03	0.04	0.00	-0.05	.	0.00	0.00	.	0.04	0.01	0.00	-0.06
.	M_LDL_C	-0.03	0.04	0.04	-0.02	-0.11	.	0.02	-0.01	.	0.06	0.02	0.00	-0.09
.	S_LDL_C	-0.03	0.04	0.04	-0.02	-0.08	.	0.02	-0.02	.	0.05	0.01	0.00	-0.08
.	XL_HDL_C	-0.04	-0.02	0.00	0.07	0.19	.	0.07	-0.05	.	-0.04	-0.03	0.01	0.08
.	L_HDL_C	-0.01	-0.02	0.00	0.05	0.12	.	0.06	-0.05	.	-0.07	-0.02	0.02	0.07
.	M_HDL_C	0.00	0.01	0.00	-0.02	-0.05	.	0.02	-0.02	.	0.00	0.02	0.00	-0.02
.	S_HDL_C	0.00	0.02	0.01	-0.04	-0.08	.	0.00	0.00	.	0.04	0.03	-0.01	-0.06
.	XXL_VLDL_FC	-0.10	0.08	0.12	0.02	-0.15	.	-0.10	0.08	.	-0.09	-0.05	0.04	0.05
.	XL_VLDL_FC	-0.12	0.09	0.13	0.01	-0.19	.	-0.02	0.01	.	-0.04	-0.04	0.03	0.00
.	L_VLDL_FC	-0.11	0.10	0.13	0.00	-0.23	.	0.02	-0.02	.	-0.01	-0.02	0.03	-0.04
.	M_VLDL_FC	-0.10	0.08	0.10	0.01	-0.11	.	0.03	-0.02	.	0.05	0.00	0.00	-0.09
.	S_VLDL_FC	-0.11	0.08	0.11	0.03	-0.07	.	0.02	-0.02	.	0.04	0.00	0.01	-0.08
.	XS_VLDL_FC	-0.19	0.07	0.15	0.14	0.11	.	-0.03	0.03	.	-0.05	-0.07	0.04	0.05
.	IDL_FC	-0.10	0.05	0.08	0.05	0.01	.	-0.05	0.04	.	0.04	0.00	0.00	-0.09
.	L_LDL_FC	-0.03	0.02	0.03	0.00	-0.04	.	-0.01	0.00	.	0.04	0.02	0.00	-0.07
.	M_LDL_FC	-0.02	0.03	0.03	-0.02	-0.07	.	0.00	0.00	.	0.06	0.03	0.00	-0.09
.	S_LDL_FC	-0.01	0.01	0.02	-0.02	-0.04	.	0.03	-0.03	.	0.06	0.03	-0.01	-0.08
.	XL_HDL_FC	-0.04	-0.01	0.00	0.07	0.16	.	0.02	-0.01	.	-0.04	-0.03	0.00	0.07
.	L_HDL_FC	-0.04	-0.01	0.01	0.07	0.14	.	0.06	-0.05	.	-0.08	-0.04	0.02	0.09
.	M_HDL_FC	-0.03	0.03	0.03	0.00	-0.03	.	0.04	-0.03	.	-0.01	0.00	0.00	0.00
.	S_HDL_FC	-0.03	0.03	0.04	0.00	-0.06	.	0.02	-0.02	.	0.01	0.00	0.00	-0.02
.	XXL_VLDL_CE	-0.11	0.09	0.13	0.01	-0.19	.	-0.03	0.02	.	-0.08	-0.05	0.04	0.03
.	XL_VLDL_CE	-0.11	0.10	0.13	0.00	-0.20	.	0.03	-0.02	.	0.01	-0.02	0.02	-0.06
.	L_VLDL_CE	-0.12	0.09	0.13	0.02	-0.16	.	0.02	-0.01	.	0.00	-0.02	0.02	-0.05

		Difference from study population mean, SD units											
		Age, years						District of residence			Educational level		
	Biomarker	35-44	45-54	55-64	65-74	75-84		Coyoacan	Iztapalapa		University	High school	Elementary	Other
.	M_VLDL_CE	-0.04	0.02	0.03	0.02	0.00	.	-0.02	0.02	.	0.06	0.02	0.00	-0.09
.	S_VLDL_CE	-0.15	0.08	0.14	0.07	-0.03	.	-0.03	0.03	.	0.00	-0.03	0.02	-0.04
.	XS_VLDL_CE	-0.15	0.06	0.11	0.10	0.09	.	-0.05	0.04	.	0.00	-0.03	0.02	0.00
.	IDL_CE	-0.09	0.06	0.08	0.02	-0.05	.	0.00	0.00	.	0.06	0.02	0.00	-0.12
.	L_LDL_CE	-0.03	0.03	0.04	0.00	-0.05	.	0.00	0.00	.	0.04	0.01	0.00	-0.06
.	M_LDL_CE	-0.04	0.04	0.05	-0.02	-0.11	.	0.02	-0.02	.	0.05	0.02	0.00	-0.09
.	S_LDL_CE	-0.04	0.04	0.05	-0.01	-0.09	.	0.01	-0.01	.	0.04	0.01	0.00	-0.07
.	XL_HDL_CE	-0.02	-0.02	0.00	0.06	0.13	.	0.05	-0.04	.	-0.04	-0.03	0.02	0.06
.	L_HDL_CE	0.00	-0.02	0.00	0.04	0.10	.	0.04	-0.03	.	-0.06	-0.02	0.02	0.06
.	M_HDL_CE	0.00	0.00	0.00	-0.02	-0.05	.	0.00	0.00	.	0.00	0.02	0.00	-0.02
.	S_HDL_CE	0.01	0.01	0.00	-0.04	-0.10	.	0.00	0.00	.	0.04	0.03	-0.01	-0.06
.	XXL_VLDL_TG	-0.05	0.04	0.07	0.00	-0.13	.	-0.08	0.06	.	-0.07	-0.03	0.03	0.04
.	XL_VLDL_TG	-0.09	0.09	0.11	-0.02	-0.26	.	0.04	-0.03	.	-0.03	-0.03	0.03	0.00
.	L_VLDL_TG	-0.04	0.08	0.08	-0.06	-0.28	.	0.13	-0.10	.	0.02	0.00	0.00	-0.05
.	M_VLDL_TG	-0.08	0.09	0.11	-0.03	-0.24	.	0.13	-0.10	.	0.00	-0.01	0.02	-0.05
.	S_VLDL_TG	-0.16	0.09	0.16	0.07	-0.10	.	0.08	-0.07	.	-0.08	-0.06	0.04	0.06
.	XS_VLDL_TG	-0.23	0.09	0.20	0.16	0.06	.	0.02	-0.01	.	-0.13	-0.11	0.06	0.15
.	IDL_TG	-0.25	0.10	0.20	0.17	0.08	.	0.00	0.00	.	-0.13	-0.11	0.06	0.15
.	L_LDL_TG	-0.23	0.10	0.20	0.15	0.03	.	0.00	0.00	.	-0.10	-0.10	0.06	0.11
.	M_LDL_TG	-0.20	0.10	0.19	0.11	-0.03	.	0.01	0.00	.	-0.08	-0.08	0.05	0.07
.	S_LDL_TG	-0.18	0.10	0.17	0.08	-0.10	.	0.02	-0.02	.	-0.08	-0.07	0.05	0.07
.	XL_HDL_TG	-0.21	0.08	0.17	0.16	0.10	.	0.02	-0.01	.	-0.13	-0.10	0.06	0.16
.	L_HDL_TG	-0.21	0.06	0.15	0.18	0.16	.	-0.07	0.05	.	-0.16	-0.10	0.06	0.19
.	M_HDL_TG	-0.19	0.09	0.16	0.12	-0.01	.	-0.08	0.06	.	-0.13	-0.08	0.05	0.13
.	S_HDL_TG	-0.19	0.11	0.18	0.08	-0.11	.	-0.05	0.04	.	-0.10	-0.07	0.05	0.08
.	XXL_VLDL_PL	-0.10	0.07	0.11	0.02	-0.13	.	-0.10	0.08	.	-0.09	-0.05	0.04	0.06
.	XL_VLDL_PL	-0.11	0.10	0.13	0.00	-0.21	.	-0.02	0.01	.	-0.04	-0.04	0.03	0.00
.	L_VLDL_PL	-0.12	0.10	0.14	0.00	-0.23	.	0.04	-0.03	.	-0.02	-0.03	0.03	0.00
.	M_VLDL_PL	-0.11	0.09	0.12	0.00	-0.15	.	0.05	-0.04	.	0.04	0.00	0.01	-0.09
.	S_VLDL_PL	-0.15	0.10	0.14	0.05	-0.10	.	0.04	-0.03	.	0.02	-0.02	0.02	-0.06
.	XS_VLDL_PL	-0.25	0.09	0.19	0.18	0.14	.	-0.02	0.02	.	-0.09	-0.10	0.05	0.12
.	IDL_PL	-0.15	0.08	0.12	0.08	0.02	.	-0.02	0.02	.	0.03	-0.02	0.01	-0.06
.	L_LDL_PL	-0.08	0.07	0.08	-0.01	-0.13	.	-0.01	0.00	.	0.09	0.03	0.00	-0.15
.	M_LDL_PL	-0.07	0.08	0.10	-0.04	-0.19	.	0.03	-0.03	.	0.10	0.03	0.00	-0.16
.	S_LDL_PL	-0.09	0.07	0.09	0.00	-0.10	.	0.09	-0.07	.	0.09	0.01	0.00	-0.13
.	XL_HDL_PL	-0.09	-0.04	0.02	0.17	0.35	.	0.09	-0.08	.	-0.14	-0.10	0.04	0.21
.	L_HDL_PL	-0.06	0.00	0.02	0.09	0.17	.	0.10	-0.08	.	-0.11	-0.04	0.03	0.13
.	M_HDL_PL	-0.06	0.06	0.06	0.00	-0.13	.	0.02	-0.01	.	-0.05	0.00	0.01	0.00
.	S_HDL_PL	-0.05	0.10	0.09	-0.07	-0.31	.	0.00	0.00	.	0.02	0.03	0.00	-0.10
.	XXL_VLDL_L	-0.09	0.07	0.11	0.00	-0.18	.	-0.08	0.06	.	-0.08	-0.05	0.04	0.05
.	XL_VLDL_L	-0.11	0.10	0.13	0.00	-0.25	.	0.02	-0.02	.	-0.03	-0.03	0.03	0.00
.	L_VLDL_L	-0.09	0.10	0.12	-0.03	-0.27	.	0.09	-0.07	.	0.00	-0.02	0.02	-0.04
.	M_VLDL_L	-0.10	0.09	0.12	-0.01	-0.20	.	0.09	-0.07	.	0.04	0.00	0.01	-0.09
.	S_VLDL_L	-0.18	0.11	0.17	0.07	-0.10	.	0.04	-0.03	.	-0.02	-0.05	0.03	0.00
.	XS_VLDL_L	-0.24	0.09	0.19	0.17	0.11	.	-0.03	0.02	.	-0.08	-0.09	0.05	0.09
.	IDL_L	-0.14	0.08	0.12	0.06	-0.02	.	-0.01	0.00	.	0.05	0.00	0.00	-0.08
.	L_LDL_L	-0.09	0.08	0.10	0.00	-0.13	.	0.00	0.00	.	0.08	0.02	0.00	-0.14

		Difference from study population mean, SD units										
		Age, years					District of residence		Educational level			
	Biomarker	35-44	45-54	55-64	65-74	75-84	Coyoacan	Iztapalapa	University	High school	Elementary	Other
.	M_LDL_L	-0.08	0.09	0.10	-0.03	-0.21	0.03	-0.03	0.10	0.03	0.00	-0.16
.	S_LDL_L	-0.09	0.09	0.11	-0.02	-0.17	0.06	-0.05	0.10	0.02	0.00	-0.14
.	XL_HDL_L	-0.11	-0.03	0.02	0.19	0.39	0.12	-0.09	-0.14	-0.10	0.04	0.22
.	L_HDL_L	-0.07	-0.02	0.01	0.12	0.23	0.12	-0.09	-0.13	-0.06	0.03	0.16
.	M_HDL_L	-0.04	0.05	0.05	-0.01	-0.13	0.02	-0.01	-0.04	0.00	0.01	0.00
.	S_HDL_L	-0.05	0.10	0.09	-0.09	-0.34	0.00	0.00	0.05	0.04	0.00	-0.14
.	VLDL_D	0.02	0.07	0.04	-0.12	-0.36	0.06	-0.05	0.02	0.02	0.00	-0.07
.	LDL_D	-0.02	-0.03	-0.01	0.06	0.18	0.00	0.00	-0.04	-0.01	0.00	0.05
.	HDL_D	-0.08	-0.04	0.00	0.17	0.37	0.09	-0.07	-0.17	-0.09	0.04	0.24
.	ApoA1	-0.06	0.06	0.06	0.00	-0.10	0.02	-0.02	-0.03	0.00	0.01	-0.01
.	ApoB	-0.13	0.09	0.13	0.03	-0.10	0.02	-0.01	0.06	0.00	0.00	-0.10
.	ApoB_ApoA1	-0.08	0.05	0.08	0.02	-0.03	0.00	0.00	0.08	0.00	0.00	-0.09
.	LDL_C	-0.03	0.03	0.04	-0.01	-0.07	0.00	0.00	0.05	0.02	0.00	-0.08
.	HDL_C	-0.01	0.02	0.01	0.00	-0.02	0.06	-0.05	-0.01	0.01	0.00	-0.02
.	Serum_TG	-0.14	0.10	0.15	0.03	-0.17	0.04	-0.04	-0.06	-0.05	0.04	0.04

To compare the NMR lipid traits between groups of interest, the table shows the difference between the standardized mean of log-transformed NMR lipid traits of the relevant comparison group and the standardized mean of log-transformed NMR lipid traits of the study population.

Table 4.3. NMR lipid traits by smoking, alcohol intake, fasting duration, and obesity

		Difference (in SD units) from the study population mean log-NMR biomarkers																
		Smoking						Alcohol intake						Fasting duration, hours				BMI
ID	Biomarker	Never	Former	Less than daily	Up to 10 cig. per day	10 cig. per day		Never	Former	Up to 3 times per month	Up to 3 times per week	3 times per week		8	8-12	12		BMI>=30	BMI<30
N		52224	20728	12338	15197	7017	.	20599	6393	25003	49717	5792	.	94177	2496	10831	.	39408	68096
1	XXL_VLDL_P	-0.03	0.02	0.03	0.02	0.05	.	-0.02	0.02	0.03	0.00	0.02	.	0.03	-0.23	-0.25	.	0.19	-0.11
2	XL_VLDL_P	-0.05	0.02	0.05	0.05	0.12	.	-0.05	0.05	0.04	-0.01	0.07	.	0.02	-0.09	-0.13	.	0.16	-0.10
3	L_VLDL_P	-0.06	0.02	0.05	0.06	0.13	.	-0.06	0.05	0.05	-0.01	0.07	.	0.01	-0.06	-0.09	.	0.14	-0.08
4	M_VLDL_P	-0.03	0.00	0.01	0.03	0.08	.	-0.03	0.04	0.01	0.00	0.00	.	-0.02	0.09	0.11	.	0.04	-0.02
5	S_VLDL_P	-0.02	0.02	0.00	0.02	0.08	.	-0.02	0.03	0.02	0.00	0.02	.	0.00	0.03	0.04	.	0.09	-0.05
6	XS_VLDL_P	0.05	0.00	-0.08	-0.08	-0.04	.	0.06	-0.02	-0.04	0.00	-0.07	.	0.00	-0.03	-0.02	.	0.11	-0.07
7	IDL_P	0.04	-0.02	-0.05	-0.05	-0.04	.	0.03	0.01	-0.05	0.02	-0.08	.	-0.02	0.07	0.12	.	0.03	-0.02
8	L_LDL_P	0.00	0.00	-0.02	-0.01	0.02	.	0.00	0.02	-0.02	0.01	-0.04	.	-0.03	0.15	0.21	.	0.00	0.00
9	M_LDL_P	-0.04	0.00	0.03	0.05	0.09	.	-0.05	0.06	0.01	0.00	0.03	.	-0.02	0.13	0.17	.	0.01	0.00
10	S_LDL_P	-0.02	0.00	0.00	0.02	0.07	.	-0.03	0.04	0.00	0.00	0.01	.	-0.02	0.11	0.16	.	0.04	-0.02
11	XL_HDL_P	0.08	0.00	-0.11	-0.10	-0.13	.	0.10	-0.10	-0.04	0.00	-0.10	.	0.03	-0.23	-0.22	.	0.00	0.00
12	L_HDL_P	0.11	-0.04	-0.10	-0.11	-0.25	.	0.11	-0.10	-0.07	0.02	-0.08	.	0.03	-0.21	-0.20	.	-0.02	0.01
13	M_HDL_P	0.08	-0.06	-0.02	-0.07	-0.25	.	0.06	-0.02	-0.08	0.02	0.03	.	0.02	-0.10	-0.11	.	0.02	-0.01
14	S_HDL_P	0.00	-0.04	0.06	0.02	-0.08	.	-0.03	0.06	-0.04	0.02	0.08	.	-0.03	0.16	0.19	.	-0.01	0.00
15	XXL_VLDL_C	-0.02	0.01	0.04	0.01	0.03	.	-0.03	0.02	0.02	0.00	0.02	.	0.02	-0.16	-0.16	.	0.17	-0.10
16	XL_VLDL_C	-0.04	0.02	0.03	0.04	0.09	.	-0.04	0.04	0.03	0.00	0.03	.	0.00	0.00	0.00	.	0.12	-0.07
17	L_VLDL_C	-0.04	0.02	0.03	0.04	0.10	.	-0.04	0.04	0.03	0.00	0.03	.	0.00	0.01	0.00	.	0.10	-0.06
18	M_VLDL_C	0.01	0.00	-0.02	0.00	0.00	.	0.00	0.02	-0.02	0.02	-0.06	.	-0.03	0.14	0.21	.	-0.02	0.01
19	S_VLDL_C	0.00	0.00	-0.02	-0.01	0.02	.	0.00	0.02	0.00	0.00	-0.04	.	-0.02	0.12	0.17	.	0.03	-0.02
20	XS_VLDL_C	0.05	0.00	-0.07	-0.08	-0.08	.	0.05	-0.02	-0.05	0.02	-0.09	.	0.00	0.00	0.03	.	0.05	-0.03
21	IDL_C	0.04	-0.02	-0.04	-0.04	-0.07	.	0.02	0.00	-0.05	0.02	-0.04	.	-0.01	0.05	0.10	.	0.00	0.00
22	L_LDL_C	0.00	-0.01	0.00	0.00	-0.01	.	0.00	0.01	0.00	0.00	0.00	.	0.00	0.06	0.07	.	-0.01	0.00
23	M_LDL_C	-0.01	0.00	0.02	0.02	0.03	.	-0.03	0.03	0.00	0.00	0.01	.	-0.01	0.09	0.10	.	0.00	0.00
24	S_LDL_C	-0.01	0.00	0.02	0.02	0.04	.	-0.03	0.03	0.00	0.00	0.01	.	-0.01	0.09	0.09	.	0.00	0.00
25	XL_HDL_C	0.05	0.00	-0.07	-0.06	-0.08	.	0.05	-0.05	-0.02	0.00	-0.08	.	0.01	-0.07	-0.08	.	-0.02	0.01
26	L_HDL_C	0.07	-0.03	-0.07	-0.08	-0.16	.	0.07	-0.08	-0.04	0.01	-0.07	.	0.01	-0.09	-0.10	.	-0.03	0.02
27	M_HDL_C	0.05	-0.03	-0.01	-0.05	-0.15	.	0.03	-0.01	-0.04	0.01	0.00	.	0.00	0.00	-0.04	.	-0.02	0.00
28	S_HDL_C	0.00	-0.01	0.00	0.00	-0.04	.	-0.02	0.02	-0.01	0.00	0.01	.	0.00	0.07	0.05	.	-0.01	0.00

		Difference (in SD units) from the study population mean log-NMR biomarkers																
		Smoking						Alcohol intake						Fasting duration, hours				BMI
ID	Biomarker	Never	Former	Less than daily	Up to 10 cig. per day	10 cig. per day		Never	Former	Up to 3 times per month	Up to 3 times per week	3 times per week		8	8-12	12		BMI>=30	BMI<30
29	XXL_VLDL_FC	-0.02	0.01	0.03	0.01	0.04	.	-0.02	0.02	0.02	0.00	0.01	.	0.03	-0.19	-0.20	.	0.17	-0.10
30	XL_VLDL_FC	-0.04	0.02	0.04	0.04	0.10	.	-0.04	0.04	0.03	0.00	0.04	.	0.01	-0.06	-0.08	.	0.14	-0.08
31	L_VLDL_FC	-0.05	0.02	0.05	0.06	0.11	.	-0.06	0.05	0.04	0.00	0.06	.	0.00	-0.01	-0.03	.	0.12	-0.07
32	M_VLDL_FC	0.00	0.00	0.00	0.01	0.05	.	-0.02	0.03	0.00	0.00	-0.02	.	-0.02	0.11	0.16	.	0.02	-0.01
33	S_VLDL_FC	0.00	0.00	-0.01	0.00	0.03	.	0.00	0.02	0.00	0.00	-0.03	.	-0.02	0.14	0.19	.	0.01	0.00
34	XS_VLDL_FC	0.04	0.00	-0.07	-0.07	-0.05	.	0.04	-0.02	-0.03	0.00	-0.07	.	0.00	0.01	0.03	.	0.08	-0.05
35	IDL_FC	0.05	-0.02	-0.05	-0.06	-0.08	.	0.04	0.00	-0.05	0.02	-0.07	.	-0.02	0.08	0.14	.	0.00	0.00
36	L_LDL_FC	0.01	-0.01	0.00	0.00	-0.03	.	0.00	0.01	-0.02	0.00	-0.02	.	-0.01	0.07	0.09	.	-0.02	0.01
37	M_LDL_FC	0.00	-0.01	0.01	0.00	0.00	.	-0.01	0.02	0.00	0.00	0.00	.	-0.02	0.10	0.12	.	-0.03	0.02
38	S_LDL_FC	0.00	0.00	0.00	0.01	0.03	.	-0.02	0.02	0.00	0.00	-0.01	.	-0.02	0.11	0.12	.	-0.03	0.02
39	XL_HDL_FC	0.03	0.00	-0.06	-0.04	-0.04	.	0.03	-0.03	-0.01	0.00	-0.07	.	0.01	-0.07	-0.09	.	0.00	0.00
40	L_HDL_FC	0.07	-0.03	-0.07	-0.08	-0.15	.	0.07	-0.08	-0.04	0.00	-0.07	.	0.02	-0.16	-0.18	.	0.00	0.00
41	M_HDL_FC	0.05	-0.03	-0.02	-0.04	-0.14	.	0.03	-0.01	-0.04	0.00	0.01	.	0.01	-0.05	-0.08	.	0.01	0.00
42	S_HDL_FC	0.01	0.00	0.00	0.00	-0.04	.	0.00	0.02	0.00	0.00	0.02	.	0.00	0.03	0.00	.	0.01	0.00
43	XXL_VLDL_CE	-0.02	0.01	0.04	0.00	0.03	.	-0.03	0.02	0.03	0.00	0.02	.	0.02	-0.13	-0.12	.	0.17	-0.10
44	XL_VLDL_CE	-0.03	0.02	0.03	0.03	0.07	.	-0.04	0.04	0.02	0.00	0.01	.	-0.01	0.06	0.08	.	0.09	-0.05
45	L_VLDL_CE	-0.03	0.02	0.02	0.03	0.08	.	-0.03	0.04	0.02	0.00	0.00	.	0.00	0.04	0.05	.	0.08	-0.04
46	M_VLDL_CE	0.03	-0.02	-0.04	-0.03	-0.04	.	0.03	0.00	-0.04	0.02	-0.10	.	-0.02	0.08	0.19	.	-0.05	0.03
47	S_VLDL_CE	0.00	0.01	-0.02	-0.02	0.01	.	0.00	0.01	0.00	0.00	-0.05	.	-0.02	0.11	0.15	.	0.05	-0.03
48	XS_VLDL_CE	0.06	-0.01	-0.07	-0.08	-0.09	.	0.05	-0.02	-0.05	0.02	-0.10	.	0.00	-0.01	0.02	.	0.04	-0.02
49	IDL_CE	0.04	-0.02	-0.03	-0.04	-0.07	.	0.02	0.00	-0.05	0.02	-0.03	.	-0.01	0.04	0.08	.	0.00	0.00
50	L_LDL_CE	0.00	-0.01	0.00	0.00	-0.01	.	0.00	0.01	0.00	0.00	0.00	.	0.00	0.06	0.07	.	-0.01	0.00
51	M_LDL_CE	-0.02	0.00	0.03	0.02	0.04	.	-0.03	0.04	0.00	0.00	0.02	.	-0.01	0.09	0.08	.	0.00	0.00
52	S_LDL_CE	-0.01	0.00	0.02	0.02	0.04	.	-0.03	0.03	0.00	0.00	0.02	.	-0.01	0.08	0.07	.	0.00	0.00
53	XL_HDL_CE	0.04	0.00	-0.06	-0.05	-0.07	.	0.04	-0.06	-0.02	0.00	-0.07	.	0.00	-0.05	-0.06	.	-0.02	0.01
54	L_HDL_CE	0.06	-0.03	-0.06	-0.07	-0.15	.	0.06	-0.08	-0.04	0.01	-0.06	.	0.00	-0.06	-0.07	.	-0.03	0.02
55	M_HDL_CE	0.04	-0.02	-0.02	-0.03	-0.12	.	0.02	0.00	-0.03	0.01	0.00	.	0.00	0.00	-0.03	.	-0.02	0.00
56	S_HDL_CE	0.00	0.00	0.00	0.00	-0.04	.	-0.01	0.01	-0.01	0.00	0.01	.	0.00	0.08	0.06	.	-0.02	0.01
57	XXL_VLDL_TG	-0.02	0.00	0.03	0.02	0.02	.	-0.02	0.00	0.02	0.00	0.00	.	0.03	-0.24	-0.23	.	0.15	-0.08
58	XL_VLDL_TG	-0.06	0.02	0.06	0.07	0.11	.	-0.06	0.04	0.04	-0.01	0.09	.	0.02	-0.12	-0.18	.	0.16	-0.09

		Difference (in SD units) from the study population mean log-NMR biomarkers																
		Smoking						Alcohol intake						Fasting duration, hours				BMI
ID	Biomarker	Never	Former	Less than daily	Up to 10 cig. per day	10 cig. per day		Never	Former	Up to 3 times per month	Up to 3 times per week	3 times per week		8	8-12	12		BMI>=30	BMI<30
59	L_VLDL_TG	-0.07	0.02	0.07	0.09	0.14	.	-0.08	0.05	0.05	-0.01	0.10	.	0.01	-0.06	-0.11	.	0.11	-0.06
60	M_VLDL_TG	-0.06	0.02	0.06	0.08	0.13	.	-0.07	0.05	0.05	-0.01	0.10	.	0.00	-0.03	-0.06	.	0.09	-0.05
61	S_VLDL_TG	-0.04	0.02	0.02	0.04	0.09	.	-0.03	0.03	0.04	-0.02	0.07	.	0.02	-0.11	-0.15	.	0.13	-0.08
62	XS_VLDL_TG	0.00	0.02	-0.03	-0.03	0.02	.	0.01	0.00	0.02	-0.02	0.03	.	0.02	-0.12	-0.17	.	0.16	-0.09
63	IDL_TG	0.02	0.02	-0.04	-0.04	0.00	.	0.02	-0.01	0.00	-0.01	0.02	.	0.02	-0.13	-0.17	.	0.17	-0.10
64	L_LDL_TG	0.00	0.02	-0.03	-0.03	0.01	.	0.00	0.00	0.01	-0.01	0.04	.	0.02	-0.08	-0.12	.	0.16	-0.09
65	M_LDL_TG	-0.02	0.03	0.00	0.00	0.05	.	-0.02	0.01	0.03	-0.02	0.07	.	0.01	-0.07	-0.10	.	0.15	-0.09
66	S_LDL_TG	-0.03	0.03	0.02	0.02	0.08	.	-0.03	0.02	0.04	-0.02	0.08	.	0.02	-0.13	-0.17	.	0.16	-0.09
67	XL_HDL_TG	0.01	0.01	-0.04	-0.03	-0.02	.	0.03	-0.02	0.01	-0.02	0.02	.	0.04	-0.27	-0.30	.	0.14	-0.08
68	L_HDL_TG	0.06	0.00	-0.06	-0.08	-0.15	.	0.07	-0.04	-0.04	0.00	-0.03	.	0.06	-0.36	-0.40	.	0.14	-0.08
69	M_HDL_TG	0.04	0.00	-0.02	-0.05	-0.10	.	0.04	0.00	-0.02	0.00	0.01	.	0.05	-0.28	-0.35	.	0.18	-0.10
70	S_HDL_TG	-0.01	0.01	0.01	0.00	0.03	.	-0.01	0.03	0.02	-0.01	0.06	.	0.03	-0.15	-0.21	.	0.18	-0.11
71	XXL_VLDL_PL	-0.02	0.01	0.03	0.00	0.04	.	-0.02	0.02	0.02	0.00	0.00	.	0.03	-0.23	-0.23	.	0.17	-0.10
72	XL_VLDL_PL	-0.04	0.02	0.04	0.04	0.11	.	-0.04	0.04	0.03	0.00	0.04	.	0.01	-0.08	-0.11	.	0.16	-0.09
73	L_VLDL_PL	-0.05	0.02	0.05	0.06	0.13	.	-0.05	0.05	0.04	-0.01	0.06	.	0.01	-0.05	-0.08	.	0.14	-0.08
74	M_VLDL_PL	-0.02	0.00	0.00	0.02	0.06	.	-0.02	0.04	0.00	0.00	0.00	.	-0.02	0.09	0.12	.	0.04	-0.02
75	S_VLDL_PL	-0.01	0.01	0.00	0.00	0.06	.	-0.02	0.03	0.00	0.00	0.00	.	-0.02	0.10	0.14	.	0.05	-0.03
76	XS_VLDL_PL	0.05	0.00	-0.09	-0.09	-0.03	.	0.06	-0.03	-0.03	0.00	-0.08	.	0.00	-0.05	-0.04	.	0.12	-0.07
77	IDL_PL	0.05	-0.02	-0.06	-0.07	-0.07	.	0.04	0.00	-0.06	0.02	-0.07	.	-0.01	0.04	0.09	.	0.02	-0.01
78	L_LDL_PL	0.01	-0.01	0.00	0.00	-0.02	.	0.00	0.03	-0.03	0.02	-0.01	.	-0.03	0.13	0.19	.	0.00	0.00
79	M_LDL_PL	-0.02	0.00	0.02	0.03	0.05	.	-0.04	0.05	0.00	0.00	0.03	.	-0.03	0.15	0.19	.	0.00	0.00
80	S_LDL_PL	-0.04	0.01	0.01	0.04	0.12	.	-0.05	0.04	0.02	0.00	0.03	.	-0.02	0.15	0.17	.	0.00	0.00
81	XL_HDL_PL	0.09	-0.01	-0.13	-0.12	-0.14	.	0.11	-0.12	-0.05	0.01	-0.18	.	0.04	-0.28	-0.29	.	0.00	0.00
82	L_HDL_PL	0.10	-0.04	-0.08	-0.10	-0.25	.	0.10	-0.09	-0.07	0.01	-0.05	.	0.04	-0.25	-0.29	.	0.02	0.00
83	M_HDL_PL	0.08	-0.06	-0.01	-0.07	-0.24	.	0.05	-0.02	-0.08	0.01	0.05	.	0.03	-0.18	-0.21	.	0.06	-0.04
84	S_HDL_PL	0.03	-0.04	0.04	-0.02	-0.14	.	0.00	0.04	-0.06	0.01	0.10	.	0.00	-0.01	-0.04	.	0.07	-0.04
85	XXL_VLDL_L	-0.03	0.01	0.04	0.02	0.05	.	-0.03	0.02	0.03	0.00	0.02	.	0.03	-0.22	-0.23	.	0.18	-0.11
86	XL_VLDL_L	-0.05	0.02	0.05	0.06	0.12	.	-0.06	0.05	0.04	-0.01	0.08	.	0.02	-0.09	-0.13	.	0.16	-0.09
87	L_VLDL_L	-0.06	0.02	0.06	0.08	0.14	.	-0.07	0.06	0.05	-0.01	0.09	.	0.00	-0.04	-0.08	.	0.12	-0.07
88	M_VLDL_L	-0.04	0.01	0.03	0.05	0.09	.	-0.05	0.04	0.02	0.00	0.04	.	0.00	0.06	0.07	.	0.06	-0.03

		Difference (in SD units) from the study population mean log-NMR biomarkers																	
		Smoking						Alcohol intake						Fasting duration, hours				BMI	
ID	Biomarker	Never	Former	Less than daily	Up to 10 cig. per day	10 cig. per day		Never	Former	Up to 3 times per month	Up to 3 times per week	3 times per week		8	8-12	12		BMI>=30	BMI<30
89	S_VLDL_L	-0.02	0.02	0.00	0.01	0.07	.	-0.02	0.03	0.02	0.00	0.02	.	0.00	0.03	0.04	.	0.09	-0.05
90	XS_VLDL_L	0.05	0.00	-0.08	-0.08	-0.05	.	0.06	-0.02	-0.04	0.00	-0.07	.	0.00	-0.05	-0.04	.	0.12	-0.07
91	IDL_L	0.05	-0.02	-0.05	-0.06	-0.07	.	0.04	0.00	-0.06	0.02	-0.05	.	-0.01	0.04	0.09	.	0.02	-0.01
92	L_LDL_L	0.01	-0.02	0.00	0.00	-0.02	.	0.00	0.03	-0.03	0.01	0.00	.	-0.02	0.11	0.17	.	0.00	0.00
93	M_LDL_L	-0.03	0.00	0.03	0.04	0.06	.	-0.05	0.06	0.00	0.00	0.04	.	-0.02	0.15	0.18	.	0.01	0.00
94	S_LDL_L	-0.03	0.00	0.02	0.04	0.10	.	-0.05	0.05	0.01	0.00	0.04	.	-0.02	0.15	0.18	.	0.02	0.00
95	XL_HDL_L	0.09	-0.02	-0.14	-0.12	-0.15	.	0.12	-0.13	-0.05	0.00	-0.15	.	0.04	-0.27	-0.27	.	0.00	0.00
96	L_HDL_L	0.13	-0.06	-0.11	-0.14	-0.31	.	0.13	-0.12	-0.09	0.02	-0.08	.	0.04	-0.29	-0.29	.	0.00	0.00
97	M_HDL_L	0.08	-0.06	-0.01	-0.08	-0.26	.	0.06	-0.02	-0.08	0.02	0.05	.	0.02	-0.14	-0.16	.	0.04	-0.02
98	S_HDL_L	0.02	-0.04	0.05	0.00	-0.11	.	-0.01	0.05	-0.05	0.01	0.10	.	0.00	0.05	0.04	.	0.05	-0.03
99	VLDL_D	-0.09	0.02	0.11	0.12	0.18	.	-0.10	0.07	0.07	-0.02	0.14	.	0.01	-0.05	-0.09	.	0.10	-0.06
100	LDL_D	0.08	-0.02	-0.09	-0.10	-0.14	.	0.09	-0.05	-0.07	0.02	-0.15	.	-0.02	0.10	0.16	.	-0.08	0.04
101	HDL_D	0.12	-0.04	-0.12	-0.14	-0.27	.	0.14	-0.13	-0.07	0.00	-0.07	.	0.05	-0.31	-0.33	.	0.00	0.00
102	ApoA1	0.09	-0.06	-0.02	-0.08	-0.25	.	0.06	-0.03	-0.08	0.01	0.04	.	0.01	-0.09	-0.10	.	0.02	0.00
103	ApoB	0.00	0.00	-0.01	0.00	0.04	.	-0.01	0.03	-0.01	0.00	-0.02	.	-0.02	0.12	0.17	.	0.02	-0.01
104	ApoB_ApoA1	-0.06	0.04	0.00	0.05	0.21	.	-0.05	0.05	0.04	0.00	-0.05	.	-0.03	0.18	0.23	.	0.01	0.00
105	LDL_C	0.00	-0.01	0.01	0.00	0.00	.	-0.02	0.02	0.00	0.00	0.00	.	-0.01	0.08	0.09	.	-0.01	0.00
106	HDL_C	0.09	-0.05	-0.04	-0.08	-0.25	.	0.07	-0.04	-0.08	0.02	0.00	.	0.00	-0.03	-0.03	.	-0.05	0.03
107	Serum_TG	-0.04	0.02	0.04	0.04	0.10	.	-0.04	0.04	0.04	-0.02	0.09	.	0.02	-0.13	-0.18	.	0.16	-0.09

Reference

1 Joshi, R. *et al.* Triglyceride-containing lipoprotein sub-fractions and risk of coronary heart disease and stroke: A prospective analysis in 11,560 adults. *European journal of preventive cardiology* **27**, 1617-1626, doi:10.1177/2047487319899621 (2020).

2 Holmes, M. V. *et al.* Lipids, Lipoproteins, and Metabolites and Risk of Myocardial Infarction and Stroke. *Journal of the American College of Cardiology* **71**, 620-632, doi:10.1016/j.jacc.2017.12.006 (2018).

3 Aday, A. W. *et al.* Lipoprotein Particle Profiles, Standard Lipids, and Peripheral Artery Disease Incidence. *Circulation* **138**, 2330-2341, doi:10.1161/circulationaha.118.035432 (2018).

4 Johansen, M. *et al.* Per-Particle Triglyceride-Rich Lipoproteins Imply Higher Myocardial Infarction Risk Than Low-Density Lipoproteins: Copenhagen General Population Study. *Arteriosclerosis, thrombosis, and vascular biology* **41**, 2063-2075, doi:10.1161/atvbaha.120.315639 (2021).

5 Mach, F. *et al.* 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk. *European heart journal* **41**, 111-188, doi:10.1093/eurheartj/ehz455 (2020).

6 Palmisano, B. T., Zhu, L., Eckel, R. H. & Stafford, J. M. Sex differences in lipid and lipoprotein metabolism. *Molecular metabolism* **15**, 45-55, doi:10.1016/j.molmet.2018.05.008 (2018).

7 Slade, E. *et al.* Age and sex are associated with the plasma lipidome: findings from the GOLDN study. *Lipids in health and disease* **20**, 30, doi:10.1186/s12944-021-01456-2 (2021).

8 Park, Y. M. *et al.* The effect of cardiorespiratory fitness on age-related lipids and lipoproteins. *Journal of the American College of Cardiology* **65**, 2091-2100, doi:10.1016/j.jacc.2015.03.517 (2015).

9 Schubert, C. M. *et al.* Lipids, lipoproteins, lifestyle, adiposity and fat-free mass during middle age: the Fels Longitudinal Study. *International journal of obesity (2005)* **30**, 251-260, doi:10.1038/sj.ijo.0803129 (2006).

10 Feng, L. *et al.* Age-related trends in lipid levels: a large-scale cross-sectional study of the general Chinese population. *BMJ open* **10**, e034226, doi:10.1136/bmjopen-2019-034226 (2020).

11 Craig, W. Y., Palomaki, G. E. & Haddow, J. E. Cigarette smoking and serum lipid and lipoprotein concentrations: an analysis of published data. *BMJ (Clinical research ed.)* **298**, 784-788, doi:10.1136/bmj.298.6676.784 (1989).

12 Brien, S. E., Ronksley, P. E., Turner, B. J., Mukamal, K. J. & Ghali, W. A.

Effect of alcohol consumption on biological markers associated with risk of coronary heart disease: systematic review and meta-analysis of interventional studies. *BMJ (Clinical research ed.)* **342**, d636, doi:10.1136/bmj.d636 (2011).

13 Millwood, I. Y. *et al.* Conventional and genetic evidence on alcohol and vascular disease aetiology: a prospective study of 500 000 men and women in China. *Lancet (London, England)* **393**, 1831-1842, doi:10.1016/s0140-6736(18)31772-0 (2019).

14 Bays, H. E. *et al.* Obesity, adiposity, and dyslipidemia: a consensus statement from the National Lipid Association. *Journal of clinical lipidology* **7**, 304-383, doi:10.1016/j.jacl.2013.04.001 (2013).

15 Borén, J. *et al.* Low-density lipoproteins cause atherosclerotic cardiovascular disease: pathophysiological, genetic, and therapeutic insights: a consensus statement from the European Atherosclerosis Society Consensus Panel. *European heart journal* **41**, 2313-2330, doi:10.1093/eurheartj/ehz962 (2020).

16 Ference, B. A. *et al.* Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. *European heart journal* **38**, 2459-2472, doi:10.1093/eurheartj/ehx144 (2017).

17 Richardson, T. G. *et al.* Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: A multivariable Mendelian randomisation analysis. *PLoS medicine* **17**, e1003062, doi:10.1371/journal.pmed.1003062 (2020).

18 Würtz, P. *et al.* Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Large-Scale Epidemiology: A Primer on -Omic Technologies. *Am J Epidemiol* **186**, 1084-1096, doi:10.1093/aje/kwx016 (2017).

19 Julkunen, H. *et al.* Atlas of plasma NMR biomarkers for health and disease in 118,461 individuals from the UK Biobank. *Nature communications* **14**, 604, doi:10.1038/s41467-023-36231-7 (2023).

20 Bell, J. A. *et al.* Sex differences in systemic metabolites at four life stages: cohort study with repeated metabolomics. *BMC Medicine* **19**, 58, doi:10.1186/s12916-021-01929-2 (2021).

21 Shohaimi, S., Boekholdt, M. S., Luben, R., Wareham, N. J. & Khaw, K.-T. Distribution of lipid parameters according to different socio-economic indicators- the EPIC-Norfolk prospective population study. *BMC Public Health* **14**, 782, doi:10.1186/1471-2458-14-782 (2014).

22 Würtz, P. *et al.* Metabolic signatures of adiposity in young adults: Mendelian randomization analysis and effects of weight change. *PLoS*

medicine **11**, e1001765, doi:10.1371/journal.pmed.1001765 (2014).

23 Bell, J. A. *et al.* Associations of Body Mass and Fat Indexes With Cardiometabolic Traits. *Journal of the American College of Cardiology* **72**, 3142-3154, doi:10.1016/j.jacc.2018.09.066 (2018).

24 Ambrose, J. A. & Barua, R. S. The pathophysiology of cigarette smoking and cardiovascular disease: an update. *Journal of the American College of Cardiology* **43**, 1731-1737, doi:10.1016/j.jacc.2003.12.047 (2004).

25 Momayyezi, M., Jambarsang, S., Fallahzadeh, H. & Sefidkar, R. Association between lipid profiles and cigarette smoke among adults in the Persian cohort (Shahedieh) study. *BMC Public Health* **24**, 1256, doi:10.1186/s12889-024-18734-0 (2024).

26 Dare, S., Mackay, D. F. & Pell, J. P. Relationship between smoking and obesity: a cross-sectional study of 499,504 middle-aged adults in the UK general population. *PLoS One* **10**, e0123579, doi:10.1371/journal.pone.0123579 (2015).

27 Holmes, M. V. *et al.* Association between alcohol and cardiovascular disease: Mendelian randomisation analysis based on individual participant data. *BMJ (Clinical research ed.)* **349**, g4164, doi:10.1136/bmj.g4164 (2014).

28 Åberg, F., Byrne, C. D., Pirola, C. J., Männistö, V. & Sookoian, S. Alcohol consumption and metabolic syndrome: Clinical and epidemiological impact on liver disease. *J Hepatol* **78**, 191-206, doi:10.1016/j.jhep.2022.08.030 (2023).

29 Sidhu, D. & Naugler, C. Fasting time and lipid levels in a community-based population: a cross-sectional study. *Arch Intern Med* **172**, 1707-1710, doi:10.1001/archinternmed.2012.3708 (2012).

30 Li-Gao, R. *et al.* Assessment of reproducibility and biological variability of fasting and postprandial plasma metabolite concentrations using 1H NMR spectroscopy. *PLoS One* **14**, e0218549, doi:10.1371/journal.pone.0218549 (2019).

CHAPTER 5: Genetics of major NMR lipid traits in MCPS: Genome-wide association studies

Chapter summary

In this chapter GWAS were conducted for five major lipid-related traits measured by NMR (LDL-C, HDL-C, TG, ApoB, and ApoA1) using 38,827,546 well-imputed variants in the MCPS study population (n=122,543 after QC and after excluding participants known to be on lipid-lowering drugs at recruitment). The additive effect of each SNP adjusted for age, sex, genetic PCs, batch, and fasting duration was estimated. For LDL-C, 135 conditionally-independent signals (including 36 potentially novel signals) at 92 loci were identified (e.g., *APOE*, *CELSR2*, *PCSK9*, *APOB*, *HMGCR*, *LPA*, *LPL*, *LDLR*). There were 202 independent signals (70 novel) at 135 loci for HDL-C (e.g., *CETP*, *SIK3*, *LIPC*, *LPL*, *APOA5*), 118 independent signals (36 potentially novel) at 75 loci for TG (e.g., *ZPR1*, *GCKR*, *APOC1P1*, *APOC3*, *LPA*, *DOCK6*), 129 independent signals (51 potentially novel) at 92 loci for ApoA1 (e.g., *SIK3*, *CETP*, *ANGPTL8*, *LIPC*, *ABCA1*), and 160 independent signals (67 potentially novel) at 113 loci for ApoB (e.g., *CELSR2*, *APOE*, *APOB*, *LDLR*, *PCSK9*, *CETP*, *LPA*). The intercepts estimated from covariate-adjusted LDSC ranged from 1.12 to 1.13, suggesting good control of genomic inflation. The covariate-adjusted LDSC demonstrated modest heritability for these major lipid traits (h^2 ranging from 0.11-0.17) in this Mexican population. The potentially novel loci warrant further replication in larger Latin American-ancestry samples. Further downstream analyses for the function and enrichment of the identified loci could provide additional information on lipoprotein biology and elucidate their potential causal roles in ASCVD.

5.1 Background

Increased plasma concentrations of lipids and lipoproteins are heritable, modifiable risk factors for ASCVD,[1,2] with disproportionate prevalence of non-optimal cholesterol patterns among populations underrepresented in genetics research, especially Latin American populations.[3] Despite advances in prevention and treatment, particularly by lowering LDL-C levels, ASCVD remains the leading cause of morbidity and premature death worldwide. The total number of deaths due to ASCVD has increased steadily since 1990 and the mortality rate is increasing in many low- and middle-income countries.[4] Genome-wide association studies (GWAS) of plasma lipids can improve the understanding of the genetic factors involved in lipid metabolism and identify potential therapeutic targets for cholesterol management and prevention of ASCVD.[5-7] However, most previous GWAS have been conducted in individuals of European ancestry,[8-13] with fewer (and smaller) studies including other ancestry groups.[6,14,15] Many genetic variants that contribute to the variation of plasma lipid levels in non-European ancestries may not yet have been identified, because of differences between ancestries in allele frequencies, effect sizes and linkage-disequilibrium (LD) patterns. The current largest multi-ancestry GWAS meta-analysis of 201 studies including 1.65 million individuals (350,000 of non-European ancestries) with lipid measures from traditional panels identified 923 lipid-associated genomic regions containing 1,750 distinct *index* variants at genome-wide significance ($P < 5 \times 10^{-8}$).[14] Although most of the genetic variants were shared among the European, African, East Asian, Hispanic, and South Asian ancestry groups, some novel coding variants were associated with lipid levels only in non-European populations.[14,16] Previous studies have also shown that increasing diversity of ancestries recruited in GWAS can improve fine-mapping of highly correlated variants and narrow down the regional windows of associations due to difference in LD structure among ancestries.[17] Therefore, studies in diverse ancestry groups are important to improve the understanding of the differences in lipid profiles across populations, the identification of

ancestry-specific variants associated with lipid traits, and finally, for fine-mapping identified loci to prioritize the causal variants or genes. However, gain in the number of discovered loci and ancestry-specific variants are still modest compared with the substantial improvements in fine-mapping, which reflect the limited power to detect an association within minority groups of small sample sizes in the previous trans-ethnic GWAS meta-analysis. Future studies based on single large populations of non-European ancestries would enhance discovery power and enable novel insights.

Notably, previous GWAS of lipid traits have also mainly focused on traditional lipid panels (LDL-C, HDL-C, TG, TC, non-HDL-C) rather than NMR lipid profile, failing to delineate the genetic determinants of functionally diverse lipoprotein subfractions and their contribution to the risk of cardiometabolic diseases. By contrast, NMR spectroscopy measures both the number and size of lipoprotein particles and their lipid contents, which independently and specifically affect manifestations of ASCVD, potentially improving risk assessment over traditional lipids (**Chapter 2**). Several GWAS of NMR-measured circulating metabolites including blood lipid profiles have been done,[7,18,19] but with sample sizes of only several thousand, which is still modest compared with studies by large disease consortium (e.g. GLGC) and results in a limit number of significant variants associated with NMR lipid profile.[19] One GWAS of NMR lipids also indicated low genetic correlations between traditional lipid panels and plasma lipid particles.[7] Estimates of variance explained by detected lipid loci ranged from 11%-23% among traditional lipids, which are substantially below family-based heritability estimates (33%-51%), suggesting that many more lipid-associated variants remain undiscovered.[15] Overall, these limitations support the need for further research into the genetic basis of the functionally and chemically diverse lipid particles in diverse populations to further provide biological insights into lipid metabolism, and inform potential therapeutic targets for cardiovascular and metabolic diseases.

As the current largest single Latin American cohort of about 120,000 individuals with

available genetic data and measures for NMR metabolites, MCPS provides a unique opportunity to overcome the challenge of small sample size and fill the current gaps in previous lipid GWAS, facilitate the identification of genetic variants influencing lipoprotein subfractions, and enhance the current understanding of genetic regulation of blood lipids. Of note, due to the historical patterns of migration, colonization, and post-colonial population expansion, the populations of modern Latin Americans exhibit complex patterns of admixture (when individuals from genetically distinct populations mate) between different continental ancestry groups including Indigenous Amerindian, European and African, leading to complex patterns of population structure. Thus, appropriate analysis methods are needed to correct for the complex admixture and extensive relatedness particularly in the Latin American population to avoid spurious genetic associations and biased test statistics due to population stratification.[20,21] The current chapter aims to test the association of more than 10 million genotyped or imputed genetic variants with five major lipids (LDL-C, HDL-C, TG) and apolipoproteins (ApoB, ApoA1). The objectives of this chapter are to: a) conduct a GWAS to identify loci and signals associated with these major lipid traits after accounting for population structure, admixture, and family relatedness; and b) Assess whether the identified loci are novel and further identify conditionally independent variants.

5.2 Methods

5.2.1 Genotyping

The overall study design and recruitment of participants is detailed in **Chapter 3**. Of the 159,755 participants recruited into MCPS, a blood sample was successfully taken, processed and stored for 155,453 (97.3%). Of these, DNA was extracted for 146,068 (94.0%) and sent for genotyping and exome sequencing, of which 145,266 (99.5%) were successfully processed. DNA was extracted at the UK Biocenter with a 260:280nm ratio of ≥ 1.8 and a 260:230nm ratio of 2.0-2.2 on UV-VIS spectroscopy was deemed uncontaminated. DNA yield was also checked with absorbance at 260nm. Pure DNA was suspended in TE buffer to generate samples containing >1.5μg at concentrations of 10-200ng/μL (with 98% of samples meeting the more stringent criteria of >2μg at concentrations of 20-200ng/μL). Genomic DNA samples were transferred to the Regeneron Genetics Center (RGC), where about 650,000 variants were genotyped using the second-generation Illumina Global Screening Array beadchip. Chip variants provide whole-genome coverage of cardiometabolic trait-influencing variants and facilitated high quality imputation of variants with MAF>1% in the 1000 Genomes Project (which included Latin American ancestry individuals from Los Angeles, Medellin, Lima and Puerto Rica, and Spanish ancestry individuals from Iberia). The average genotype call rate per sample was 98.4% and 98.4% of samples had a call rate above 90%. Over 9,430,000 autosomal variants were imputed using the diverse TOPMed reference panel (within the TOPMED imputation server) including 47,159 European, 24,267 African, and 17,085 admixed American genomes, with imputation info score >0.3.

5.2.2 Genetic quality control and exclusions

Multiple quality control (QC) steps on individuals and SNPs were performed prior to inclusion in the present study. Samples failed initial QC steps if they met one or more of the following criteria: disagreement between genetically-determined and reported sex

(n=1,827), low quality samples with call rates below 90% (n=2,276), genotyping chip variants discordant with exome data (n=44), genetically-identified duplicated sample (n=1,063), samples flagged by MCPS for exclusion (n=1,122), and other reasons (n=223). In total, 4,435 unique samples were excluded and data on 650,380 variants in 140,831 individuals remained. Following additional QC steps established in conjunction with the initial QC by the RGC, data on 580,437 variants in the remaining 138,511 individuals were used for downstream analysis. Additional QC steps implemented in PLINK 1.9 included: checks for consistency of genotypes in sex chromosomes (steps 1-4), sample- and variant-level missingness filters (step 5-6), the Hardy-Weinberg equilibrium exact test applied to a set of 81,747 third degree unrelated samples identified from the initial relatedness analysis by PLINK and PRIMUS (step 7), setting genotypes with Mendel errors in nuclear families to missing (step 8), and the second round of steps 5-7 (step 9).

In addition to the participant exclusions applied in the above QC steps, participants were further excluded if they were recorded at recruitment as taking a lipid-lowering medication (0.6% of participants) or if they had, at the time of analysis, missing values of NMR lipid measures. After applying the above QC and sample selection criteria, a total of 122,543 MCPS participants were included in the GWAS analyses.

5.2.3 Phenotype definitions

As detailed in **Chapter 3**, an array of 104 measurements related to lipid traits were quantified by the Nightingale Health Ltd NMR metabolomics platform. Of these, the non-fasting concentration of derived major lipoprotein-related traits including LDL-C, HDL-C, TG, ApoB, and ApoA1 were used as the main exposures. TG values were natural-log transformed. For each phenotype, residuals were obtained after accounting for age, sex, and the first seven genotype-based PCs (see below) as covariates, and inverse-normal-transform residuals were created for GWAS analyses.

5.2.4 Genotype-based principal components

To control for population stratification (the presence of multiple genetically distinct subpopulations that differ in their mean phenotypic values), principal components (PCs) were computed to define clusters of individuals with similar genotypes, and were then included as covariates in the subsequent GWAS regression models.[22] Variants with MAF >1% were clumped at LD r^2 threshold of 0.01 using bigsnpr. Singular value decomposition generated 20 PCs in 58,239 participants unrelated at third degree or closer identified by KING, and were then projected in the remaining participants. The first seven PCs were used in the subsequent analyses to limit the effect of population stratification.

5.2.5 Primary association analyses

As MCPS recruited members of households in the two neighboring geographical districts of the Mexico City, extensive family relatedness was observed in this admixed Latin American population. Excluding participants related at third degree or closer (kinship co-efficient <0.0442) using KING would reduce the sample size from ~150,000 to ~58,000. Thus, rather than restricting the population to unrelated participants, robust statistical methods accounting for relatedness were used to improve statistical power for genomic discovery and control type 1 error. REGENIE is a two-step whole-genome regression approach that accounts for potential population stratification and family relatedness, providing better statistical power than restricting the association analysis to unrelated individuals.[23] The first step contains two stages. The trait of interest is first ridge-regressed on variants in non-overlapping blocks to reduce dimensionality. The trait is then regressed again on the lambda coefficients from the initial ridge regression before cross-validation. Leave one chromosome out (LOCO) estimates (the variance of adjusted phenotypes explained by the genetic variants in all-but-one chromosome) that avoid proximal contamination are computed by combining blocks into sections of the

genome corresponding to chromosomes. In the second step, each chromosome's residual variance is regressed on the larger set of genetic markers (e.g., imputed genotypes) in it. Step one used variants with MAF ≥1% and 200 variants per block and step two used variants with MAF ≥0.1% and 200 variants per block. Both steps adjusted for age, sex, and the first seven genetic PCs. Single-variant associations with $P < 5 \times 10^{-8}$ were considered significant.

The primary output of GWAS analyses include a list of P values, effect sizes and their directions generated from the association tests of included genetic variants with each major NMR lipid trait (under the additive model). The GWAS results were first visualized using Manhattan and quantile-quantile plots, generated using R software (v 4.2). Manhattan plots show the significance of each genetic variant's association with major lipid traits and their genomic positions. Quantile-quantile plots show the distribution of expected P values under the null model of no significance verses the observed P values. Each variant identified by GWAS was mapped to the nearest gene (herein referred to as genetic loci) based on its genomic position using the reference genome. This approach ensured that all significant variants were systematically linked to their nearest genes, facilitating biological interpretation of the genetic associations.

To identify independent signals of association at each genetic locus identified by GWAS, stepwise conditional association analysis was conducted using GCTA-COJO software.[24] This method uses the summary statistics from the GWAS results and LD correlations between SNPs estimated from a reference sample (~79K participants unrelated to the third degree) to perform a conditional association analysis. It implements the association analysis conditioning on the selected SNP(s) that show the strongest evidence of association in each locus to search for the top SNPs iteratively via a stepwise model selection procedure until no SNP have a conditional P value that passes the genome-wide significance level of $P < 5 \times 10^{-8}$.

Variants were classified as known or novel through manual cross-referencing with entries in the NHGRI-EBI GWAS Catalog for the five major lipid traits studied (details on studies included and data used for each trait were shown in: EFO_0004612 for HDL-C; EFO_0004611 for LDL-C; EFO_0004530 for TG; EFO_0004615 for ApoB; EFO_0004614 for ApoA1). In brief, the significant summary statistics reported in the database were used for comparison.

5.2.6 Genomic inflation factor and Covariate-adjusted LD score regression

The genomic inflation factor (λ_{GC}) was computed as the ratio of the median observed χ^2 statistics to the median expected χ^2 value (p=0.5) for a χ^2 distribution with one degree of freedom. This established metric assesses genomic inflation but does not differentiate between various sources of inflation: population structure, linkage disequilibrium (LD) between causal and non-causal variants, and true polygenicity in the genetic architecture of complex traits. Given that all three factors may contribute to inflation in GWAS of blood lipid traits within a population with complex admixture such as MCPS, we further utilized covariate-adjusted linkage disequilibrium score regression (Cov-LDSC) to evaluate inflation attributable to residual population structure.

Cov-LDSC is an extension of LDSC[25] designed for populations with complex admixture, such as those of African American and Latin American ancestries. It adjusts genotype counts for population structure when constructing LD scores, thus extending LD score regression to admixed and heterogeneous populations. Given that admixture proportions can vary between samples, Cov-LD scores were derived from 9,764 whole-genome sequenced MCPS participants unrelated at the third degree or closer (at a KING kinship threshold of 0.0442).

5.2.7 Power calculation

When the GWAS sample includes related individuals (as in the case of MCPS) the effective sample size is lower than the number of genotyped individuals and statistical

power will typically be lower, depending on the heritability of the phenotypes and on the relatedness of individuals (there will be a greater loss in power for more heritable traits and more closely related individuals). Power was calculated based on the significance level (0.05/the number of independent markers examined) and the variance explained by additive effects at the marker of interest under a linear model. If the GWAS analyses were restricted to the ~58,000 unrelated individuals, the study would have 99% and 82% power to detect variants that explain 0.1% and 0.07% phenotypic variation, respectively, in lipid traits at a p-value threshold of 5×10^{-8}. However, power is higher when using robust methods to account for relatedness in the full cohort (as in the analyses described above and presented below).

5.3 Results

5.3.1 Associations in GWAS of LDL-C

GWAS for LDL-C identified 1,269 variants with $p<5x10^{-8}$ at 447 loci (**Figure 5.1)**. Step-wise approximate conditional analysis identified 135 conditionally-independent signals at 92 loci (**Table 5.1**). Of these, 31 were previously reported, and 68 were ±500kb from variants that have been previously reported as associated with LDL-C, with 36 potentially novel variants (i.e., not located in any previously reported locus for the trait), according to the published results on LDL-C measurements in NHGRI-EBI GWAS Catalog. The top ten significant signals were chr19:44908822:C:T at the *APOE* locus, chr1:109274968:G:T at *CELSR2*, chr8:9323885:A:G at *PPP1R3B*, chr9:104858554:G:A at *ABCA1*, chr19:11086585:G:A at *LDLR*, chr2:21070463:T:TAG at *NUTF2P8*, chr11:61820833:A:G at *FADS2*, chr19:11236981:C:T at *DOCK6*, chr11:116778201:G:C at *ZPR1*, and chr2:94961415:C:T at *MAL*. Other signals with acknowledged biological functions in metabolism of LDL-C were also identified, such as chr1:55039974:G:T at *PCSK9*, chr2:21011127:T:C at *APOB*, chr5:75343719:CTTGTA:C at *HMGCR*, chr6:160508872:A:G at *LPA*, chr8:20054859:G:A at *LPL*, and chr19:11120527:G:A at

LDLR.

The overall λ_{GC} was 1.17 for 38,827,546 variants tested (**Figure 5.2**). The λ_{GC} was higher among more common variants (1.42 for variants with 0.05 < MAF ≤0.5). This increase is attributable to these variants being in linkage disequilibrium with common causal variants, which inflates the test statistics. The λ_{GC} was lower in rarer variants (1.11 for variants with MAF <0.001). The use of a large population and high numbers of rare variants lower the overall λ_{GC}. For 9,634,325 variants with MAF>0.1% and cov-LD scores, the λ_{GC} was 1.27. The cov-LDSC intercept was 1.12, suggesting that 56% of genomic inflation was attributable to LD tagging between causal and non-causal variants and LDL-C as a polygenic trait, with only 44% was due to residual population structure and cryptic relatedness. The heritability in LDL-C estimated from the slope of the cov-LDSC regression line was 0.12.

5.3.2 Associations in GWAS of HDL-C

GWAS for HDL-C identified 1,384 variants with $p<5x10^{-8}$ at 434 loci (**Figure 5.3)**. Step-wise approximate conditional analysis identified 202 conditionally-independent signals at 135 loci (**Table 5.2**). Of these, 39 were previously reported, and 93 were ±500kb from variants that have been previously reported as associated with HDL-C, with 70 potentially novel variants, according to the published results on HDL-C measurements in GWAS Catalog. The top ten significant signals were chr16:56959412:C:A at *CETP*, chr11:116869766:C:T at *SIK3*, chr16:56973534:T:G at *CETP*, chr9:104858554:G:A at *ABCA1*, chr19:11235479:G:A at *DOCK6*, chr15:58431740:G:A at *LIPC*, chr8:9323885:A:G at *PPP1R3B*, chr15:58391167:A:G at *ALDH1A2*, and chr7:80664401:T:C at *CD36*. Other signals with acknowledged biological functions in metabolism of HDL-C were also identified, such as chr8:19973410:C:T at *LPL* and chr11:116791691:G:C at *APOA5*

The overall λ_{GC} was 1.17 for 38,827,546 variants tested (**Figure 5.4**). The λ_{GC} was higher

among more common variants (1.40 for variants with 0.05 < MAF ≤0.5). The λ_{GC} was lower in rarer variants (1.11 for variants with MAF <0.001). For 9,634,325 variants with MAF>0.1% and cov-LD scores, the λ_{GC} was 1.26. The cov-LDSC intercept was 1.13, suggesting that 50% of genomic inflation was attributable to LD tagging between causal and non-causal variants and HDL-C as a polygenic trait, with 50% was due to residual population structure and cryptic relatedness. The heritability in HDL-C estimated from the slope of the cov-LDSC regression line was 0.11.

5.3.3 Associations in GWAS of TG

GWAS for TG identified 2,001 variants with $p<5x10^{-8}$ at 481 loci (**Figure 5.5)**. Step-wise approximate conditional analysis identified 118 conditionally-independent signals at 75 loci (**Table 5.3**). Of these, 28 were previously reported, and 99 were ±500kb from variants that have been previously reported as associated with TG, with 36 potentially novel variants, according to the published results on TG measurements in GWAS Catalog. The top ten significant signals were chr11:116778201:G:C at *ZPR1*, chr2:27508073:T:C at *GCKR*, chr19:44927023:C:G at *APOC1P1*, chr2:21014672:G:A at *APOB*, chr8:125465736:G:C at *TRIB1*, chr15:58431740:G:A at *PNPLA3*, chr8:125495066:T:C at *TRIB1AL*, chr11:116830638:G:A at *APOC3*, and chr7:73623626:A:G at *LXIPL*. Other signals with acknowledged biological functions in metabolism of TG were also identified, such as chr6:160690668:C:T at *LPA*, chr15:43464986:T:C at *TP53BP1*, and chr19:11235479:G:A at *DOCK6*.

The overall λ_{GC} was 1.19 for 38,827,546 variants tested (**Figure 5.6)**. The λ_{GC} was higher among more common variants (1.55 for variants with 0.05 < MAF ≤0.5). The λ_{GC} was lower in rarer variants (1.11 for variants with MAF <0.001). For 9,634,325 variants with MAF>0.1% and cov-LD scores, the λ_{GC} was 1.33. The cov-LDSC intercept was 1.13, suggesting that 61% of genomic inflation was attributable to LD tagging between causal and non-causal variants and TG as a polygenic trait, with 39% was due to residual

population structure and cryptic relatedness. The heritability in TG estimated from the slope of the cov-LDSC regression line was 0.17.

5.3.4 Associations in GWAS of ApoA1

GWAS for ApoA1 identified 1,786 variants with $p<5x10^{-8}$ at 477 loci (**Figure 5.7)**. Step-wise approximate conditional analysis identified 129 conditionally-independent signals at 92 loci (**Table 5.4**). Of these, 14 were previously reported, and 64 were ±500kb from variants that have been previously reported as associated with ApoA1, with 51 potentially novel variants, according to the published results on ApoA1 measurements in GWAS Catalog. The top ten significant signals were chr11:116869766:C:T at *SIK3*, chr16:56959412:C:A at *CETP*, chr19:11239812:C:T at *ANGPTL8*, chr15:58431740:G:A at *LIPC*, chr9:104858554:G:A at *ABCA1*, chr22:43928847:C:G at *PNPLA3*, chr1:62620326:T:A at *DOCK7*, chr8:9323885:A:G at *PPP1R3B*, chr7:80664401:T:C at *CD36*, and chr18:49594129:T:G at *LIPG*. Other signals with acknowledged biological functions in metabolism of ApoA1 were also identified, such as chr2:27508073:T:C at *GCKR*, chr20:44413724:C:T at *HNF4A*, chr2:210675783:C:A at *CPS1*, chr19:44909976:G:T at *APOE*, and hr8:19966981:T:C at *LPL*.

The overall λ_{GC} was 1.15 for 38,827,546 variants tested (**Figure 5.8**). The λ_{GC} was higher among more common variants (1.38 for variants with $0.05 < MAF \leq 0.5$). The λ_{GC} was lower in rarer variants (1.10 for variants with MAF <0.001). For 9,634,325 variants with MAF>0.1% and cov-LD scores, the λ_{GC} was 1.24. The cov-LDSC intercept was 1.12, suggesting that 50% of genomic inflation was attributable to LD tagging between causal and non-causal variants and ApoA1 as a polygenic trait, with only 50% was due to residual population structure and cryptic relatedness. The heritability in ApoA1 estimated from the slope of the cov-LDSC regression line was 0.11.

5.3.5 Associations in GWAS of ApoB

GWAS for ApoB identified 1,786 variants with $p<5x10^{-8}$ at 641 loci (**Figure 5.9)**. Step-

wise approximate conditional analysis identified 160 conditionally-independent signals at 113 loci (**Table 5.5**). Of these, 24 were previously reported, and 69 were ±500kb from variants that have been previously reported as associated with ApoB, with 67 potentially novel variants, according to the published results on ApoB measurements in GWAS Catalog. The top 10 significant signals included chr1:109274968:G:T at *CELSR2*, chr19:44908822:C:T at *APOE*, chr2:21043902:GGCAGCGCCA:G at *APOB*, chr19:11081874:G:A at *SMARCA4*, chr19:11086585:G:A at *LDLR*, chr9:133273983:A:G at *ABO*, chr11:61820833:A:G at *FADS2*, chr1:55039974:G:T at *PCSK9*, chr19:11236981:C:T at *DOCK6*, and chr5:75360714:T:C at *HMGCR*. Other signals with acknowledged biological functions in metabolism of ApoB were also identified, such as chr16:56983407:G:A at *CETP*, chr6:160585045:C:T at *LPA*, chr2:27508073:T:C at GCKR, chr13:32412082:A:G at *N4BP2L1*, and chr20:44413724:C:T at *HNF4A*.

The overall λ_{GC} was 1.18 for 38,827,546 variants tested (**Figure 5.10**). The λ_{GC} was higher among more common variants (1.45 for variants with 0.05 < MAF ≤0.5). The λ_{GC} was lower in rarer variants (1.11 for variants with MAF <0.001). For 9,634,325 variants with MAF>0.1% and cov-LD scores, the λ_{GC} was 1.29. The cov-LDSC intercept was 1.12, suggesting that 59% of genomic inflation was attributable to LD tagging between causal and non-causal variants and ApoB as a polygenic trait, with 41% was due to residual population structure and cryptic relatedness. The heritability in ApoB estimated from the slope of the cov-LDSC regression line was 0.14.

5.4 Discussion

In this chapter GWAS were conducted for five major lipid-related traits measured by NMR (LDL-C, HDL-C, TG, ApoB, and ApoA1) using 38,827,546 well-imputed variants in a large Latin-American population (*n*=122,543). Participants known to be on lipid-lowering drugs were excluded and the additive effect of each SNP after adjusting for age, sex, genetic PCs , batch, and fasting duration was estimated. For LDL-C, 135

conditionally-independent signals (36 potentially novel) at 92 loci were identified (e.g., *APOE, CELSR2, PCSK9, APOB, HMGCR, LPA, LPL, LDLR*). There were 202 independent signals (70 novel) at 135 loci for HDL-C (e.g., *CETP, SIK3, LIPC, LPL, APOA5*), 118 independent signals (36 novel) at 75 loci for TG (e.g., *ZPR1, GCKR, APOC1P1, APOC3, LPA, DOCK6*), 129 independent signals (51 novel) at 92 loci for ApoA1 (e.g., *SIK3, CETP, ANGPTL8, LIPC, ABCA1*), and 160 independent signals (67 novel) at 113 loci for ApoB (e.g., *CELSR2, APOE, APOB, LDLR, PCSK9, CETP, LPA*). The intercepts estimated from covariate-adjusted LDSC ranged from 1.12 to 1.13, suggesting reasonable control of genomic inflation. The covariate-adjusted LDSC demonstrated modest heritability for these major lipid traits (h^2 ranging from 0.11-0.17) in this large, related sample of an admixed Latin-American population.

The numbers of independent, significant signals identified for each lipid trait in the current study are generally higher than findings from previous studies, given the similar sample size. The current largest multi-ancestry GWAS on blood lipids (LDL-C, HDL-C, TG, TC, and non HDL-C) including 1.65 million individuals (350,000 of non-European ancestries) by GLGC identified a total of 1,750 index variants at 923 loci (±500 kb regions) that reached genome-wide significance for at least one lipid trait.[26] This study revealed that ancestry-specific analyses identified more loci per genotyped individual (e.g., ~50 loci per 100,000 individuals for Admixed African and Hispanic ancestry-specific analyses) than European and multi-ancestry analyses (~30 loci per 100,000 individuals). Similar to Admixed African populations, the current Admixed Mexican population has greater genetic diversity than other populations.[27] These previous findings, together with the results in this chapter, suggest that including underrepresented populations with greater genetic diversity (rather than studying additional individuals of European ancestry) may improve the genetic discovery and enhance the fine-mapping of functional variants. However, the number of signals and loci still appears to be higher than the GLGC results[26] and other ancestry-specific

GWAS[28]. A recent GWAS meta-analysis of ~125,000 individuals of African ancestry identified 63, 68, 48, and 92 independent genetic loci (500 kb around lead SNP) associated with HDL, LDL, TG and TC at genome-wide significance.[28] Considering the lack of representation of diverse non-European ancestry population (e.g., Latin-American) in current GWAS of blood lipid, future investigations should further evaluate the improvement of genetic discovery by including multi-ancestry populations to increase genetic diversity.

The heritability for major lipid traits (0.11-0.17) estimated by cov-LDSC in the current study is lower than previous studies (0.3-0.7).[29] Results from the largest multi-ancestry GWAS by GLGC indicated that lipid-related variants accounted for 9% to 13% of the variance in lipid traits overall, and 10% to 13% in the Hispanic population analysis; however, heritability estimates were not reported.[26] Heritability estimated by cov-LDSC in the current study may be underestimated due to measurement error, differences in allele frequencies across population, and/or the adjustment for population structure (as well as some of environmental differences across sub-populations) by cov-LDSC[30] in the admixed population. These differences in heritability further highlight the importance of studying diverse populations.

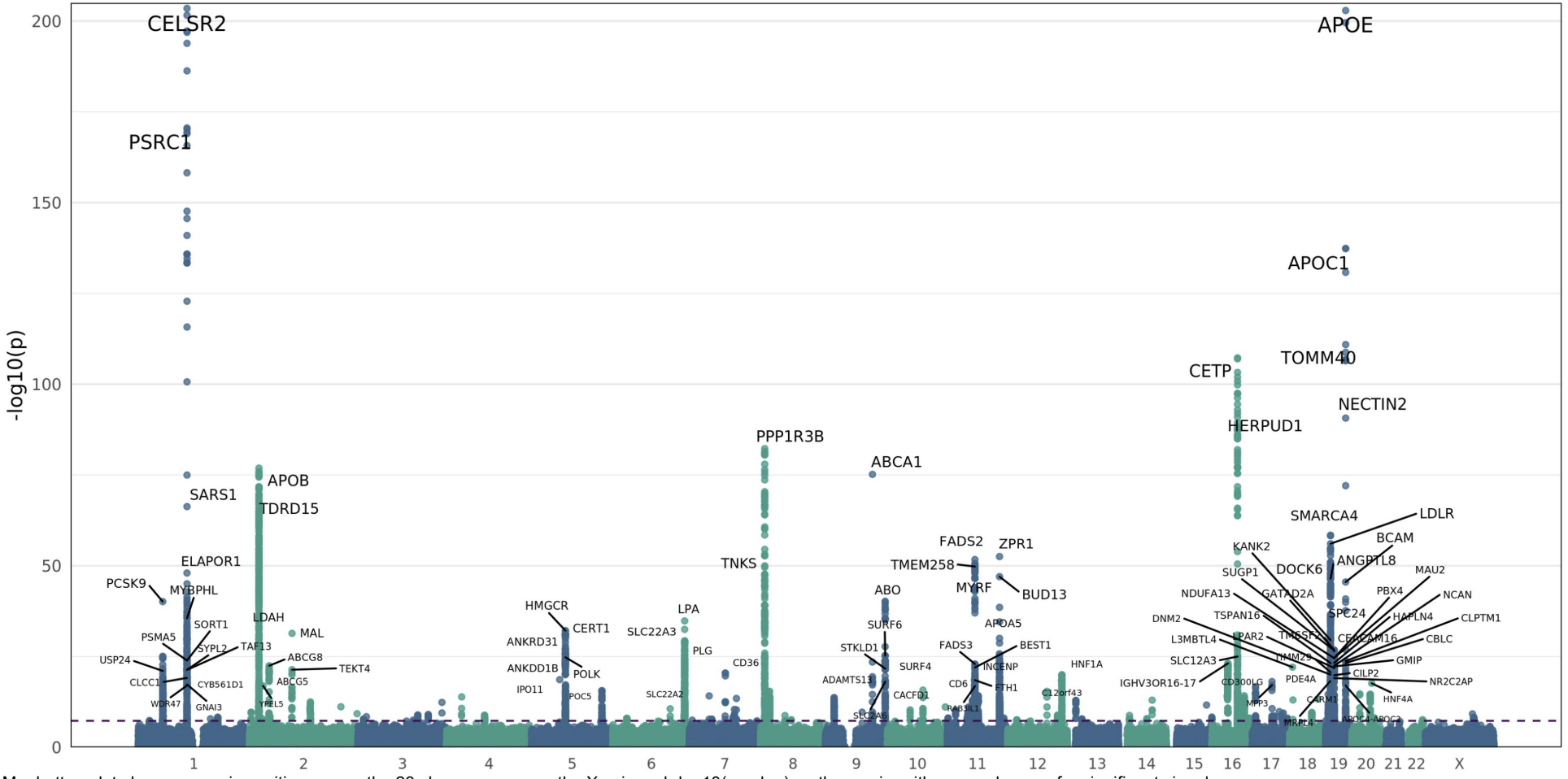
The major strength of the current GWAS on major lipid traits include a large sample size from a population of under-studied Latin-American ancestry and inclusion of major lipids and apolipoproteins. The current sample size is six times that of Hispanic participants in the MVP lipid GWAS[31] (n = 24,000) and three times that of the latest GLGC GWAS[26] (n = 48,057). The current analysis also includes two major apolipoproteins (ApoB and ApoA1) which were largely unavailable in previous large meta-analysis of lipid GWAS. [26,31] Participants taking lipid-lowering medications were excluded at baseline to better capture untreated lipid levels, which were also largely unavailable in previous GWAS based on electronic health records or multiple small cohorts. Cov-LDSC was used where cov-LD scores were estimated based on

participants with WGS data (9,634,325 variants with MAF>0.1%) to better account for the population structure when assessing genomic inflation and heritability in a population with complex admixture patterns. Several previously known genetic loci involved in lipid metabolism were identified, including *CELSR2*, *APOE*, *APOB*, *LDLR*, *PCSK9*, *CETP*, and *LPA*, some of which are direct targets of lipid-lowering medications. Several potentially novel loci were also identified. However, the analyses aimed at identifying these loci should be considered preliminary. Replicating these novel loci in other studies is beyond the scope of the current thesis but would be necessary, along with further downstream analyses to explore gene function and enrichment. The variants included in the current study are more common in individuals of Central and Southern Mexican ancestry than analyses in cohorts recruited from other regions of the Americas.[27] The findings may therefore not be generalizable to other regions of Mexico or other Latin-American populations, where genetic backgrounds can vary significantly (although the future validation of novel genetic loci could lead to translatable insights that could benefit all populations). Other limitations include the lack of insight into the detailed NMR lipoprotein subfractions, while further work needs to be done to control for the residual confounding observed (which may require the development of novel approaches, such as haplotype-based methods).

5.5 Conclusion

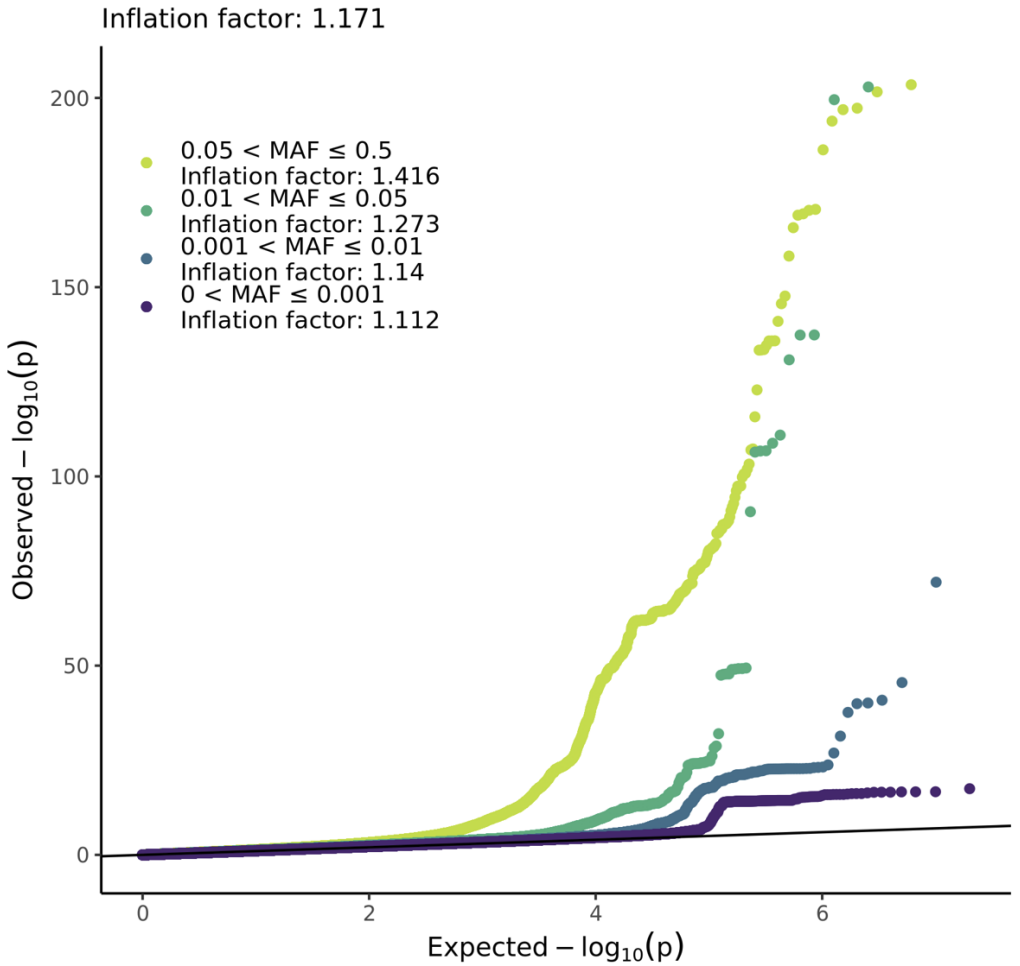
GWAS of five major lipid-related traits (LDL-C, HDL-C, TG, ApoB, and ApoA1) using 38,827,546 well-imputed variants in this Mexican population identified multiple previously known genetic loci involved in lipid metabolism as well as several potentially novel loci. Future replication of these novel loci in other Latin American-ancestry populations and downstream analyses for the function and enrichment of the identified loci may provide additional information on lipoprotein biology and elucidate the causal roles of lipid traits in ASCVD that may offer benefits for all populations.

Figure 5.1. Manhattan plot for the GWAS of NMR-measured LDL-C



Manhattan plot shows genomic position across the 23 chromosomes on the X axis and -log10(p-value) on the y axis, with mapped genes for significant signals.

Figure 5.2. QQ plot for GWAS of NMR-measured LDL-C



QQ plot split by the MAF of genetic variants. QQ: quantile-quantile; MAF: minor allele frequency.

Table 5.1. Conditionally independent signals from fine-mapped results of the LDL-C GWAS

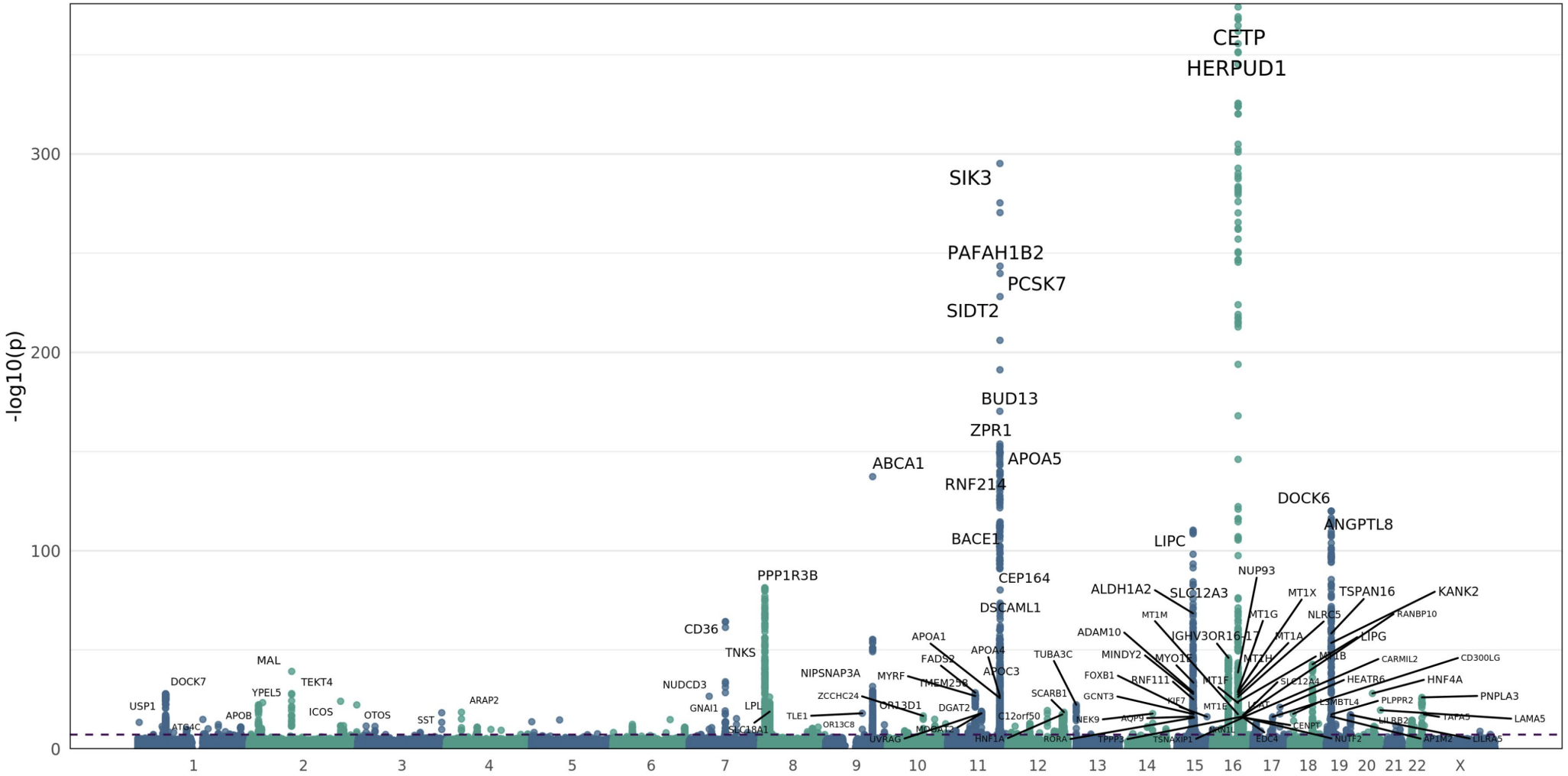
Chr	Variant ID	bp	ref	b	se	p	gene	rsID	Reported variant	Known locus
1	chr1:55021673:C:G	55021673	C	0.04903	0.006257	4.63E-15	*BSND*	rs12117661	1	1
1	chr1:55039974:G:T	55039974	G	0.438927	0.032793	7.43E-41	*PCSK9*	rs11591147	1	1
1	chr1:109274968:G:T	1.09E+08	G	0.164555	0.005396	3.22E-204	*CELSR2*	rs12740374	1	1
1	chr1:161642443:T:C	1.62E+08	T	-0.09589	0.016902	1.40E-08	*FCGR3B*	NA	NA	1
2	chr2:3594771:T:C	3594771	T	0.028243	0.004448	2.16E-10	*COLEC11*	rs3820897	1	1
2	chr2:21043902:GGCAGCGCCA:G	21043902	GGCAGCGCCA	-0.08099	0.004343	1.30E-77	*APOB*	NA	NA	1
2	chr2:30193959:T:G	30193959	T	-0.34665	0.040497	1.13E-17	*YPEL5*	NA	NA	NA
2	chr2:43864927:C:T	43864927	C	0.073447	0.007403	3.38E-23	*ABCG8*	rs4603816	1	1
2	chr2:94961415:C:T	94961415	C	-0.3181	0.026969	4.14E-32	*MAL*	NA	NA	NA
2	chr2:135950412:T:C	1.36E+08	T	-0.04788	0.006567	3.09E-13	*DARS1*	NA	NA	1
2	chr2:204107315:C:T	2.04E+08	C	-0.50023	0.072874	6.68E-12	*ICOS*	NA	NA	NA
2	chr2:240217492:G:A	2.4E+08	G	-0.19787	0.0319	5.54E-10	*OTOS*	NA	NA	NA
3	chr3:12362332:G:A	12362332	G	-0.03405	0.006036	1.69E-08	*PPARG*	NA	NA	1
3	chr3:18939493:T:C	18939493	T	-0.49616	0.08413	3.69E-09	*KCNH8*	NA	NA	NA
3	chr3:133729317:T:G	1.34E+08	T	-0.04655	0.007669	1.29E-09	*TF*	NA	NA	1
3	chr3:157079820:T:C	1.57E+08	T	-0.02582	0.004213	8.87E-10	*LEKR1*	NA	NA	NA
3	chr3:187669234:T:G	1.88E+08	T	-0.20792	0.028683	4.21E-13	*SST*	NA	NA	NA
4	chr4:33595172:C:T	33595172	C	-0.20736	0.026906	1.29E-14	*ARAP2*	NA	NA	NA
4	chr4:84647685:C:T	84647685	C	0.071301	0.011779	1.42E-09	*CDS1*	NA	NA	NA
5	chr5:62615128:G:A	62615128	G	-0.30686	0.034107	2.32E-19	*IPO11*	NA	NA	NA
5	chr5:75343719:CTTGTA:C	75343719	CTTGTA	-0.05045	0.004224	6.96E-33	*HMGCR*	NA	NA	1
5	chr5:156968992:C:T	1.57E+08	C	-0.04894	0.00597	2.46E-16	*TIMD4*	rs12657266	1	1
6	chr6:31304888:A:G	31304888	A	-0.03624	0.006335	1.06E-08	*HLA-C*	NA	NA	1
6	chr6:128048785:G:A	1.28E+08	G	-0.10972	0.016429	2.42E-11	*PTPRK*	NA	NA	NA
6	chr6:160508872:A:G	1.61E+08	A	0.054437	0.004375	1.53E-35	*LPA*	NA	NA	1
7	chr7:6581140:T:C	6581140	T	0.259068	0.046869	3.25E-08	*ZDHHC4*	NA	NA	1
7	chr7:44498864:T:C	44498864	T	-0.34645	0.044502	6.97E-15	*NUDCD3*	NA	NA	1
7	chr7:44539580:C:A	44539580	C	-0.0467	0.008283	1.72E-08	*NPC1L1*	NA	NA	1
7	chr7:80604075:C:T	80604075	C	-0.10854	0.011489	3.48E-21	*CD36*	NA	NA	NA
7	chr7:98312758:C:CTG	98312758	C	-0.03071	0.005229	4.24E-09	*BAIAP2L1*	NA	NA	1
7	chr7:100814752:T:C	1.01E+08	T	-0.02839	0.00429	3.67E-11	*EPHB4*	NA	NA	1
7	chr7:105394429:C:T	1.05E+08	C	-0.17563	0.023179	3.53E-14	*SRPK2*	NA	NA	1

8	chr8:9323885:A:G	9323885	A	-0.08786	0.004552	5.43E-83	*PPP1R3B*	rs2169387	1	1
8	chr8:20054859:G:A	20054859	G	-0.08184	0.010046	3.76E-16	*LPL*	NA	NA	1
8	chr8:54524077:C:A	54524077	C	-0.04374	0.00727	1.78E-09	*RP1*	NA	NA	1
8	chr8:68422706:C:T	68422706	C	-0.6108	0.109786	2.64E-08	*C8orf34*	NA	NA	NA
9	chr9:19228003:T:A	19228003	T	0.041903	0.005475	1.96E-14	*DENND4C*	NA	NA	1
9	chr9:81781896:G:A	81781896	G	-0.49238	0.0775	2.11E-10	*TLE1*	NA	NA	NA
9	chr9:104858554:G:A	1.05E+08	G	0.123765	0.006713	6.80E-76	*ABCA1*	NA	NA	1
9	chr9:131057085:G:A	1.31E+08	G	-0.31493	0.051009	6.66E-10	*LAMC3*	NA	NA	NA
9	chr9:133273983:A:G	1.33E+08	A	0.093495	0.006976	5.87E-41	*ABO*	rs507666	1	1
10	chr10:50814156:T:C	50814156	T	-0.07072	0.010754	4.83E-11	*ASAH2B*	NA	NA	1
10	chr10:72904011:G:A	72904011	G	0.057564	0.010516	4.40E-08	*OIT3*	NA	NA	1
10	chr10:79473458:T:G	79473458	T	-0.5608	0.068081	1.76E-16	*ZCCHC24*	NA	NA	NA
10	chr10:129475679:G:A	1.29E+08	G	-0.18659	0.027271	7.82E-12	*MGMT*	NA	NA	NA
11	chr11:1863112:G:A	1863112	G	0.025771	0.004469	8.11E-09	*LSP1*	NA	NA	1
11	chr11:18624296:C:A	18624296	C	-0.0279	0.004259	5.68E-11	*SPTY2D1*	NA	NA	1
11	chr11:47925284:G:A	47925284	G	0.03014	0.004404	7.71E-12	*PTPRJ*	NA	NA	1
11	chr11:49900650:C:T	49900650	C	-0.11904	0.018371	9.17E-11	*OR4C13*	NA	NA	NA
11	chr11:61820833:A:G	61820833	A	0.078923	0.005179	1.98E-52	*FADS2*	rs174564	1	1
11	chr11:63490651:C:T	63490651	C	-0.12545	0.017426	6.06E-13	*PLAAT5*	NA	NA	1
11	chr11:116778201:G:C	1.17E+08	G	-0.06623	0.004311	2.90E-53	*ZPR1*	rs964184	1	1
12	chr12:20279369:G:A	20279369	G	-0.02591	0.004655	2.58E-08	*PDE3A*	NA	NA	NA
12	chr12:87714165:A:G	87714165	A	-0.21042	0.02584	3.85E-16	*C12orf50*	NA	NA	NA
12	chr12:111834972:C:T	1.12E+08	C	-0.03569	0.006116	5.38E-09	*MAPKAPK5*	NA	NA	1
12	chr12:120985583:A:G	1.21E+08	A	0.04066	0.004356	1.01E-20	*HNF1A*	NA	NA	1
13	chr13:18926972:T:TAA	18926972	T	-0.51512	0.069436	1.18E-13	*TUBA3C*	NA	NA	NA
13	chr13:32367579:A:AGATGGCTTG	32367579	A	0.023608	0.004167	1.46E-08	*BRCA2*	NA	NA	1
14	chr14:24402720:C:T	24402720	C	-0.03566	0.005912	1.62E-09	*NYNRIN*	rs11621792	1	1
14	chr14:63134329:G:A	63134329	G	-0.10215	0.017933	1.22E-08	*KCNH5*	NA	NA	NA
14	chr14:75093953:T:C	75093953	T	-0.15007	0.020176	1.02E-13	*NEK9*	NA	NA	1
15	chr15:89642262:C:G	89642262	C	-0.81072	0.115509	2.24E-12	*KIF7*	NA	NA	NA
15	chr15:101527397:A:G	1.02E+08	A	0.030869	0.005264	4.52E-09	*PCSK6*	NA	NA	1
16	chr16:35442020:G:T	35442020	G	-0.77941	0.077407	7.58E-24	*IGHV3OR16-17*	NA	NA	NA
16	chr16:56972466:A:G	56972466	A	0.11225	0.005085	5.63E-108	*CETP*	NA	NA	1
16	chr16:72017470:A:G	72017470	A	-0.0396	0.005076	6.14E-15	*DHODH*	NA	NA	1

16	chr16:83945717:A:G	83945717	A	0.02548	0.004432	8.95E-09	*OSGIN1*	NA	NA	1
17	chr17:7176750:A:G	7176750	A	0.037644	0.004428	1.88E-17	*ASGR1*	rs12945299	1	1
17	chr17:43848758:C:T	43848758	C	0.227883	0.025712	7.81E-19	*CD300LG*	rs72836561	1	1
17	chr17:47659909:A:T	47659909	A	0.027703	0.004639	2.35E-09	*KPNB1*	rs4264433	1	1
17	chr17:60094576:T:C	60094576	T	-0.14447	0.022543	1.47E-10	*HEATR6*	NA	NA	1
17	chr17:66199313:A:G	66199313	A	-0.09958	0.016128	6.65E-10	*CEP112*	NA	NA	1
17	chr17:69086821:C:G	69086821	C	0.034676	0.005526	3.48E-10	*ABCA6*	rs740516	1	1
17	chr17:78393095:C:G	78393095	C	-0.02683	0.004665	8.84E-09	*PGS1*	NA	NA	1
18	chr18:6512865:G:A	6512865	G	-0.31977	0.032535	8.47E-23	*L3MBTL4*	NA	NA	NA
18	chr18:7432877:G:A	7432877	G	-0.19811	0.026574	8.98E-14	*PTPRM*	NA	NA	NA
18	chr18:49594230:G:A	49594230	G	-0.03192	0.005038	2.38E-10	*LIPG*	rs9953437	1	1
18	chr18:63195322:G:A	63195322	G	-0.24113	0.032022	5.06E-14	*BCL2*	NA	NA	1
19	chr19:1115299:G:A	1115299	G	0.030396	0.005535	3.98E-08	*SBNO2*	NA	NA	1
19	chr19:5840608:C:T	5840608	C	-0.03353	0.004239	2.59E-15	*FUT6*	rs708686	1	1
19	chr19:8364444:G:C	8364444	G	-0.07156	0.010606	1.51E-11	*ANGPTL4*	NA	NA	1
19	chr19:11236981:C:T	11236981	C	0.067165	0.004572	7.36E-49	*DOCK6*	NA	NA	1
19	chr19:19492132:G:T	19492132	G	0.080507	0.007405	1.56E-27	*GATAD2A*	NA	NA	1
19	chr19:44908822:C:T	44908822	C	0.45628	0.014985	1.24E-203	*APOE*	rs7412	1	1
19	chr19:44935906:C:G	44935906	C	0.045724	0.006387	8.11E-13	*APOC4*	rs35136575	1	1
20	chr20:17863848:G:T	17863848	G	0.036388	0.004576	1.84E-15	*SNX5*	rs2618567	1	1
20	chr20:41169778:G:C	41169778	G	-0.03325	0.004202	2.50E-15	*PLCG1*	NA	NA	1
20	chr20:44413724:C:T	44413724	C	0.087362	0.009998	2.37E-18	*HNF4A*	rs1800961	1	1
20	chr20:50845709:G:A	50845709	G	-0.11536	0.019962	7.53E-09	*BCAS4*	NA	NA	NA
22	chr22:46965675:C:G	46965675	C	-0.92529	0.165964	2.47E-08	*TBC1D22A*	NA	NA	NA
22	chr22:48386401:G:A	48386401	G	-0.1617	0.029651	4.94E-08	*TAFA5*	NA	NA	NA
23	chrX:107821976:C:T	1.08E+08	C	0.147959	0.023917	6.15E-10	*MID2*	NA	NA	NA

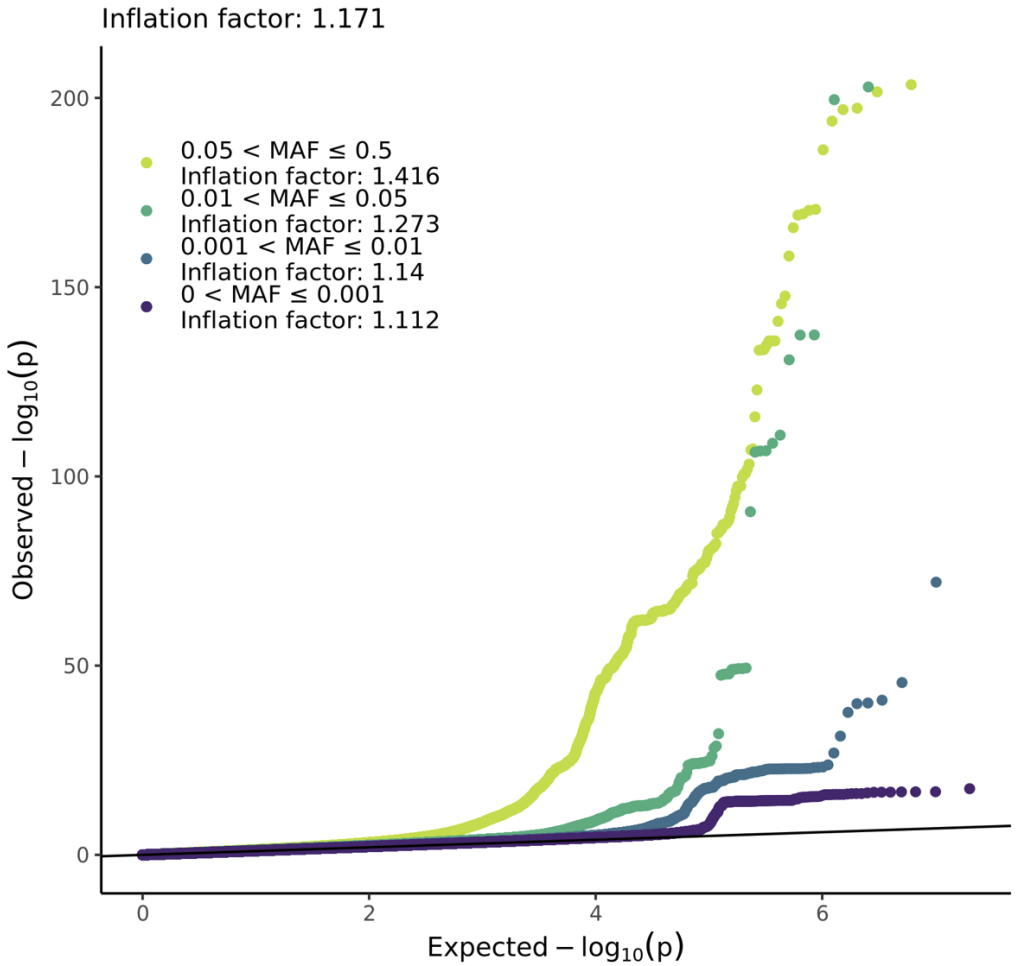
Beta, se and p-value reported were from conditional analysis for conditionally-independent signal. The most significant signal for each mapped gene was reported. SNP and PUBMED ID for reported variants were showed. BP: base position; Chr: chromosome; b, beta; se: standard error; NA: data not available; rsID: reference SNP (single nucleotide polymorphism) cluster ID.

Figure 5.3. Manhattan plot for the GWAS of NMR-measured HDL-C



Manhattan plot shows genomic position across the 23 chromosomes on the X axis and -log10(p-value) on the y axis, with mapped genes for significant signals.

Figure 5.4. QQ plot for GWAS of NMR-measured HDL-C



QQ plot split by the MAF of genetic variants. QQ: quantile-quantile; MAF: minor allele frequency.

Table 5.2. Conditionally independent signals from fine-mapped results of the HDL-C GWAS

Chr	Variant ID	bp	ref	b	se	p	gene	rsID	Reported variant	Known locus
1	chr1:80467956:C:T	80467956	C	-0.17563	0.02991	4.31E-09	*ADGRL2*	NA	NA	NA
1	chr1:150870220:C:T	1.51E+08	C	-0.02475	0.004405	1.92E-08	*ARNT*	rs72992015	1	1
1	chr1:110811781:G:A	1.11E+08	G	-0.27572	0.04366	2.70E-10	*CD53*	NA	NA	NA
1	chr1:109275216:T:C	1.09E+08	T	-0.03424	0.005244	6.64E-11	*CELSR2*	rs660240	1	1
1	chr1:68622607:C:T	68622607	C	-0.19246	0.031878	1.57E-09	*DEPDC1*	NA	NA	1
1	chr1:62668259:G:A	62668259	G	0.046106	0.004155	1.31E-28	*DOCK7*	NA	NA	1
1	chr1:157406268:T:G	1.57E+08	T	-0.23704	0.036177	5.67E-11	*FCRL5*	NA	NA	NA
1	chr1:230159944:G:A	2.3E+08	G	-0.0281	0.004104	7.56E-12	*GALNT2*	rs4846914	1	1
1	chr1:145655702:C:T	1.46E+08	C	-0.17042	0.021272	1.13E-15	*GPR89A*	NA	NA	1
1	chr1:195916851:A:C	1.96E+08	A	-0.23423	0.038866	1.67E-09	*KCNT2*	NA	NA	NA
1	chr1:3593253:C:T	3593253	C	-0.28096	0.037052	3.38E-14	*MEGF6*	NA	NA	NA
1	chr1:54703968:T:C	54703968	T	-0.18239	0.026225	3.53E-12	*MROH7-TTC4*	NA	NA	1
1	chr1:147062045:C:T	1.47E+08	C	-0.13413	0.023876	1.93E-08	*NBPF12*	NA	NA	NA
1	chr1:56615655:T:C	56615655	T	-0.09298	0.014123	4.58E-11	*PLPP3*	NA	NA	1
1	chr1:180182877:C:T	1.8E+08	C	-0.28396	0.039159	4.13E-13	*QSOX1*	NA	NA	NA
1	chr1:205668373:G:C	2.06E+08	G	0.026199	0.004711	2.68E-08	*SLC45A3*	NA	NA	1
1	chr1:108752530:C:T	1.09E+08	C	-0.12242	0.021339	9.65E-09	*STXBP3*	NA	NA	NA
2	chr2:21008652:G:A	21008652	G	-0.04474	0.00452	4.26E-23	*APOB*	rs676210	1	1
2	chr2:210675783:C:A	2.11E+08	C	0.031507	0.004459	1.60E-12	*CPS1*	rs1047891	1	1
2	chr2:4526958:G:T	4526958	G	-0.61314	0.110267	2.69E-08	*DCDC2C*	NA	NA	NA
2	chr2:27508073:T:C	27508073	T	0.02918	0.004386	2.88E-11	*GCKR*	rs1260326	1	1
2	chr2:204107315:C:T	2.04E+08	C	-0.72892	0.070838	7.82E-25	*ICOS*	NA	NA	1
2	chr2:94961415:C:T	94961415	C	-0.34657	0.026215	6.71E-40	*MAL*	NA	NA	NA
2	chr2:240217492:G:A	2.4E+08	G	-0.30619	0.031008	5.37E-23	*OTOS*	NA	NA	NA
2	chr2:30193959:T:G	30193959	T	-0.39902	0.039365	3.80E-24	*YPEL5*	NA	NA	1
3	chr3:62398777:G:A	62398777	G	-0.56684	0.103445	4.26E-08	*CADPS*	NA	NA	1
3	chr3:18939493:T:C	18939493	T	-0.5744	0.082052	2.55E-12	*KCNH8*	NA	NA	NA
3	chr3:139717549:C:A	1.4E+08	C	-0.80928	0.137199	3.67E-09	*NMNAT3*	NA	NA	NA
3	chr3:38690037:C:T	38690037	C	-0.28879	0.041494	3.41E-12	*SCN10A*	NA	NA	NA
3	chr3:187669234:T:G	1.88E+08	T	-0.2495	0.027975	4.71E-19	*SST*	NA	NA	NA
4	chr4:99323162:G:A	99323162	G	-0.29334	0.045597	1.25E-10	*ADH1B*	NA	NA	1
4	chr4:33595172:C:T	33595172	C	-0.23559	0.026223	2.61E-19	*ARAP2*	NA	NA	NA

4	chr4:66149466:G:A	66149466	G	-0.19637	0.035662	3.66E-08	*EPHA5*	NA	NA	NA
4	chr4:122131188:C:T	1.22E+08	C	-0.22631	0.036004	3.26E-10	*KIAA1109*	NA	NA	NA
4	chr4:166044830:A:C	1.66E+08	A	-0.28069	0.051129	4.02E-08	*TLL1*	NA	NA	NA
4	chr4:68672462:G:A	68672462	G	-0.0295	0.004325	9.04E-12	*UGT2B15*	NA	NA	1
5	chr5:5701961:G:C	5701961	G	0.975887	0.165728	3.90E-09	*ICE1*	NA	NA	NA
5	chr5:62615128:G:A	62615128	G	-0.26382	0.033216	1.98E-15	*IPO11*	NA	NA	NA
5	chr5:523946:C:T	523946	C	-1.10188	0.143566	1.65E-14	*SLC9A3*	NA	NA	1
5	chr5:147231394:G:A	1.47E+08	G	-0.13094	0.023314	1.95E-08	*STK32A*	NA	NA	NA
6	chr6:119008344:G:T	1.19E+08	G	-0.11388	0.019591	6.14E-09	*FAM184A*	NA	NA	NA
6	chr6:1592690:C:T	1592690	C	-0.10649	0.018135	4.30E-09	*FOXC1*	NA	NA	NA
6	chr6:160504823:G:T	1.61E+08	G	0.028008	0.00411	9.50E-12	*LPA*	NA	NA	1
6	chr6:128048785:G:A	1.28E+08	G	-0.12795	0.016016	1.36E-15	*PTPRK*	NA	NA	NA
6	chr6:43796814:C:T	43796814	C	0.030572	0.0042	3.34E-13	*VEGFA*	rs1358980	1	1
7	chr7:80664401:T:C	80664401	T	-0.19518	0.01146	4.85E-65	*CD36*	NA	NA	1
7	chr7:50294364:C:T	50294364	C	-0.02413	0.0043	2.00E-08	*IKZF1*	NA	NA	1
7	chr7:130484073:C:T	1.3E+08	C	0.513377	0.087068	3.72E-09	*MEST*	NA	NA	1
7	chr7:26076003:C:T	26076003	C	-0.0285	0.005193	4.06E-08	*NFE2L3*	NA	NA	1
7	chr7:44498864:T:C	44498864	T	-0.47004	0.04339	2.41E-27	*NUDCD3*	NA	NA	1
7	chr7:105394429:C:T	1.05E+08	C	-0.1838	0.0226	4.20E-16	*SRPK2*	NA	NA	NA
8	chr8:104456502:T:G	1.04E+08	T	-0.12896	0.021409	1.70E-09	*DPYS*	NA	NA	1
8	chr8:9054627:C:T	9054627	C	-0.07589	0.013411	1.53E-08	*ERI1*	NA	NA	1
8	chr8:19973410:C:T	19973410	C	-0.08845	0.008226	5.78E-27	*LPL*	rs10096633	1	1
8	chr8:143176727:G:T	1.43E+08	G	-0.09901	0.018122	4.66E-08	*LY6H*	NA	NA	1
8	chr8:128337012:G:A	1.28E+08	G	-0.15681	0.022954	8.41E-12	*MYC*	NA	NA	NA
8	chr8:125266703:G:A	1.25E+08	G	0.580739	0.099542	5.41E-09	*NSMCE2*	NA	NA	1
8	chr8:9323885:A:G	9323885	A	-0.0852	0.004439	4.23E-82	*PPP1R3B*	rs2169387	1	1
8	chr8:115483682:T:A	1.15E+08	T	-0.02774	0.0042	3.97E-11	*TRPS1*	NA	NA	1
9	chr9:104858554:G:A	1.05E+08	G	0.163987	0.006556	4.58E-138	*ABCA1*	rs9282541	1	1
9	chr9:131057085:G:A	1.31E+08	G	-0.35947	0.049817	5.36E-13	*LAMC3*	NA	NA	NA
9	chr9:81781896:G:A	81781896	G	-0.66975	0.075689	8.86E-19	*TLE1*	NA	NA	NA
10	chr10:45442791:T:C	45442791	T	-0.03107	0.005335	5.75E-09	*ALOX5*	rs2242332	1	1
10	chr10:112172941:T:C	1.12E+08	T	0.028754	0.004867	3.47E-09	*GPAM*	NA	NA	1
10	chr10:129475679:G:A	1.29E+08	G	-0.19592	0.026603	1.77E-13	*MGMT*	NA	NA	1
10	chr10:105053006:G:A	1.05E+08	G	-0.27498	0.034215	9.22E-16	*SORCS3*	NA	NA	NA

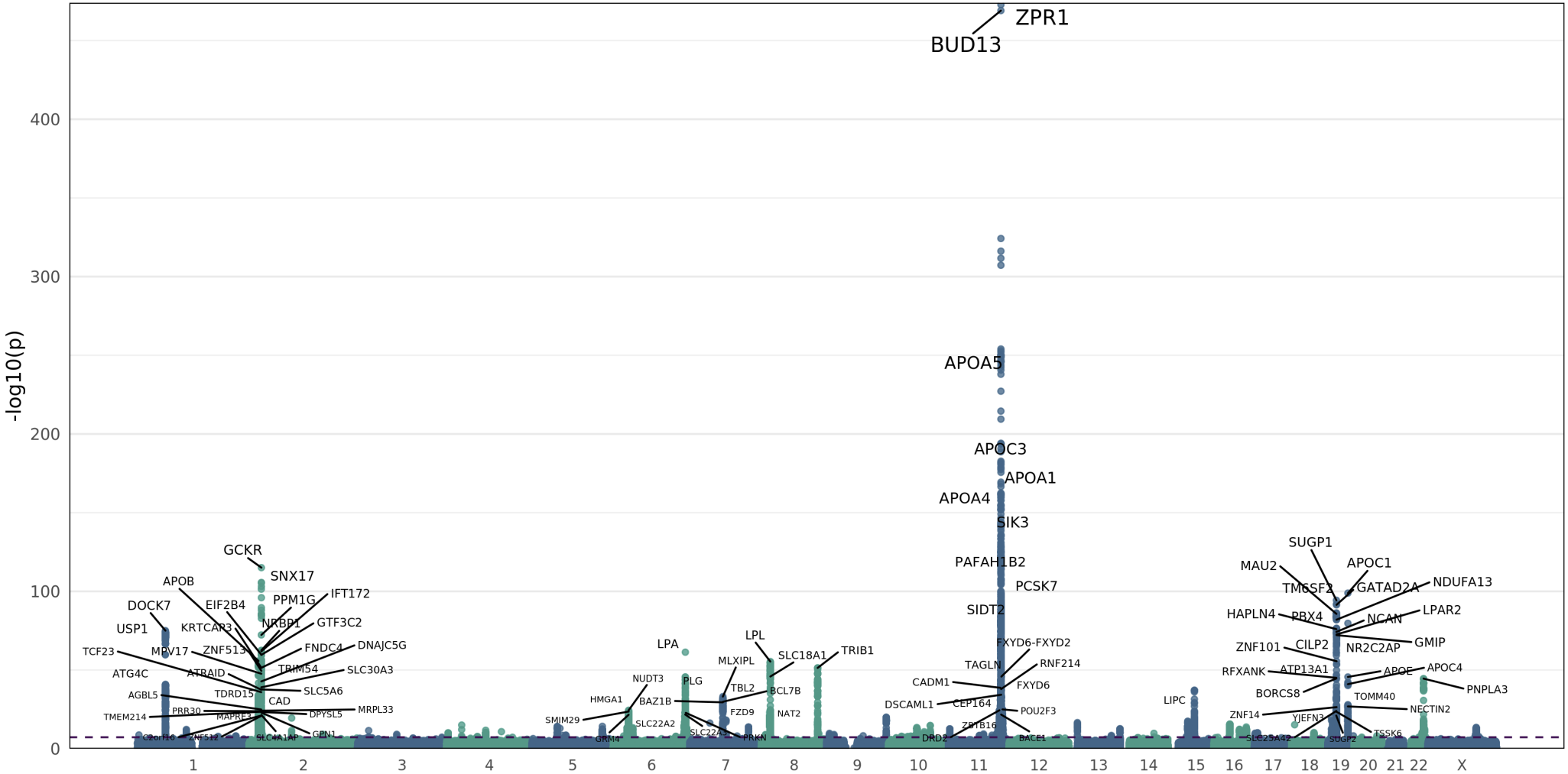
10	chr10:79473458:T:G	79473458	T	-0.5648	0.066411	1.82E-17	*ZCCHC24*	NA	NA	1
10	chr10:38133001:T:A	38133001	T	-0.37551	0.060796	6.55E-10	*ZNF37A*	NA	NA	NA
11	chr11:116791691:G:C	1.17E+08	G	0.046064	0.005463	3.41E-17	*APOA5*	rs3135506	1	1
11	chr11:116830638:G:A	1.17E+08	G	-0.86479	0.079524	1.52E-27	*APOC3*	rs138326449	1	1
11	chr11:75757381:C:CAG	75757381	C	0.059113	0.006497	9.12E-20	*DGAT2*	NA	NA	1
11	chr11:61820833:A:G	61820833	A	0.056774	0.005064	3.60E-29	*FADS2*	rs174564	1	1
11	chr11:49900650:C:T	49900650	C	-0.12601	0.017962	2.30E-12	*OR4C13*	NA	NA	1
11	chr11:63490651:C:T	63490651	C	-0.12019	0.017038	1.74E-12	*PLAAT5*	NA	NA	NA
11	chr11:116869766:C:T	1.17E+08	C	-0.23701	0.006446	6.21E-296	*SIK3*	rs188287950	1	1
11	chr11:130173078:T:G	1.3E+08	T	-0.16253	0.028668	1.43E-08	*ST14*	NA	NA	NA
12	chr12:103129174:C:T	1.03E+08	C	0.025992	0.004219	7.27E-10	*C12orf42*	NA	NA	1
12	chr12:87714165:A:G	87714165	A	-0.2318	0.025199	3.62E-20	*C12orf50*	NA	NA	NA
12	chr12:7500717:T:C	7500717	T	0.029236	0.004986	4.52E-09	*CD163*	NA	NA	1
12	chr12:15673479:T:C	15673479	T	-0.15437	0.022127	3.03E-12	*EPS8*	NA	NA	NA
12	chr12:41859885:T:C	41859885	T	0.648633	0.114282	1.38E-08	*GXYLT1*	NA	NA	1
12	chr12:120978847:A:C	1.21E+08	A	-0.03665	0.004215	3.46E-18	*HNF1A*	NA	NA	1
12	chr12:48604724:A:G	48604724	A	-0.28161	0.037898	1.08E-13	*LALBA*	NA	NA	1
12	chr12:109549113:C:G	1.1E+08	C	-0.02497	0.004129	1.48E-09	*MMAB*	rs888194	1	1
12	chr12:44867039:C:T	44867039	C	-0.1645	0.027957	4.00E-09	*NELL2*	NA	NA	NA
12	chr12:20310515:C:T	20310515	C	-0.02523	0.004465	1.61E-08	*PDE3A*	NA	NA	1
12	chr12:124780655:G:A	1.25E+08	G	0.037813	0.004207	2.51E-19	*SCARB1*	rs921919	1	1
12	chr12:126562177:T:C	1.27E+08	T	-0.10135	0.016911	2.06E-09	*TMEM132B*	NA	NA	NA
13	chr13:18882182:C:A	18882182	C	-0.72576	0.073362	4.47E-23	*TUBA3C*	NA	NA	NA
13	chr13:51682578:C:T	51682578	C	-0.5345	0.096036	2.61E-08	*WDFY2*	NA	NA	NA
14	chr14:63134329:G:A	63134329	G	-0.12051	0.017491	5.60E-12	*KCNH5*	NA	NA	NA
14	chr14:75093953:T:C	75093953	T	-0.17319	0.019679	1.36E-18	*NEK9*	NA	NA	1
14	chr14:30960098:A:G	30960098	A	-0.1455	0.025746	1.59E-08	*STRN3*	NA	NA	NA
14	chr14:104801891:G:A	1.05E+08	G	0.027622	0.004223	6.09E-11	*ZBTB42*	rs4983387	1	1
15	chr15:58391167:A:G	58391167	A	0.073557	0.004237	1.63E-67	*ALDH1A2*	rs1532085	1	1
15	chr15:89642262:C:G	89642262	C	-0.94783	0.113092	5.24E-17	*KIF7*	NA	NA	1
15	chr15:58431740:G:A	58431740	G	-0.10748	0.004799	4.16E-111	*LIPC*	rs2070895	1	1
15	chr15:59228450:A:G	59228450	A	-0.05327	0.004368	3.28E-34	*MYO1E*	NA	NA	1
15	chr15:101527638:C:T	1.02E+08	C	-0.02663	0.004132	1.17E-10	*PCSK6*	NA	NA	1
16	chr16:65820986:T:C	65820986	T	-0.17633	0.022278	2.48E-15	*CDH5*	NA	NA	1

16	chr16:56973534:T:G	56973534	T	0.157607	0.004391	3.71E-282	*CETP*	rs11076176	1	1
16	chr16:81501185:T:C	81501185	T	-0.03394	0.00622	4.83E-08	*CMIP*	rs2925979	1	1
16	chr16:56959412:C:A	56959412	C	-0.17591	0.004248	0	*HERPUD1 - CETP*	rs3764261	1	1
16	chr16:35442020:G:T	35442020	G	-1.09252	0.076106	9.87E-47	*IGHV3OR16-17*	NA	NA	NA
16	chr16:67940360:G:T	67940360	G	0.456157	0.053001	7.53E-18	*LCAT*	NA	NA	1
16	chr16:67795860:TAAAC:T	67795860	TAAAC	-0.04956	0.005804	1.35E-17	*RANBP10*	NA	NA	1
16	chr16:19267303:G:A	19267303	G	0.037147	0.005669	5.66E-11	*SYT17*	NA	NA	1
16	chr16:56963437:C:CA	56963437	C	0.164857	0.004687	5.02E-271	*NA*	NA	NA	1
17	chr17:68876704:C:G	68876704	C	0.062814	0.011274	2.52E-08	*ABCA8*	NA	NA	1
17	chr17:43848758:C:T	43848758	C	0.210181	0.025093	5.47E-17	*CD300LG*	rs72836561	1	1
17	chr17:39727784:C:G	39727784	C	-0.02573	0.004157	5.99E-10	*ERBB2*	NA	NA	1
17	chr17:60094576:T:C	60094576	T	-0.21117	0.022	8.10E-22	*HEATR6*	NA	NA	1
17	chr17:17182089:C:T	17182089	C	-0.15789	0.025549	6.42E-10	*MPRIP*	NA	NA	1
17	chr17:7058953:G:A	7058953	G	-0.03181	0.004295	1.30E-13	*SLC16A11*	NA	NA	1
17	chr17:79952582:C:T	79952582	C	-0.20426	0.036594	2.38E-08	*TBC1D16*	NA	NA	NA
18	chr18:63195322:G:A	63195322	G	-0.22992	0.031213	1.76E-13	*BCL2*	NA	NA	1
18	chr18:6512865:G:A	6512865	G	-0.278	0.031712	1.84E-18	*L3MBTL4*	NA	NA	NA
18	chr18:49590804:T:C	49590804	T	-0.06817	0.00493	1.74E-43	*LIPG*	NA	NA	1
18	chr18:7432877:G:A	7432877	G	-0.20236	0.025902	5.61E-15	*PTPRM*	NA	NA	1
19	chr19:8364439:G:A	8364439	G	-0.19409	0.024195	1.04E-15	*ANGPTL4*	rs116843064	1	1
19	chr19:44945208:T:G	44945208	T	-0.02568	0.004079	3.05E-10	*APOC4-APOC2*	rs5167	1	1
19	chr19:44908783:C:T	44908783	C	0.396987	0.062731	2.48E-10	*APOE*	NA	NA	1
19	chr19:11235479:G:A	11235479	G	0.10363	0.004435	9.58E-121	*DOCK6*	rs56322906	1	1
19	chr19:54290350:C:T	54290350	C	0.036643	0.004232	4.80E-18	*LILRB2*	NA	NA	1
19	chr19:2269693:G:T	2269693	G	-0.28868	0.041674	4.30E-12	*OAZ1*	NA	NA	1
19	chr19:21315777:G:A	21315777	G	-0.15072	0.020565	2.32E-13	*ZNF708*	NA	NA	NA
20	chr20:50845709:G:A	50845709	G	-0.13234	0.019469	1.07E-11	*BCAS4*	NA	NA	NA
20	chr20:44413724:C:T	44413724	C	0.108542	0.009751	8.79E-29	*HNF4A*	rs1800961	1	1
20	chr20:62324457:C:A	62324457	C	-0.67231	0.072713	2.33E-20	*LAMA5*	NA	NA	NA
20	chr20:35001284:A:C	35001284	A	-0.15539	0.026059	2.47E-09	*MYH7B*	NA	NA	1
20	chr20:47719851:G:A	47719851	G	-0.03218	0.004621	3.32E-12	*SULF2*	NA	NA	1
21	chr21:41942186:C:T	41942186	C	-0.12118	0.018786	1.11E-10	*C2CD2*	NA	NA	1
21	chr21:32593327:G:A	32593327	G	-0.20088	0.02829	1.24E-12	*CFAP298-TCP10L*	NA	NA	NA

21	chr21:29631462:G:A	29631462	G	-0.90431	0.165177	4.38E-08	*GRIK1*	NA	NA	1
22	chr22:49458153:T:G	49458153	T	-0.12771	0.018093	1.68E-12	*BRD1*	NA	NA	NA
22	chr22:43945024:C:T	43945024	C	0.046007	0.004294	8.76E-27	*PNPLA3*	rs2294915	1	1
22	chr22:48386401:G:A	48386401	G	-0.25179	0.028915	3.10E-18	*TAFA5*	NA	NA	1
22	chr22:26576825:G:A	26576825	G	-0.66495	0.096358	5.17E-12	*TPST2*	NA	NA	NA
22	chr22:21622758:C:T	21622758	C	0.033311	0.004228	3.31E-15	*UBE2L3*	rs7445	1	1
23	chrX:122718884:A:C	1.23E+08	A	-0.92618	0.151958	1.09E-09	*GRIA3*	NA	NA	NA
23	chrX:145999162:G:A	1.46E+08	G	0.633761	0.115349	3.92E-08	*SLITRK2*	NA	NA	NA

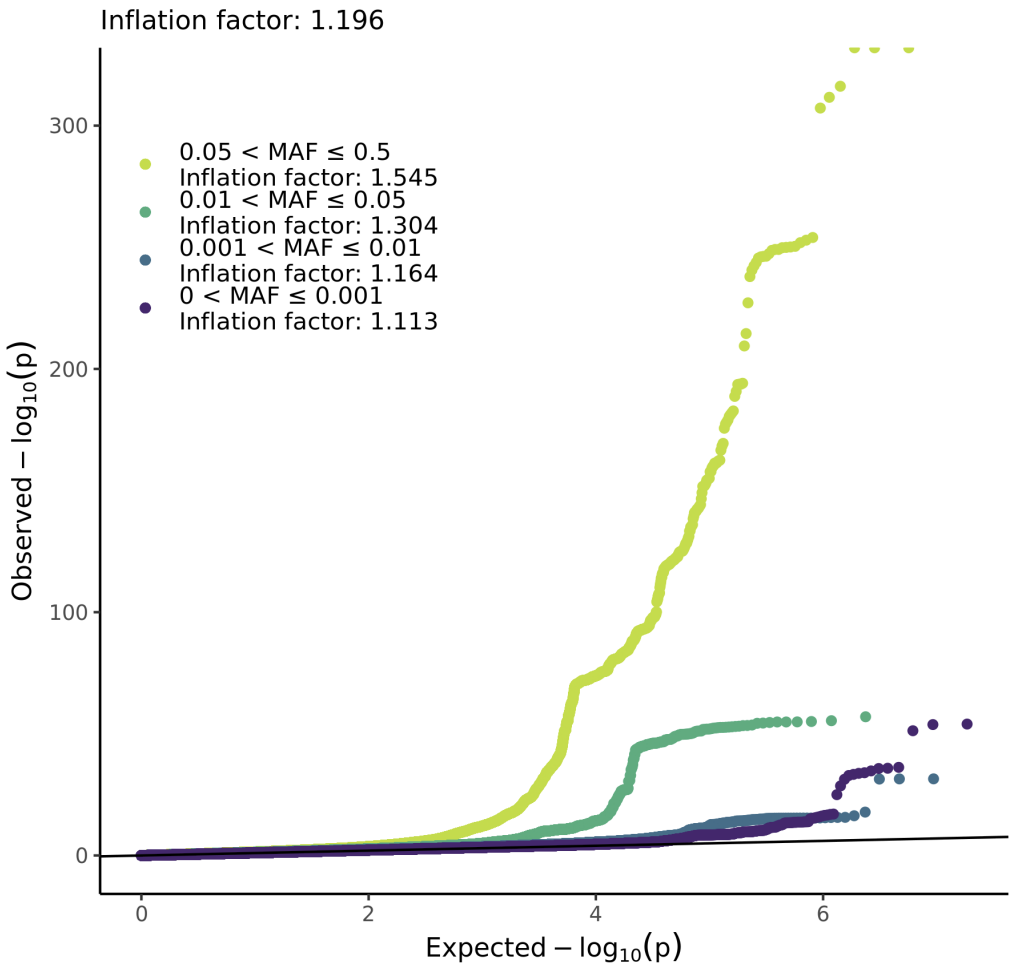
Beta, se and p-value reported were from conditional analysis for conditionally-independent signal. The most significant signal for each mapped gene was reported. SNP and PUBMED ID for reported variants were showed. BP: base position; Chr: chromosome; b, beta; se: standard error; NA: data not available; rsID: reference SNP (single nucleotide polymorphism) cluster ID.

Figure 5.5. Manhattan plot for the GWAS of NMR-measured TG



Manhattan plot shows genomic position across the 23 chromosomes on the X axis and -log10(p-value) on the y axis, with mapped genes for significant signals.

Figure 5.6. QQ plot for GWAS of NMR-measured TG



QQ plot split by the MAF of genetic variants. QQ: quantile-quantile; MAF: minor allele frequency.

Table 5.3. Conditionally independent signals from fine-mapped results of the TG GWAS

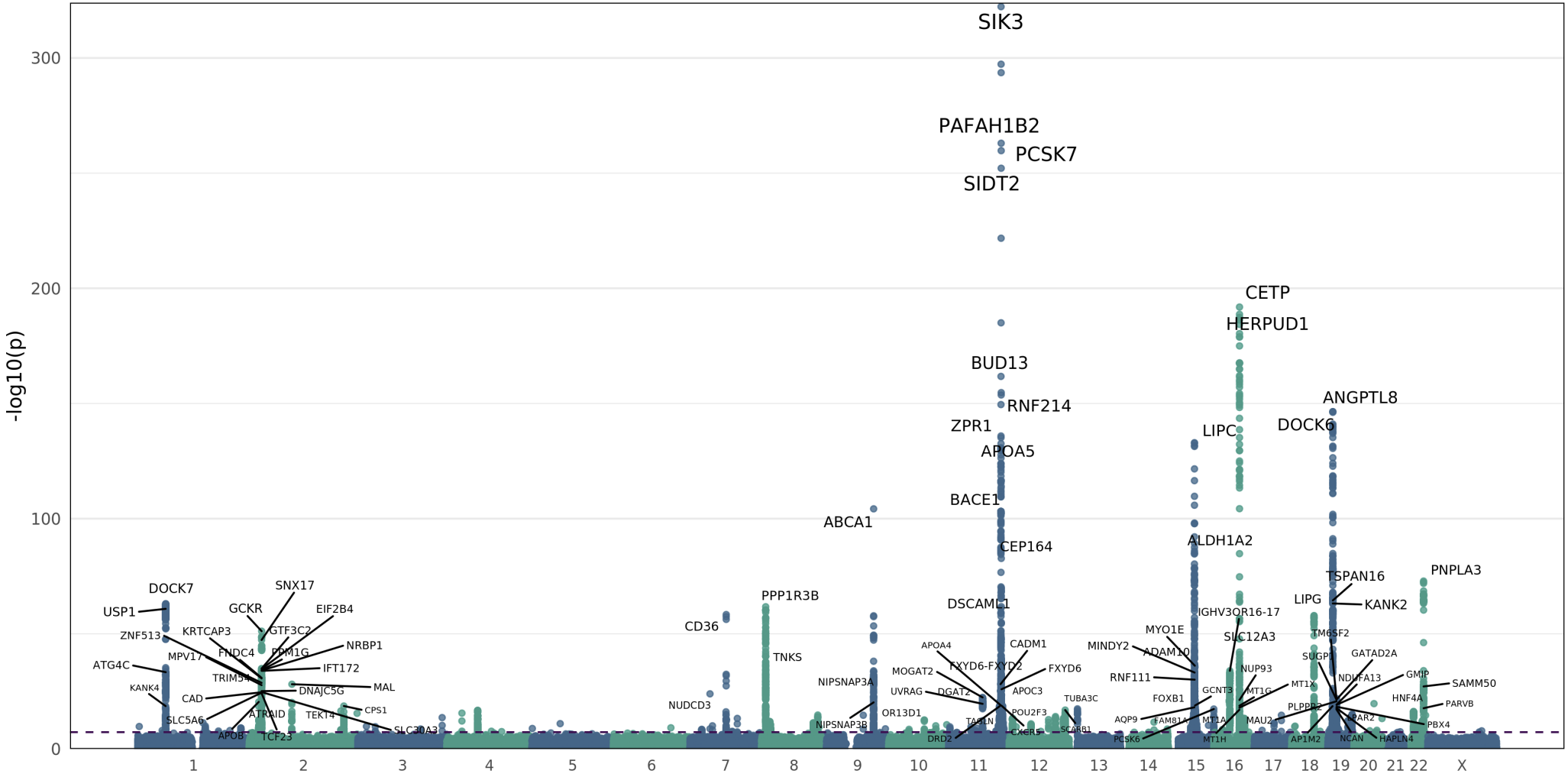
Chr	SNP	bp	ref	b	se	p	gene	rsID	Reported variant	Known locus
1	chr1:109274968:G:T	1.09E+08	G	0.038702	0.005369	5.68E-13	CELSR2	rs12740374	1	1
1	chr1:62620326:T:A	62620326	T	0.078427	0.004257	8.81E-76	DOCK7	NA	NA	1
1	chr1:3593253:C:T	3593253	C	-0.22828	0.037951	1.80E-09	MEGF6	NA	NA	NA
1	chr1:180182877:C:T	1.8E+08	C	-0.22797	0.04011	1.32E-08	QSOX1	NA	NA	NA
1	chr1:219508359:AT:A	2.2E+08	AT	-0.031	0.005358	7.17E-09	ZC3H11B	NA	NA	1
2	chr2:21014672:G:A	21014672	G	0.073316	0.004627	1.53E-56	APOB	rs673548	1	1
2	chr2:164645417:T:TG	1.65E+08	T	0.036115	0.005856	6.96E-10	COBLL1	NA	NA	1
2	chr2:27508073:T:C	27508073	T	0.102924	0.0045	8.73E-116	GCKR	rs1260326	1	1
2	chr2:94961415:C:T	94961415	C	-0.24759	0.026894	3.38E-20	MAL	NA	NA	NA
2	chr2:30193959:T:G	30193959	T	-0.24267	0.040384	1.86E-09	YPEL5	NA	NA	1
3	chr3:187669234:T:G	1.88E+08	T	-0.15665	0.028605	4.35E-08	SST	NA	NA	1
3	chr3:24478792:A:G	24478792	A	0.030689	0.004394	2.86E-12	THRB	rs6792725	1	1
3	chr3:86988393:A:G	86988393	A	0.07023	0.011377	6.69E-10	VGLL3	rs13066793	1	1
4	chr4:87088468:G:A	87088468	G	-0.03152	0.004471	1.80E-12	AFF1	NA	NA	1
4	chr4:33595172:C:T	33595172	C	-0.21581	0.026856	9.29E-16	ARAP2	NA	NA	NA
4	chr4:122131188:C:T	1.22E+08	C	-0.24924	0.036873	1.39E-11	KIAA1109	NA	NA	NA
4	chr4:3411224:C:T	3411224	C	0.028063	0.004309	7.36E-11	RGS12	NA	NA	1
4	chr4:68626601:A:G	68626601	A	-0.02901	0.00513	1.57E-08	UGT2B15	NA	NA	1
5	chr5:56512648:C:T	56512648	C	-0.03215	0.004853	3.45E-11	ANKRD55	rs465002	1	1
5	chr5:157045814:G:A	1.57E+08	G	-0.03464	0.004436	5.78E-15	HAVCR1	rs17054136	1	1
5	chr5:62615128:G:A	62615128	G	-0.25262	0.034017	1.12E-13	IPO11	NA	NA	NA
5	chr5:56572270:C:T	56572270	C	-0.03957	0.005077	6.48E-15	MAP3K1	NA	NA	1
5	chr5:68418419:A:G	68418419	A	-0.02532	0.004281	3.34E-09	PIK3R1	rs4976033	1	1
5	chr5:75741261:A:G	75741261	A	-0.02576	0.004301	2.10E-09	POC5	NA	NA	1
6	chr6:139519556:A:C	1.4E+08	A	0.023607	0.004145	1.23E-08	CITED2	NA	NA	1
6	chr6:160668275:C:T	1.61E+08	C	0.137117	0.008247	4.52E-62	LPA	NA	NA	1
6	chr6:160690668:C:T	1.61E+08	C	0.328709	0.037398	1.50E-18	LPA - PLG	rs186696265	1	1
6	chr6:34276468:C:A	34276468	C	0.046333	0.004469	3.51E-25	NUDT3	NA	NA	1
6	chr6:43790159:C:A	43790159	C	-0.0312	0.004318	4.95E-13	VEGFA - LINC02537	rs998584	1	1
7	chr7:80604075:C:T	80604075	C	0.10173	0.011458	6.76E-19	CD36	NA	NA	1
7	chr7:130739314:G:A	1.31E+08	G	0.033074	0.004308	1.63E-14	KLF14	NA	NA	1
7	chr7:73623626:A:G	73623626	A	0.107292	0.008842	6.95E-34	MLXIPL	rs34060476	1	1

7	chr7:44498864:T:C	44498864	T	-0.37244	0.044379	4.77E-17	NUDCD3	NA	NA	1
7	chr7:105394429:C:T	1.05E+08	C	-0.13332	0.023115	8.05E-09	SRPK2	NA	NA	NA
8	chr8:19969590:CTTG:C	19969590	CTTG	0.164944	0.01045	4.04E-56	LPL	NA	NA	1
8	chr8:128337012:G:A	1.28E+08	G	-0.14796	0.023498	3.04E-10	MYC	NA	NA	1
8	chr8:18425842:C:G	18425842	C	-0.04369	0.00434	7.75E-24	NAT2	NA	NA	1
8	chr8:9315699:G:A	9315699	G	0.024608	0.004224	5.71E-09	PPP1R3B-DT	rs7012637	1	1
8	chr8:14438847:G:A	14438847	G	0.215019	0.037004	6.22E-09	SGCZ	NA	NA	NA
8	chr8:125465736:G:C	1.25E+08	G	0.064305	0.004234	4.21E-52	TRIB1	rs2001945	1	1
8	chr8:125495066:T:C	1.25E+08	T	0.072425	0.005185	2.41E-44	TRIB1AL	rs2980888	1	1
9	chr9:133273680:C:CA	1.33E+08	C	0.051744	0.005548	1.09E-20	ABO	NA	NA	1
9	chr9:16848792:C:T	16848792	C	-0.03324	0.005429	9.13E-10	BNC2	NA	NA	1
9	chr9:6651319:C:T	6651319	C	-0.04214	0.006587	1.59E-10	GLDC	NA	NA	1
9	chr9:81781896:G:A	81781896	G	-0.44969	0.077177	5.65E-09	TLE1	NA	NA	NA
10	chr10:49617885:C:G	49617885	C	-0.02635	0.00459	9.37E-09	CHAT	NA	NA	NA
10	chr10:93079885:G:A	93079885	G	0.033295	0.004183	1.73E-15	CYP26A1	rs2068888	1	1
10	chr10:69216946:G:T	69216946	G	0.028372	0.004503	2.96E-10	HKDC1	NA	NA	1
10	chr10:29048722:A:G	29048722	A	0.362775	0.06361	1.18E-08	LYZL1	NA	NA	NA
10	chr10:129475679:G:A	1.29E+08	G	-0.1488	0.027208	4.53E-08	MGMT	NA	NA	1
10	chr10:63128096:C:T	63128096	C	0.040983	0.005403	3.32E-14	NRBF2	NA	NA	1
10	chr10:79473458:T:G	79473458	T	-0.4574	0.067922	1.65E-11	ZCCHC24	NA	NA	1
11	chr11:116792412:C:T	1.17E+08	C	-0.25936	0.057154	5.68E-06	APOA5	rs45611741	1	1
11	chr11:116830638:G:A	1.17E+08	G	1.03184	0.081812	1.80E-36	APOC3	rs138326449	1	1
11	chr11:68815246:T:C	68815246	T	0.030605	0.005178	3.41E-09	CPT1A	NA	NA	1
11	chr11:105614117:T:C	1.06E+08	T	0.030783	0.004992	7.00E-10	GRIA4	NA	NA	NA
11	chr11:2837210:G:A	2837210	G	0.032998	0.004506	2.42E-13	KCNQ1	NA	NA	1
11	chr11:78063901:G:A	78063901	G	-0.04504	0.007206	4.12E-10	THRSP	NA	NA	1
11	chr11:116778201:G:C	1.17E+08	G	0.201979	0.004337	0	ZPR1	rs964184	1	1
11	chr11:117179800:CT:C	1.17E+08	CT	1.2733	0.08188	1.57E-54	NA	NA	NA	1
13	chr13:113825176:T:G	1.14E+08	T	-0.03335	0.004545	2.15E-13	GAS6	NA	NA	1
13	chr13:94584571:A:G	94584571	A	-0.02694	0.004556	3.35E-09	TGDS	NA	NA	1
13	chr13:18920004:C:T	18920004	C	-0.65889	0.077759	2.38E-17	TUBA3C	NA	NA	NA
14	chr14:75093953:T:C	75093953	T	-0.12861	0.020115	1.62E-10	NEK9	NA	NA	1
15	chr15:58391167:A:G	58391167	A	0.040362	0.004319	9.15E-21	ALDH1A2	rs1532085	1	1
15	chr15:58430391:G:GC	58430391	G	-0.06114	0.004749	6.22E-38	LIPC	NA	NA	1

15	chr15:43464986:T:C	43464986	T	-0.16232	0.018648	3.20E-18	TP53BP1	rs190543502	1	1
16	chr16:35442020:G:T	35442020	G	-0.63228	0.076981	2.15E-16	IGHV3OR16-17	NA	NA	NA
17	chr17:7018644:G:A	7018644	G	0.028063	0.004309	7.35E-11	C17orf49	NA	NA	1
18	chr18:6512865:G:A	6512865	G	-0.26234	0.032436	6.07E-16	L3MBTL4	NA	NA	NA
18	chr18:49594129:T:G	49594129	T	-0.03338	0.005088	5.38E-11	LIPG	NA	NA	1
19	chr19:8364444:G:C	8364444	G	0.096891	0.010529	3.50E-20	ANGPTL4	NA	NA	1
19	chr19:44927023:C:G	44927023	C	-0.09393	0.004433	1.29E-99	APOC1	rs5112	1	1
19	chr19:11235479:G:A	11235479	G	0.039289	0.004539	4.85E-18	DOCK6	rs56322906	1	1
19	chr19:57175959:G:A	57175959	G	-0.41111	0.070808	6.40E-09	DUXA	NA	NA	1
19	chr19:19301236:G:A	19301236	G	0.152583	0.007376	4.56E-95	SUGP1	NA	NA	1
20	chr20:50845709:G:A	50845709	G	-0.11536	0.019895	6.70E-09	BCAS4	NA	NA	NA
20	chr20:17232761:C:G	17232761	C	-0.03706	0.006767	4.32E-08	PCSK2	NA	NA	1
22	chr22:33401376:C:T	33401376	C	-0.03712	0.006798	4.76E-08	LARGE1	NA	NA	1
22	chr22:43928847:C:G	43928847	C	0.062049	0.004394	2.79E-45	PNPLA3	rs738409	1	1
22	chr22:46219983:T:C	46219983	T	0.063477	0.008689	2.76E-13	PPARA	rs1800234	1	1
23	chrX:110440911:G:C	1.1E+08	G	0.044722	0.005883	2.92E-14	RTL9	NA	NA	NA

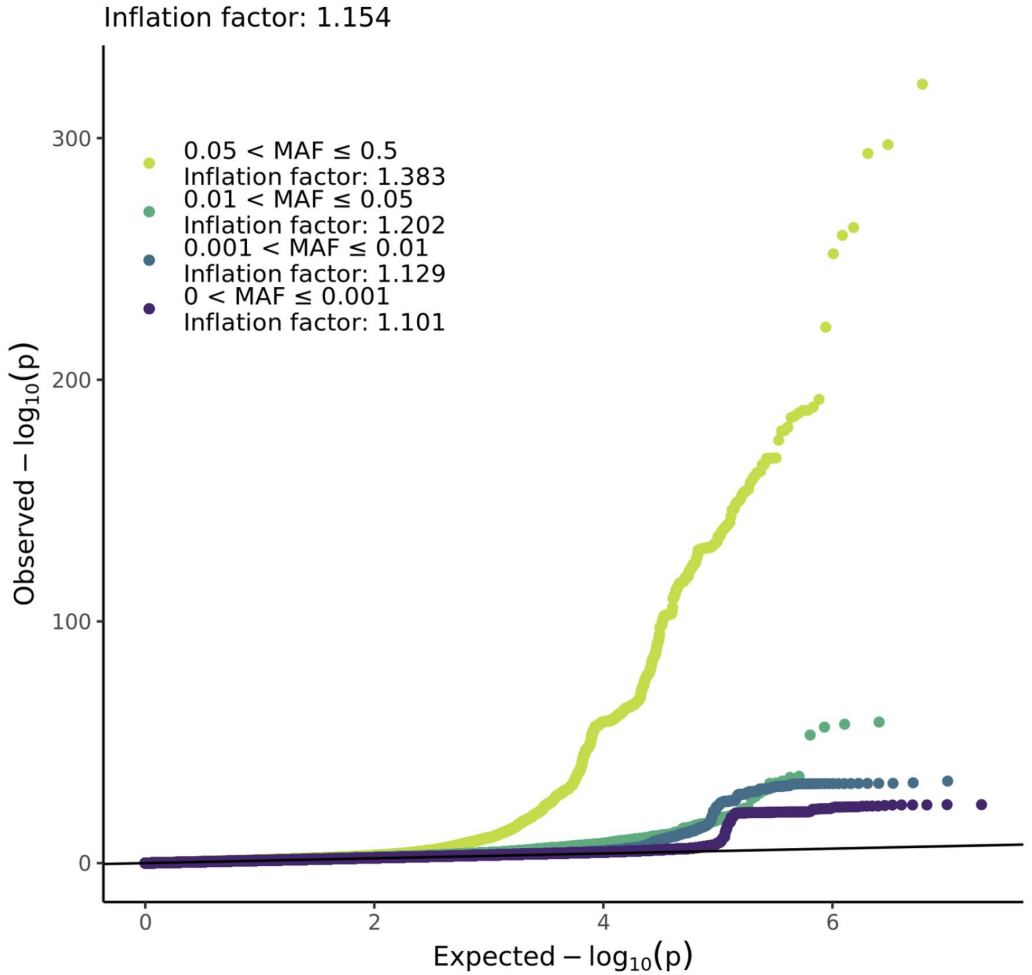
Beta, se and p-value reported were from conditional analysis for conditionally-independent signal. The most significant signal for each mapped gene was reported. SNP and PUBMED ID for reported variants were showed. BP: base position; Chr: chromosome; b, beta; se: standard error; NA: data not available; rsID: reference SNP (single nucleotide polymorphism) cluster ID.

Figure 5.7. Manhattan plot for the GWAS of NMR-measured ApoA1



Manhattan plot shows genomic position across the 23 chromosomes on the X axis and -log10(p-value) on the y axis, with mapped genes for significant signals.

Figure 5.8. QQ plot for GWAS of NMR-measured ApoA1



QQ plot split by the MAF of genetic variants. QQ: quantile-quantile; MAF: minor allele frequency.

Table 5.4. Conditionally independent signals from fine-mapped results of the ApoA1 GWAS

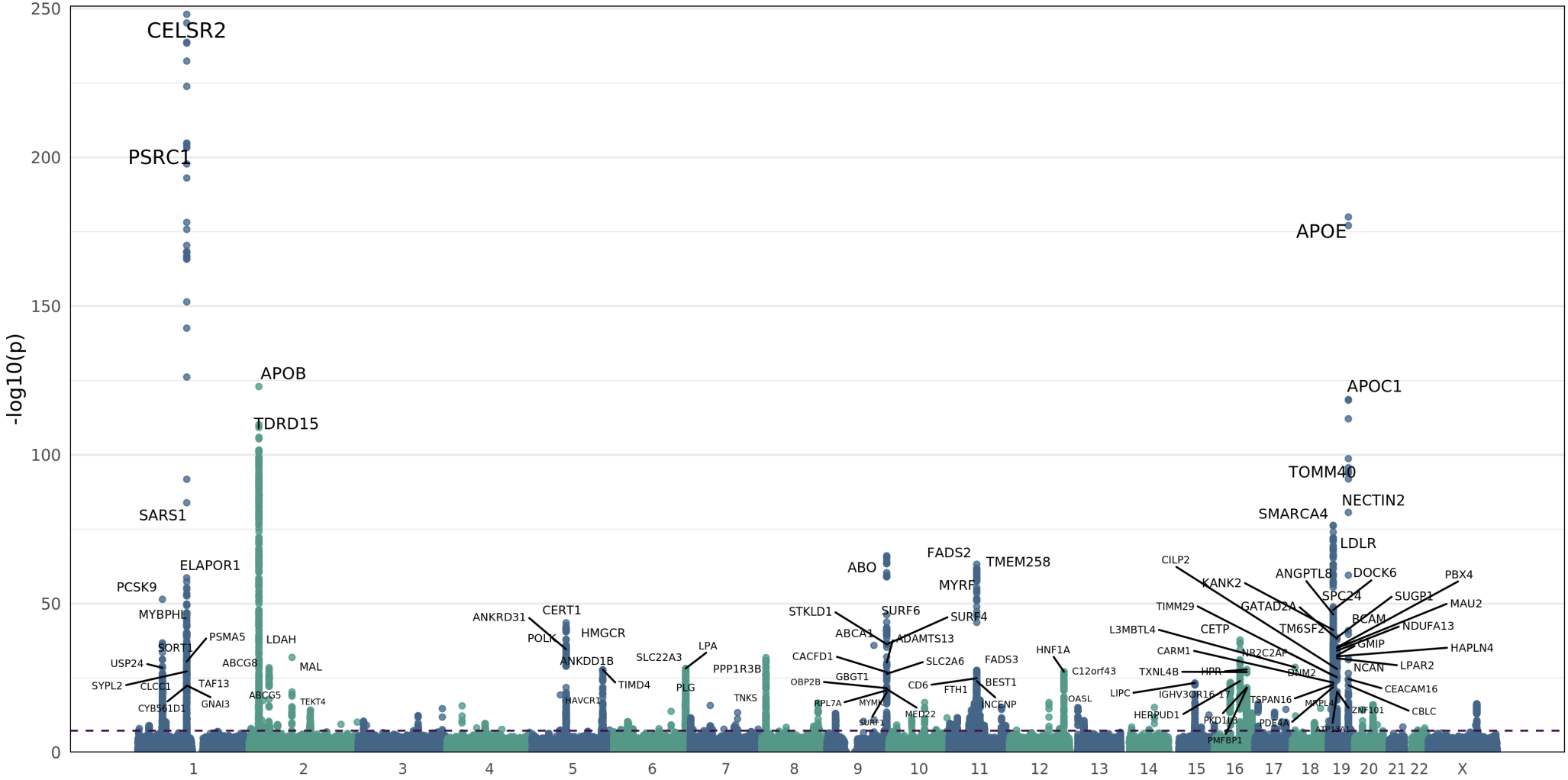
Chr	Variant ID	bp	ref	b	se	p	gene	rsID	Reported variant	Known locus
1	chr1:150870220:C:T	1.51E+08	C	-0.02864	0.004436	1.07E-10	*ARNT*	NA	NA	1
1	chr1:62620326:T:A	62620326	T	0.070564	0.004185	8.99E-64	*DOCK7*	NA	NA	1
1	chr1:205637044:A:G	2.06E+08	A	0.025413	0.004435	1.01E-08	*ELK4*	NA	NA	1
1	chr1:230159944:G:A	2.3E+08	G	-0.02539	0.004133	8.00E-10	*GALNT2*	NA	NA	1
1	chr1:145655702:C:T	1.46E+08	C	-0.14841	0.021421	4.26E-12	*GPR89A*	NA	NA	1
1	chr1:3593253:C:T	3593253	C	-0.23814	0.037311	1.74E-10	*MEGF6*	NA	NA	NA
1	chr1:54703968:T:C	54703968	T	-0.15165	0.026408	9.33E-09	*MROH7-TTC4*	NA	NA	1
1	chr1:180182877:C:T	1.8E+08	C	-0.22294	0.039433	1.57E-08	*QSOX1*	NA	NA	NA
2	chr2:21011127:T:C	21011127	T	0.072575	0.007686	3.65E-21	*APOB*	rs1801699	1	1
2	chr2:210675783:C:A	2.11E+08	C	0.040426	0.004492	2.25E-19	*CPS1*	rs1047891	1	1
2	chr2:27508073:T:C	27508073	T	0.066948	0.004418	7.35E-52	*GCKR*	rs1260326	1	1
2	chr2:204107315:C:T	2.04E+08	C	-0.60222	0.071356	3.18E-17	*ICOS*	NA	NA	NA
2	chr2:94961415:C:T	94961415	C	-0.29427	0.026407	7.71E-29	*MAL*	NA	NA	NA
2	chr2:240217492:G:A	2.4E+08	G	-0.25493	0.031235	3.31E-16	*OTOS*	NA	NA	NA
2	chr2:30193959:T:G	30193959	T	-0.32067	0.039653	6.12E-16	*YPEL5*	NA	NA	1
3	chr3:18939493:T:C	18939493	T	-0.52891	0.082573	1.50E-10	*KCNH8*	NA	NA	NA
3	chr3:139717549:C:A	1.4E+08	C	-0.82693	0.13807	2.11E-09	*NMNAT3*	NA	NA	NA
3	chr3:38690037:C:T	38690037	C	-0.26522	0.041757	2.13E-10	*SCN10A*	NA	NA	1
3	chr3:187669234:T:G	1.88E+08	T	-0.21477	0.028152	2.37E-14	*SST*	NA	NA	NA
4	chr4:99323162:G:A	99323162	G	-0.2579	0.045931	1.97E-08	*ADH1B*	NA	NA	1
4	chr4:33595172:C:T	33595172	C	-0.21621	0.026416	2.73E-16	*ARAP2*	NA	NA	NA
4	chr4:122131188:C:T	1.22E+08	C	-0.20125	0.036268	2.88E-08	*KIAA1109*	NA	NA	NA
4	chr4:68672015:T:G	68672015	T	-0.03731	0.004387	1.81E-17	*UGT2B15*	NA	NA	1
5	chr5:62615128:G:A	62615128	G	-0.22769	0.033444	9.90E-12	*IPO11*	NA	NA	NA
5	chr5:523946:C:T	523946	C	-0.87791	0.14455	1.25E-09	*SLC9A3*	NA	NA	NA
6	chr6:128048785:G:A	1.28E+08	G	-0.09964	0.016114	6.28E-10	*PTPRK*	NA	NA	NA
7	chr7:94170434:GT:G	94170434	GT	0.126006	0.022485	2.10E-08	*BET1*	NA	NA	NA
7	chr7:80664401:T:C	80664401	T	-0.18708	0.011542	4.35E-59	*CD36*	NA	NA	1
7	chr7:26076003:C:T	26076003	C	-0.03101	0.005229	3.03E-09	*NFE2L3*	NA	NA	1
7	chr7:44498864:T:C	44498864	T	-0.44802	0.043698	1.15E-24	*NUDCD3*	NA	NA	1
7	chr7:105394429:C:T	1.05E+08	C	-0.17125	0.022761	5.32E-14	*SRPK2*	NA	NA	NA
8	chr8:19966981:T:C	19966981	T	-0.03398	0.005072	2.07E-11	*LPL*	rs13702	1	1

8	chr8:128337012:G:A	1.28E+08	G	-0.14518	0.023092	3.24E-10	*MYC*	NA	NA	NA
8	chr8:125266703:G:A	1.25E+08	G	0.611039	0.100141	1.05E-09	*NSMCE2*	NA	NA	1
8	chr8:9323885:A:G	9323885	A	-0.0745	0.004466	1.79E-62	*PPP1R3B*	NA	NA	1
8	chr8:125479491:G:T	1.25E+08	G	0.034642	0.004356	1.83E-15	*TRIB1*	NA	NA	1
9	chr9:104858554:G:A	1.05E+08	G	0.143508	0.006595	5.53E-105	*ABCA1*	NA	NA	1
9	chr9:131057085:G:A	1.31E+08	G	-0.30159	0.05011	1.76E-09	*LAMC3*	NA	NA	NA
9	chr9:81781896:G:A	81781896	G	-0.60333	0.076134	2.29E-15	*TLE1*	NA	NA	NA
10	chr10:45443688:T:C	45443688	T	-0.03424	0.00577	2.96E-09	*ALOX5*	NA	NA	1
10	chr10:129475679:G:A	1.29E+08	G	-0.1781	0.026773	2.89E-11	*MGMT*	NA	NA	NA
10	chr10:105053006:G:A	1.05E+08	G	-0.22343	0.034435	8.66E-11	*SORCS3*	NA	NA	NA
10	chr10:79473458:T:G	79473458	T	-0.48874	0.066837	2.62E-13	*ZCCHC24*	NA	NA	1
11	chr11:116830638:G:A	1.17E+08	G	-0.60077	0.080138	6.54E-14	*APOC3*	rs138326449	1	1
11	chr11:105536945:T:G	1.06E+08	T	-0.0257	0.004535	1.45E-08	*CARD18*	NA	NA	NA
11	chr11:75745089:C:T	75745089	C	0.060315	0.006092	4.15E-23	*MOGAT2*	NA	NA	1
11	chr11:49900650:C:T	49900650	C	-0.11464	0.018101	2.40E-10	*OR4C13*	NA	NA	1
11	chr11:63490651:C:T	63490651	C	-0.1048	0.01717	1.04E-09	*PLAAT5*	NA	NA	NA
11	chr11:116869766:C:T	1.17E+08	C	-0.2496	0.006496	0	*SIK3*	NA	NA	1
11	chr11:2900598:CAGG:C	2900598	CAGG	0.038222	0.006035	2.39E-10	*SLC22A18*	NA	NA	1
12	chr12:103129174:C:T	1.03E+08	C	0.032974	0.004247	8.27E-15	*C12orf42*	NA	NA	1
12	chr12:87714165:A:G	87714165	A	-0.18682	0.025366	1.77E-13	*C12orf50*	NA	NA	NA
12	chr12:7490731:G:A	7490731	G	0.037919	0.005031	4.78E-14	*CD163*	NA	NA	1
12	chr12:15673479:T:C	15673479	T	-0.1413	0.022274	2.25E-10	*EPS8*	NA	NA	NA
12	chr12:120978847:A:C	1.21E+08	A	-0.03203	0.004243	4.36E-14	*HNF1A*	NA	NA	1
12	chr12:48592500:C:T	48592500	C	-0.22636	0.03337	1.17E-11	*LALBA*	NA	NA	NA
12	chr12:124779741:T:G	1.25E+08	T	0.03682	0.004309	1.28E-17	*SCARB1*	NA	NA	1
13	chr13:18882182:C:A	18882182	C	-0.64129	0.073837	3.78E-18	*TUBA3C*	NA	NA	NA
14	chr14:75093953:T:C	75093953	T	-0.139	0.019807	2.26E-12	*NEK9*	NA	NA	1
14	chr14:104801891:G:A	1.05E+08	G	0.025417	0.00425	2.23E-09	*ZBTB42*	NA	NA	1
15	chr15:89642262:C:G	89642262	C	-0.74837	0.114009	5.24E-11	*KIF7*	NA	NA	1
15	chr15:58431740:G:A	58431740	G	-0.11904	0.004838	1.07E-133	*LIPC*	NA	NA	1
15	chr15:101527638:C:T	1.02E+08	C	-0.03609	0.004166	4.55E-18	*PCSK6*	NA	NA	1
16	chr16:67645445:C:T	67645445	C	-0.04281	0.005689	5.29E-14	*CARMIL2*	NA	NA	1
16	chr16:65820986:T:C	65820986	T	-0.14964	0.022298	1.93E-11	*CDH5*	NA	NA	NA
16	chr16:56959412:C:A	56959412	C	-0.12586	0.004252	1.37E-192	*CETP*	NA	NA	1

16	chr16:35442020:G:T	35442020	G	-0.93553	0.076175	1.14E-34	*IGHV3OR16-17*	NA	NA	NA
16	chr16:67940360:G:T	67940360	G	0.387731	0.053049	2.69E-13	*LCAT*	NA	NA	1
16	chr16:19267303:G:A	19267303	G	0.033488	0.005674	3.60E-09	*SYT17*	NA	NA	NA
16	chr16:56973534:T:G	56973534	T	0.118097	0.004395	4.68E-159	*NA*	NA	NA	1
17	chr17:43848758:C:T	43848758	C	0.174114	0.025249	5.35E-12	*CD300LG*	rs72836561	1	1
17	chr17:60094576:T:C	60094576	T	-0.17549	0.022137	2.23E-15	*HEATR6*	NA	NA	NA
17	chr17:17182089:C:T	17182089	C	-0.14816	0.025708	8.25E-09	*MPRIP*	NA	NA	1
18	chr18:63195322:G:A	63195322	G	-0.17639	0.031426	1.99E-08	*BCL2*	NA	NA	NA
18	chr18:49594129:T:G	49594129	T	-0.08086	0.005009	1.25E-58	*LIPG*	NA	NA	1
18	chr18:7432877:G:A	7432877	G	-0.16782	0.02608	1.24E-10	*PTPRM*	NA	NA	NA
19	chr19:11239812:C:T	11239812	C	0.114961	0.00445	3.49E-147	*ANGPTL8*	rs2278426	1	1
19	chr19:44945208:T:G	44945208	T	-0.02573	0.004111	3.88E-10	*APOC4-APOC2*	rs5167	1	1
19	chr19:44909976:G:T	44909976	G	-0.10464	0.014481	4.98E-13	*APOE*	rs1065853	1	1
19	chr19:54290350:C:T	54290350	C	0.03438	0.004265	7.59E-16	*LILRB2*	NA	NA	1
19	chr19:2269693:G:T	2269693	G	-0.23402	0.041999	2.52E-08	*OAZ1*	NA	NA	NA
19	chr19:19271131:G:C	19271131	G	0.071576	0.007309	1.21E-22	*TM6SF2*	NA	NA	1
19	chr19:21315777:G:A	21315777	G	-0.12912	0.020726	4.67E-10	*ZNF708*	NA	NA	NA
20	chr20:50845709:G:A	50845709	G	-0.11326	0.01959	7.41E-09	*BCAS4*	NA	NA	NA
20	chr20:44413724:C:T	44413724	C	0.090836	0.009811	2.08E-20	*HNF4A*	rs1800961	1	1
20	chr20:62324457:C:A	62324457	C	-0.54913	0.073166	6.13E-14	*LAMA5*	NA	NA	NA
20	chr20:35001284:A:C	35001284	A	-0.14938	0.026221	1.22E-08	*MYH7B*	NA	NA	1
21	chr21:38466134:T:C	38466134	T	-0.03471	0.006282	3.30E-08	*ERG*	NA	NA	NA
22	chr22:43928847:C:G	43928847	C	0.078577	0.004329	1.27E-73	*PNPLA3*	rs738409	1	1
22	chr22:31737857:A:G	31737857	A	0.031595	0.005075	4.80E-10	*PRR14L*	NA	NA	NA
22	chr22:48386401:G:A	48386401	G	-0.23065	0.029129	2.41E-15	*TAFA5*	NA	NA	NA
22	chr22:26576825:G:A	26576825	G	-0.58802	0.097069	1.38E-09	*TPST2*	NA	NA	NA
22	chr22:21622758:C:T	21622758	C	0.035664	0.004259	5.60E-17	*UBE2L3*	NA	NA	1
23	chrX:117086716:C:T	1.17E+08	C	0.745793	0.136237	4.39E-08	*CT83*	NA	NA	NA
23	chrX:122718884:A:C	1.23E+08	A	-0.86941	0.152942	1.31E-08	*GRIA3*	NA	NA	NA

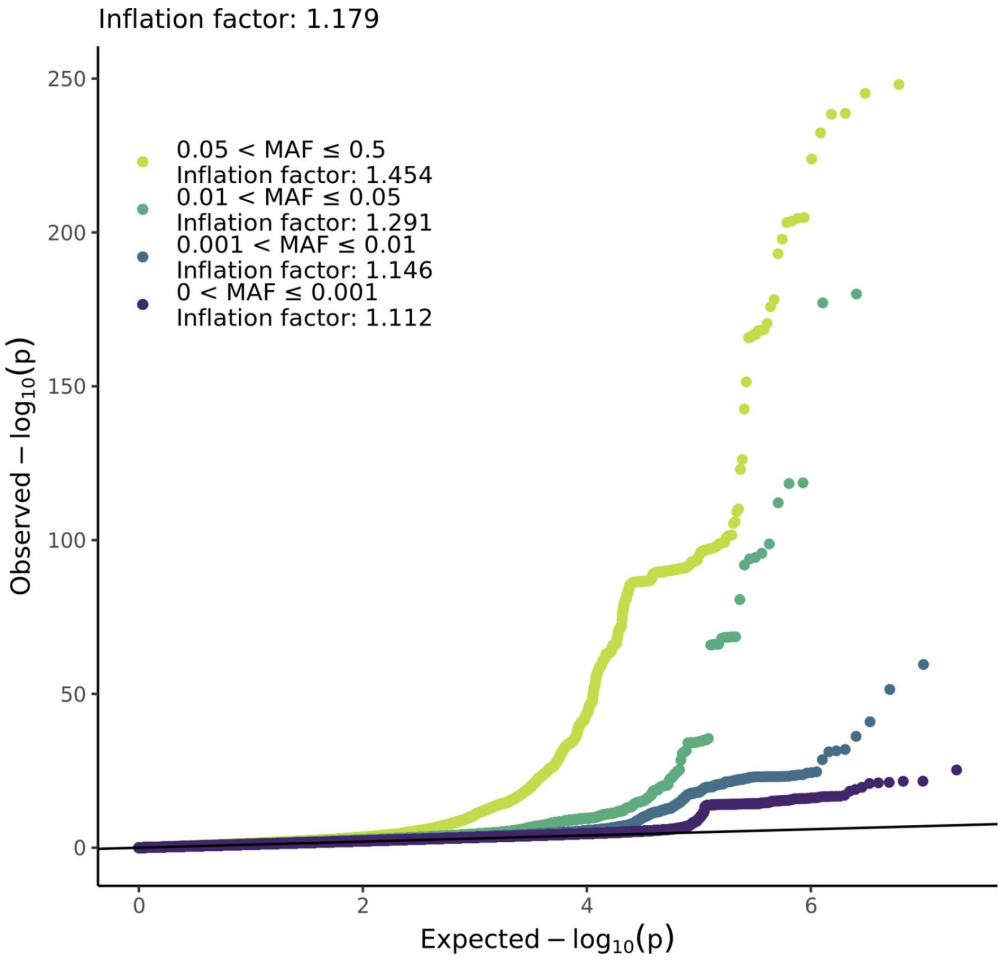
Beta, se and p-value reported were from conditional analysis for conditionally-independent signal. The most significant signal for each mapped gene was reported. SNP and PUBMED ID for reported variants were showed. BP: base position; Chr: chromosome; b, beta; se: standard error; NA: data not available; rsID: reference SNP (single nucleotide polymorphism) cluster ID.

Figure 5.9. Manhattan plot for the GWAS of NMR-measured ApoB



Manhattan plot shows genomic position across the 23 chromosomes on the X axis and -log10(p-value) on the y axis, with mapped genes for significant signals.

Figure 5.10. QQ plot for GWAS of NMR-measured ApoB



QQ plot split by the MAF of genetic variants. QQ: quantile-quantile; MAF: minor allele frequency.

Table 5.5. Conditionally independent signals from fine-mapped results of the ApoB GWAS

Chr	Variant ID	bp	ref	b	se	p	gene	rsID	Reported variant	Known locus
1	chr1:55021673:C:G	55021673	C	0.058148	0.006256	1.48E-20	*BSND*	NA	NA	1
1	chr1:109274968:G:T	1.09E+08	G	0.181796	0.005396	8.18E-249	*CELSR2*	rs12740374	1	1
1	chr1:3593253:C:T	3593253	C	-0.21938	0.038141	8.83E-09	*MEGF6*	NA	NA	NA
1	chr1:55039974:G:T	55039974	G	0.498422	0.032791	3.55E-52	*PCSK9*	rs11591147	1	1
2	chr2:43847292:C:T	43847292	C	0.076996	0.006863	3.30E-29	*ABCG8*	NA	NA	1
2	chr2:21043902:GGCAGCGCCA:G	21043902	GGCAGCGCCA	-0.10275	0.004344	1.09E-123	*APOB*	NA	NA	1
2	chr2:3600300:G:A	3600300	G	0.02615	0.004554	9.34E-09	*COLEC11*	NA	NA	NA
2	chr2:135950412:T:C	1.36E+08	T	-0.05129	0.006569	5.80E-15	*DARS1*	NA	NA	1
2	chr2:62642879:CA:C	62642879	CA	-0.02996	0.004807	4.62E-10	*EHBP1*	NA	NA	1
2	chr2:27508073:T:C	27508073	T	0.030425	0.004514	1.58E-11	*GCKR*	rs1260326	1	1
2	chr2:204107315:C:T	2.04E+08	C	-0.45688	0.072896	3.67E-10	*ICOS*	NA	NA	NA
2	chr2:94961415:C:T	94961415	C	-0.32109	0.026977	1.15E-32	*MAL*	NA	NA	NA
2	chr2:240217492:G:A	2.4E+08	G	-0.20839	0.031909	6.54E-11	*OTOS*	NA	NA	NA
2	chr2:30193959:T:G	30193959	T	-0.35177	0.040509	3.83E-18	*YPEL5*	NA	NA	NA
2	chr2:21070463:T:TAG	21070463	T	-0.15282	0.007132	7.26E-102	*NA*	NA	NA	1
3	chr3:18939493:T:C	18939493	T	-0.51646	0.084037	7.97E-10	*KCNH8*	NA	NA	NA
3	chr3:187669234:T:G	1.88E+08	T	-0.22757	0.028652	1.98E-15	*SST*	NA	NA	NA
3	chr3:12226148:A:G	12226148	A	0.036938	0.005555	2.93E-11	*SYN2*	NA	NA	1
3	chr3:133759510:G:A	1.34E+08	G	-0.05532	0.007643	4.56E-13	*TF*	NA	NA	NA
4	chr4:33595172:C:T	33595172	C	-0.22062	0.026873	2.22E-16	*ARAP2*	NA	NA	NA
4	chr4:66149466:G:A	66149466	G	-0.21284	0.036545	5.74E-09	*EPHA5*	NA	NA	NA
4	chr4:122131188:C:T	1.22E+08	C	-0.20759	0.036896	1.84E-08	*KIAA1109*	NA	NA	NA
4	chr4:84799165:T:G	84799165	T	0.071762	0.01124	1.72E-10	*WDFY3*	NA	NA	NA
5	chr5:75360714:T:C	75360714	T	-0.05839	0.004182	2.67E-44	*CERT1*	rs12916	1	1
5	chr5:62615128:G:A	62615128	G	-0.31252	0.03407	4.61E-20	*IPO11*	NA	NA	NA
5	chr5:172148664:A:T	1.72E+08	A	0.48668	0.087691	2.86E-08	*STK10*	NA	NA	NA
5	chr5:156968992:C:T	1.57E+08	C	-0.06607	0.005964	1.60E-28	*TIMD4*	NA	NA	1
6	chr6:42957217:T:G	42957217	T	0.028511	0.005016	1.31E-08	*GNMT*	NA	NA	1
6	chr6:101016200:G:A	1.01E+08	G	-0.02399	0.004199	1.11E-08	*GRIK2*	NA	NA	NA
6	chr6:31253574:T:C	31253574	T	-0.06423	0.009763	4.75E-11	*HLA-C*	NA	NA	1
6	chr6:160585045:C:T	1.61E+08	C	-0.06919	0.006196	5.93E-29	*LPA*	NA	NA	1
6	chr6:16126703:G:A	16126703	G	0.027916	0.004967	1.90E-08	*MYLIP*	rs7746081	1	1

6	chr6:128048785:G:A	1.28E+08	G	-0.12601	0.016412	1.62E-14	*PTPRK*	NA	NA	NA
7	chr7:98357118:T:C	98357118	T	-0.03128	0.005223	2.10E-09	*BAIAP2L1*	NA	NA	1
7	chr7:75008657:C:T	75008657	C	0.68854	0.121788	1.57E-08	*CASTOR2*	NA	NA	NA
7	chr7:983981:C:A	983981	C	0.03072	0.004383	2.39E-12	*CYP2W1*	NA	NA	NA
7	chr7:44574799:C:G	44574799	C	-0.03881	0.006367	1.09E-09	*DDX56*	NA	NA	1
7	chr7:100814752:T:C	1.01E+08	T	-0.02669	0.004284	4.69E-10	*EPHB4*	NA	NA	1
7	chr7:151327429:C:A	1.51E+08	C	-0.10917	0.019279	1.49E-08	*NUB1*	NA	NA	NA
7	chr7:44498864:T:C	44498864	T	-0.36631	0.044437	1.68E-16	*NUDCD3*	NA	NA	1
7	chr7:105394429:C:T	1.05E+08	C	-0.17478	0.023145	4.30E-14	*SRPK2*	NA	NA	NA
7	chr7:6581140:T:C	6581140	T	0.262827	0.046801	1.96E-08	*ZDHHC4*	NA	NA	NA
8	chr8:128337012:G:A	1.28E+08	G	-0.13369	0.023469	1.23E-08	*MYC*	NA	NA	NA
8	chr8:9323885:A:G	9323885	A	-0.05397	0.004539	1.32E-32	*PPP1R3B*	NA	NA	1
8	chr8:54524077:C:A	54524077	C	-0.04279	0.007249	3.58E-09	*RP1*	NA	NA	1
8	chr8:125475993:C:T	1.25E+08	C	0.037178	0.004396	2.76E-17	*TRIB1*	NA	NA	1
8	chr8:115645684:T:G	1.16E+08	T	0.026007	0.004301	1.48E-09	*TRPS1*	rs2721961	1	1
9	chr9:104858554:G:A	1.05E+08	G	0.084779	0.006701	1.09E-36	*ABCA1*	NA	NA	1
9	chr9:133273983:A:G	1.33E+08	A	0.120242	0.006963	8.12E-67	*ABO*	rs507666	1	1
9	chr9:19225102:G:C	19225102	G	0.041268	0.005521	7.71E-14	*DENND4C*	NA	NA	NA
9	chr9:128009032:G:A	1.28E+08	G	-0.04195	0.007674	4.59E-08	*FAM102A*	NA	NA	NA
9	chr9:131057085:G:A	1.31E+08	G	-0.32264	0.050914	2.34E-10	*LAMC3*	NA	NA	NA
9	chr9:81781896:G:A	81781896	G	-0.51166	0.077355	3.73E-11	*TLE1*	NA	NA	NA
10	chr10:50814156:T:C	50814156	T	-0.07552	0.010737	2.00E-12	*ASAH2B*	NA	NA	1
10	chr10:129475679:G:A	1.29E+08	G	-0.19066	0.027226	2.51E-12	*MGMT*	NA	NA	NA
10	chr10:105314447:G:A	1.05E+08	G	-0.44535	0.079434	2.06E-08	*SORCS3*	NA	NA	NA
10	chr10:17213096:G:A	17213096	G	-0.02926	0.004724	5.91E-10	*TRDMT1*	NA	NA	NA
10	chr10:79473458:T:G	79473458	T	-0.57857	0.067968	1.70E-17	*ZCCHC24*	NA	NA	NA
11	chr11:46515498:A:C	46515498	A	0.035125	0.00497	1.59E-12	*AMBRA1*	NA	NA	NA
11	chr11:61820833:A:G	61820833	A	0.08732	0.005171	5.61E-64	*FADS2*	rs174564	1	1
11	chr11:78417001:C:G	78417001	C	0.026319	0.004282	7.90E-10	*GAB2*	NA	NA	NA
11	chr11:123298810:G:A	1.23E+08	G	-0.08854	0.015296	7.10E-09	*GRAMD1B*	NA	NA	NA
11	chr11:2818521:C:T	2818521	C	0.027888	0.004416	2.71E-10	*KCNQ1*	NA	NA	NA
11	chr11:1863112:G:A	1863112	G	0.028534	0.004462	1.61E-10	*LSP1*	NA	NA	NA
11	chr11:50097023:A:G	50097023	A	-0.13521	0.017252	4.58E-15	*OR4C12*	NA	NA	NA
11	chr11:49900650:C:T	49900650	C	-0.14016	0.018341	2.13E-14	*OR4C13*	NA	NA	NA

11	chr11:63490651:C:T	63490651	C	-0.14052	0.017397	6.64E-16	*PLAAT5*	NA	NA	NA
11	chr11:116869766:C:T	1.17E+08	C	0.053126	0.006582	6.94E-16	*SIK3*	NA	NA	1
11	chr11:18624296:C:A	18624296	C	-0.02981	0.004252	2.36E-12	*SPTY2D1*	NA	NA	1
11	chr11:126367386:G:A	1.26E+08	G	-0.0329	0.005363	8.56E-10	*ST3GAL4*	NA	NA	1
11	chr11:5679844:C:A	5679844	C	-0.03968	0.006362	4.48E-10	*TRIM5*	rs11601507	1	1
11	chr11:122717441:G:A	1.23E+08	G	0.025813	0.004288	1.75E-09	*UBASH3B*	NA	NA	NA
12	chr12:111466567:T:A	1.11E+08	T	-0.03311	0.005793	1.09E-08	*ATXN2*	rs4766578	1	1
12	chr12:87714165:A:G	87714165	A	-0.21959	0.025802	1.73E-17	*C12orf50*	NA	NA	NA
12	chr12:15673479:T:C	15673479	T	-0.12528	0.022657	3.21E-08	*EPS8*	NA	NA	NA
12	chr12:120985583:A:G	1.21E+08	A	0.047571	0.004349	7.57E-28	*HNF1A*	NA	NA	1
12	chr12:115091610:T:G	1.15E+08	T	0.026273	0.00475	3.17E-08	*TBX3*	NA	NA	NA
13	chr13:32412082:A:G	32412082	A	-0.03158	0.004706	1.95E-11	*N4BP2L1*	rs206319	1	1
13	chr13:18926972:T:TAA	18926972	T	-0.55859	0.06935	7.97E-16	*TUBA3C*	NA	NA	NA
14	chr14:63134329:G:A	63134329	G	-0.10072	0.017904	1.85E-08	*KCNH5*	NA	NA	NA
14	chr14:75093953:T:C	75093953	T	-0.16235	0.020144	7.65E-16	*NEK9*	NA	NA	NA
14	chr14:24402720:C:T	24402720	C	-0.03508	0.005902	2.79E-09	*NYNRIN*	rs11621792	1	1
15	chr15:72690632:A:G	72690632	A	-0.03161	0.005321	2.83E-09	*BBS4*	NA	NA	NA
15	chr15:89642262:C:G	89642262	C	-0.84412	0.115401	2.58E-13	*KIF7*	NA	NA	NA
15	chr15:58430391:G:GC	58430391	G	-0.04812	0.004754	4.38E-24	*LIPC*	NA	NA	1
15	chr15:101527397:A:G	1.02E+08	A	0.032468	0.005259	6.67E-10	*PCSK6*	NA	NA	1
16	chr16:56983407:G:A	56983407	G	0.082148	0.006329	1.59E-38	*CETP*	rs1800777	1	1
16	chr16:56955678:C:T	56955678	C	-0.04008	0.004324	1.86E-20	*HERPUD1 - CETP*	rs247616	1	1
16	chr16:72072066:T:C	72072066	T	0.047525	0.004283	1.30E-28	*HPR*	NA	NA	1
16	chr16:35496176:C:T	35496176	C	-0.83787	0.082567	3.39E-24	*IGHV3OR16-17*	NA	NA	NA
16	chr16:83945717:A:G	83945717	A	0.033156	0.004416	5.96E-14	*OSGIN1*	NA	NA	1
17	chr17:69086821:C:G	69086821	C	0.04344	0.005516	3.41E-15	*ABCA6*	NA	NA	1
17	chr17:66214462:A:C	66214462	A	-0.09851	0.015945	6.49E-10	*APOH*	rs1801689	1	1
17	chr17:7176750:A:G	7176750	A	0.037272	0.004421	3.41E-17	*ASGR1*	NA	NA	1
17	chr17:43848758:C:T	43848758	C	0.195148	0.025669	2.91E-14	*CD300LG*	rs72836561	1	1
17	chr17:60094576:T:C	60094576	T	-0.14792	0.022505	4.95E-11	*HEATR6*	NA	NA	NA
17	chr17:47658340:T:C	47658340	T	0.025247	0.004551	2.90E-08	*KPNB1*	NA	NA	1
18	chr18:63195322:G:A	63195322	G	-0.25499	0.031975	1.53E-15	*BCL2*	NA	NA	NA
18	chr18:6512865:G:A	6512865	G	-0.36534	0.032487	2.43E-29	*L3MBTL4*	NA	NA	NA
18	chr18:49594230:G:A	49594230	G	-0.03268	0.005031	8.25E-11	*LIPG*	NA	NA	1

19	chr19:44908822:C:T	44908822	C	0.42921	0.014974	1.07E-180	*APOE*	rs7412	1	1
19	chr19:11236981:C:T	11236981	C	0.066948	0.004568	1.26E-48	*DOCK6*	NA	NA	1
19	chr19:5840608:C:T	5840608	C	-0.03342	0.004236	3.03E-15	*FUT6*	NA	NA	NA
19	chr19:11120527:G:A	11120527	G	0.337175	0.039125	6.82E-18	*LDLR*	rs72658867	1	1
19	chr19:1115299:G:A	1115299	G	0.036765	0.005531	2.98E-11	*SBNO2*	NA	NA	NA
19	chr19:11081874:G:A	11081874	G	0.100282	0.005397	4.70E-77	*SMARCA4*	NA	NA	1
19	chr19:11086585:G:A	11086585	G	0.172394	0.009595	3.58E-72	*SMARCA4 - LDLR*	rs12151108	1	1
19	chr19:19301236:G:A	19301236	G	0.097419	0.007424	2.49E-39	*SUGP1*	NA	NA	1
20	chr20:50845709:G:A	50845709	G	-0.12365	0.019931	5.50E-10	*BCAS4*	NA	NA	NA
20	chr20:44413724:C:T	44413724	C	0.054679	0.009982	4.31E-08	*HNF4A*	rs1800961	1	1
20	chr20:40555894:C:CA	40555894	C	-0.02883	0.004762	1.41E-09	*MAFB*	NA	NA	1
20	chr20:17616086:C:T	17616086	C	0.027141	0.004511	1.78E-09	*RRBP1*	NA	NA	1
20	chr20:17863848:G:T	17863848	G	0.035914	0.004569	3.83E-15	*SNX5*	NA	NA	1
20	chr20:41126514:G:A	41126514	G	-0.03487	0.004197	9.78E-17	*TOP1*	NA	NA	1
21	chr21:43290732:G:A	43290732	G	-0.0553	0.009298	2.72E-09	*CRYAA*	NA	NA	NA
22	chr22:30218448:G:A	30218448	G	-0.02684	0.004846	3.02E-08	*LIF*	NA	NA	NA
22	chr22:45398055:A:G	45398055	A	0.024771	0.004258	5.99E-09	*SMC1B*	NA	NA	NA
23	chrX:110648522:C:T	1.11E+08	C	0.048476	0.005778	4.84E-17	*CHRDL1*	NA	NA	NA

Beta, se and p-value reported were from conditional analysis for conditionally-independent signal. The most significant signal for each mapped gene was reported. SNP and PUBMED ID for reported variants were showed. BP: base position; Chr: chromosome; b, beta; se: standard error; NA: data not available; rsID: reference SNP (single nucleotide polymorphism) cluster ID.

Reference

1 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 (London, England)* **396**, 1223-1249, doi:10.1016/s0140-6736(20)30752-2 (2020).

2 Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet (London, England)* **396**, 1204-1222, doi:10.1016/s0140-6736(20)30925-9 (2020).

3 Repositioning of the global epicentre of non-optimal cholesterol. *Nature* **582**, 73-77, doi:10.1038/s41586-020-2338-1 (2020).

4 Roth, G. A. *et al.* Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019: Update From the GBD 2019 Study. *Journal of the American College of Cardiology* **76**, 2982-3021, doi:10.1016/j.jacc.2020.11.010 (2020).

5 Willer, C. J. *et al.* Discovery and refinement of loci associated with lipid levels. *Nature genetics* **45**, 1274-1283, doi:10.1038/ng.2797 (2013).

6 Liu, D. J. *et al.* Exome-wide association study of plasma lipids in >300,000 individuals. *Nature genetics* **49**, 1758-1766, doi:10.1038/ng.3977 (2017).

7 Tabassum, R. *et al.* Genetic architecture of human plasma lipidome and its link to cardiovascular disease. *Nature communications* **10**, 4329, doi:10.1038/s41467-019-11954-8 (2019).

8 Richardson, T. G. *et al.* Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: A multivariable Mendelian randomisation analysis. *PLoS medicine* **17**, e1003062, doi:10.1371/journal.pmed.1003062 (2020).

9 Teslovich, T. M. *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707-713, doi:10.1038/nature09270 (2010).

10 Kathiresan, S. *et al.* Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nature genetics* **40**, 189-197, doi:10.1038/ng.75 (2008).

11 Aulchenko, Y. S. *et al.* Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nature genetics* **41**, 47-55, doi:10.1038/ng.269 (2009).

12 Sabatti, C. *et al.* Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nature genetics* **41**, 35-46, doi:10.1038/ng.271 (2009).

13 Willer, C. J. *et al.* Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nature genetics* **40**, 161-169,

doi:10.1038/ng.76 (2008).
14 Graham, S. E. *et al.* The power of genetic diversity in genome-wide association studies of lipids. *Nature* **600**, 675-679, doi:10.1038/s41586-021-04064-3 (2021).
15 Hoffmann, T. J. *et al.* A large electronic-health-record-based genome-wide study of serum lipids. *Nature genetics* **50**, 401-413, doi:10.1038/s41588-018-0064-5 (2018).
16 Klarin, D. *et al.* Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program. *Nature genetics* **50**, 1514-1523, doi:10.1038/s41588-018-0222-9 (2018).
17 Schaid, D. J., Chen, W. & Larson, N. B. From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nature reviews. Genetics* **19**, 491-504, doi:10.1038/s41576-018-0016-z (2018).
18 Chasman, D. I. *et al.* Forty-three loci associated with plasma lipoprotein size, concentration, and cholesterol content in genome-wide analysis. *PLoS genetics* **5**, e1000730, doi:10.1371/journal.pgen.1000730 (2009).
19 Kettunen, J. *et al.* Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. *Nature communications* **7**, 11122, doi:10.1038/ncomms11122 (2016).
20 Marchini, J., Cardon, L. R., Phillips, M. S. & Donnelly, P. The effects of human population structure on large genetic association studies. *Nature genetics* **36**, 512-517, doi:10.1038/ng1337 (2004).
21 Morris, T. T., Davies, N. M., Hemani, G. & Smith, G. D. Population phenomena inflate genetic associations of complex social traits. *Science advances* **6**, eaay0328, doi:10.1126/sciadv.aay0328 (2020).
22 Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics* **38**, 904-909, doi:10.1038/ng1847 (2006).
23 Mbatchou, J. *et al.* Computationally efficient whole-genome regression for quantitative and binary traits. *Nat Genet* **53**, 1097-1103, doi:10.1038/s41588-021-00870-7 (2021).
24 Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nature genetics* **44**, 369-375, s361-363, doi:10.1038/ng.2213 (2012).
25 Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature genetics* **47**, 291-295, doi:10.1038/ng.3211 (2015).
26 Graham, S. E. *et al.* The power of genetic diversity in genome-wide association studies of lipids. *Nature* **600**, 675-679, doi:10.1038/s41586-021-04064-3 (2021).
27 Ziyatdinov, A. *et al.* Genotyping, sequencing and analysis of 140,000

adults from Mexico City. *Nature* **622**, 784-793, doi:10.1038/s41586-023-06595-3 (2023).

28 Kamiza, A. B. *et al.* Multi-trait discovery and fine-mapping of lipid loci in 125,000 individuals of African ancestry. *Nature communications* **14**, 5403, doi:10.1038/s41467-023-41271-0 (2023).

29 Cadby, G. *et al.* Heritability of 596 lipid species and genetic correlation with cardiovascular traits in the Busselton Family Heart Study[S]. *Journal of lipid research* **61**, 537-545, doi:https://doi.org/10.1194/jlr.RA119000594 (2020).

30 Luo, Y. *et al.* Estimating heritability and its enrichment in tissue-specific gene sets in admixed populations. *Human molecular genetics* **30**, 1521-1534, doi:10.1093/hmg/ddab130 (2021).

31 Klarin, D. *et al.* Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program. *Nature genetics* **50**, 1514-1523, doi:10.1038/s41588-018-0222-9 (2018).

CHAPTER 6: Observational associations between major lipid traits and atherosclerotic cardiovascular mortality: Prospective cohort study

Chapter summary

Previous studies on the association between blood lipids and CVD have tended to rely on traditional lipid panels in Western populations, failing to distinguish the size, compositions, and subclass of lipoprotein particles. Using data from 107,504 individuals with NMR biomarkers who were aged 35-84 years when recruited into MCPS, this chapter assesses the prospective associations of a range of composite and subclass lipid-related traits with ASCVD. For clinically relevant, composite lipid traits (LDL-C,HDL-C, TG, ApoB, and ApoA1), when assessed individually, *usual* level of ApoB and TG, but *not* LDL-C, were positively associated with risk, while usual level of ApoA1 and HDL-C were inversely associated with risk in women but positively associated with risk in men. After accounting for the effect of other lipids, the key findings are that, given LDL-C, usual ApoB was strongly *positively* associated with risk (per 2SD change: HR, 1.76, 1.71-1.81) while, given ApoB, usual LDL-C was strongly *negatively* associated with risk (0.58, 0.56-0.59). These results both implicate small dense LDL particles as being particularly atherogenic. The strong, log-linear association with TG was decreased after accounting for ApoB. Further investigation on lipid-related subfractions confirmed that concentration of LDL particles rather than the cholesterol in them *per se* was more closely related to ASCVD mortality risk, with each ApoB-containing particle (VLDL-P, L-LDL-P and S-LDL-P) that has exactly one ApoB molecule related to similar risk of ASCVD. By contrast, smaller HDL particles (S-HDL-P) and their cholesterol contents (S-HDL-C) were negatively associated with risk, whereas larger ones (XL-HDL-P and XL-HDL-C) were positively associated with risk. Overall, the findings suggest that: 1) the concentration/number of ApoB containing particles is closely correlated with the risk of ASCVD; 2) for specific types of lipoprotein particles (LDL and HDL), their subfractions

with varying sizes differ in their relevance to risk; 3) direct measures of the number of different lipoproteins and their subtypes may be more relevant to the risk than the total amount of cholesterol within these particles.

6.1 Background

Atherosclerotic cardiovascular disease (ASCVD) and its principal clinical manifestations, including coronary heart disease (CHD) and ischaemic stroke, is the leading cause of morbidity, premature death and rising health costs worldwide.[1,2] The total number of deaths due to cardiovascular disease (CVD) has increased steadily since 1990 and reached 18.6 million worldwide in 2019, corresponding to 380 million years of lost life and 34.4 million person-years of disability in 2019.[2] Although the global age-standardized mortality rate of CVD declined over the past three decades, the rate is increasing in many locations of low- and middle-income countries, indicating intensified regional efforts to the control of potential risk factors are needed to address this persistent global health threat.[2]

A variety of modifiable risk factors have been identified to be associated with an increased risk of cardiovascular events, such as high levels of blood pressure, fasting glucose, and low density lipoprotein (LDL) cholesterol, obesity, smoking, alcohol consumption, and low physical activity.[3,4] Among these risk factors, LDL and other ApoB-containing particles play a particularly central role in the pathophysiology of atherosclerosis and incident ASCVD events.[5,6] In 2019, about two-thirds of global CVD deaths were caused by atherosclerotic conditions in which LDL and other cholesterol-rich ApoB-containing lipoproteins are primary drivers.[2] The comprehensive assessment of the associations of major blood lipids and apolipoproteins with the risk of ASCVD and the management of dyslipidemias, accompanying other public health interventions, are critical to reduce the burgeoning burden of CVD in the world.

Evidence from genetic studies, prospective cohort studies, Mendelian randomization studies, and randomized clinical trials of lipid-lowering therapies in predominantly Western populations have suggested LDL and other ApoB-containing lipoproteins (VLDL remnants, IDL, and Lp(a)) have causal effect on ASCVD.[6] However, the role of other

lipids such as HDL-C and TG in the etiology of ASCVD is not so well established and uncertainty persists about the merits of measurement of apoAI and ApoB.[7,8] Most of the evidence has tended to rely on traditional lipid panels (TG, TC, LDL-C, and HDL-C) measured by conventional chemical assays, which fail to distinguish the size, compositions, and subclass concentration of lipoprotein particles. Given the complexities and inter-relatedness of lipid traits, a comprehensive NMR approach might be more preferable and informative than traditional chemical assays to quantify the concentration and size of lipoprotein subfractions and their lipid contents that may have contrasting effects on ASCVD. Previous observational studies of NMR-measured lipid related traits and ASCVD (**Chapter 2**) support the evidence based on traditional lipid panels that LDL and other ApoB-containing lipoproteins are directly implicated in the development of ASCVD, as introduced in **Chapter 1**. However, these NMR studies are limited to participants of primarily European ancestry and the sample size is still modest compared with studies of traditional lipids.

As the current largest study with the availability of both genetic and NMR lipid data in about 120,000 Latin American individuals, MCPS provides a unique opportunity to assess the association between lipid traits and ASCVD risk with a larger sample size and in greater detail than previously possible. The current chapter reports the prospective associations between NMR lipid traits (lipids, lipoproteins, and apolipoproteins) and the risk of ASCVD (fatal ischaemic heart disease, ischaemic stroke, or peripheral arterial disease) in a large Mexican population after taking into account confounding and reverse causality, and assesses whether these associations differ by age, sex, demographic, physical, and lifestyle characteristics.

6.2 Methods

6.2.1 Study design and study population

The overall study design and recruitment of participants are detailed in **Chapter 3**. Briefly,

the MCPS is a prospective cohort of 159,755 Mexican adults aged 35 years and older at enrollment, recruited between 1998 and 2004 from two districts of Mexico City. For this chapter, participants aged 35 to 84 years at baseline who had no prior history of self-reported chronic diseases except diabetes (cardiovascular disease, cancer, or chronic renal, hepatic, or pulmonary disease), were not using lipid-lowering treatment, with complete data on NMR metabolomics and covariates, and with reliable mortality linkage were included. A random subset of about 10,000 surviving participants who were resurveyed between 2015 and 2019 with repeated measures of NMR metabolomics were also included in the analyses to account for regression dilution bias.

6.2.2 Definitions of exposure and outcome

In this study investigating the association between blood lipid levels and the risk of atherosclerotic vascular mortality, three sets of lipid-related phenotypes were used as the main exposures. All lipid-related traits were measured by high-throughput NMR metabolomics platform (the Nightingale Health Ltd.), as detailed in **Chapter 3**. Three 'arrays' of measurements were considered. The first array includes *composite lipid measures* including major blood lipids (TG, LDL-C, and HDL-C) and apolipoproteins (ApoB and ApoA1), which are consistent with traditional lipid panels used in clinical practice and research. The second array relates to *lipoprotein subfractions,* and includes the absolute concentrations (particle number) of 14 lipoprotein subclasses (6 VLDL, IDL, 3 LDL, and 4 HDL), the lipids contained in each lipoprotein subclass (cholesterol [total, esterified, and free], triglycerides, phospholipids, and total lipids), mean lipoprotein particle size (VLDL, LDL, and HDL), concentrations of ApoB-100 and ApoA1, and the ApoB/ApoA1 ratio. The second array was used to supplement the significant discovery from first array in in a more targeted and tailored way. For example, cholesterol in LDL subfractions were further assessed given the findings from the composite LDL-C. The third array includes transformations of composite lipid measures, including 'clinical' LDL-C (estimated from total cholesterol, HDL cholesterol and triglycerides according to the

Friedwald formula,[9] rather than LDL-C directly measured by NMR as above), non-HDL-C (TC - HDL-C), ratio of ApoB to ApoA1 (ApoB/ApoA1), ratio of non-HDL-C to HDL-C (non-HDL-C/HDL-C), remnant cholesterol (TC - HDL-C - LDL-C, representing VLDL-C and IDL-C; remnant-C). This third array is used for sensitivity analyses only.

Previous studies have suggested that the association between NMR lipid traits and cardiovascular events vary by vascular domain (**Chapter 2**). Therefore, in this thesis the outcome studied was limited to cardiovascular mortality primarily caused by atherosclerosis (ASCVD mortality, including death due to ischaemic heart disease, ischaemic stroke, or peripheral arterial disease) (**Table 6.S1**). Since the majority (91%) of such deaths were from ischemic heart disease, the separate components are not analyzed individually. The use of only ischemic heart disease by excluding other causes lead to similar risk estimates (data not shown).

6.2.3 Statistical analysis

Cox proportional hazards regression was used to estimate the hazard ratios (HRs) and 95% confidence intervals (CIs) for the associations of lipid-related traits with the risk of ASCVD mortality. The main analyses included deaths occurring at ages 35 to 84 years. To assess the shape of associations with ASCVD mortality eligible participants were categorized into sex-specific fifths of baseline concentrations of blood lipid-related traits. In figures, group-specific confidence intervals around the HRs (including for the reference group with HR of 1·0) were estimated using the variance of the log risk in each group.

The floating absolute risk method was used to calculate HRs for all exposure categories after the first group was selected as the reference group, with an HR of 1. The variance of the log risk in each category, including the reference, was calculated (from the variances and covariances of the log HRs in all categories except the reference group) and used to obtain group-specific 95% CIs.[10] Participants who did not die during the age-

at-risk range were censored at the end of the risk period (e.g. on their 75th birthday for analyses of deaths occurring before age 75 years), or at the end of follow-up (December 31, 2020). Participants who died from a cause other than ASCVD deaths were censored at their date of death. Potential confounding factors were selected based on *a priori* knowledge and were progressively added into the models. All analyses were stratified by 5-year age-at-risk (age attained at the time of death or censoring), and were adjusted simultaneously (except where the variable was included in subsidiary analyses to assess effect modification) for sex, study region (Coyoacán and Iztapalapa), and education level (university or college, high school, elementary school, others).

To assess the average strength of the association with ASCVD mortality, baseline blood lipid-related traits were included in analyses as continuous variables. HRs were reported for 2 SD higher *usual* lipoprotein subfractions, lipids and apolipoproteins (to allow for comparisons across lipid traits), with the same adjustment for covariates as described earlier. However, because the baseline *measured* levels of NMR lipid traits may underestimate the associations of long-term "*usual*" levels of blood lipids (due to random measurement error and the variation in blood lipids within individuals over time),[11] the log HRs (and their standard errors) were divided by an estimate of the regression dilution ratio (RDR) for that lipid marker. These RDRs were conservatively estimated as the ratio of the range of mid-point to baseline means using the non-parametric method.[12] Similarly, for the categorical analyses described earlier, the HRs were plotted against the expected mean usual level in each category (calculated as $M + RDR \times [b - M]$, where b is the category-specific mean baseline level and M is the overall mean).

As several lipid traits are highly correlated according to their intrinsic chemical characteristics (e.g., ApoB and LDL-C; ApoA1 and HDL-C), and one lipid trait is also influenced by another in the biological pathway of lipid metabolism (e.g., high TG levels are related to low HDL-C level), analyses of mutual adjustments were conducted for correlated lipid pairs of interest. For example, more than 90% of ApoB is on LDL particles

(each particle has one ApoB molecule) and most plasma cholesterol is in LDL particles. Consequently, levels of ApoB, LDL particles, and LDL-C are closely correlated, as are levels of apoA1 and HDL-C. HDL particles can transfer cholesterol to TG-rich VLDL particles in exchange of TG, thus high TG levels are associated with low HDL-C level. The independent association of HDL-C (or cholesterol in different subtypes of HDL particle) with ASCVD risk should account for any effect of LDL particles on risk. However, as the random errors for the highly correlated lipid traits are also correlated, the independent relationship of two correlated factors with risk may not be simply derived as for a single-factor relationship (e.g., by entering both factors into the regression model). Instead, to assess the independent effects of two highly correlated traits, such as LDL-C and ApoB, the residuals of each trait after regression on the other were extracted and used in analyses. For example, the residuals from a linear regression of *measured* LDL-C on ApoB represent the difference between LDL-C and the *prediction of LDL-C* based only on ApoB (referred to as LDL-C given ApoB). Similarly, the residuals from a linear regression of *measured* ApoB on LDL-C represent difference between ApoB and the prediction of ApoB based only on LDL-C (referred to as ApoB given LDL-C). For each set of residuals, their repeatability over time was examined as for the individual traits to estimate and then adjust for regression dilution bias. The pairs of correlated lipids of interest included ApoB and LDL-C, ApoA1 and HDL-C, HDL-C and ApoB, TG and ApoB and LDL-C.

In the main analyses, the associations of lipid-related traits as continuous variables with ASCVD were reported by sex (men and women) and age-at-risk groups (35-59 years, 60-74 years, and 75-84 years). The associations of lipid-related traits as categorical variables with ASCVD were reported by sex (men and women). For both analyses, the associations during different follow-up periods (all follow-up, and excluding first 5/10 years of follow-up) were reported to illustrate the extent to which the associations of interest may have been distorted by reverse causality bias.

To investigate the consistency of the associations with ASCVD mortality across the study population (besides age and sex groups), the adjusted HRs were further estimated and compared across strata of risk factors associated with blood lipids and ASCVD (with mutual adjustment for the other shown risk factors). These population subgroups were selected *a priori* based on theoretical importance, previous evidence identified in the literature review (**Chapter 2**) and the correlation of blood lipids with baseline characteristics (**Chapter 3**), and included district, education level (university, high school, elementary, others), smoking (never, former, current), alcohol consumption (never, former, current), physical activity (none, up to twice weekly, at least 3 times weekly), SBP (<125, 125-145, ≥145 mmHg), BMI (<25, 25-30, ≥30 kg/m^2), and fasting duration (<8, 8-12, and >12 hours). Chi-square tests for trend and heterogeneity were used to assess the consistency of HRs across population subgroups.

Multiple sensitivity analyses were conducted to assess the robustness of the risk estimates in main analyses. First, participants with diabetes at baseline (self-reported diagnosis, use of glucose-lowering medications, or HbA1c ≥ 6.5%) were excluded to examine the extent to which the associations may be distorted by diabetes at baseline that could affect lipid levels and the risk of ASCVD. Second, transformed lipid measures as complementary to major lipid traits, including non-HDL-C, ratio of ApoB to ApoA1, ratio of non-HDL-C to HDL-C (non-HDL-C/HDL-C), and remnant cholesterol, were used to estimate the association with the risk of ASCVD. Third, for the main significant findings, additional adjustments were made for BMI (<25, 25-30, ≥30 kg/m2), smoking (never, former, less than daily, less than 10 cigarettes per day, 10 or more cigarettes per day), alcohol consumption (never, former, up to 3 times per month, up to 3 times per week, more than 3 times per week), physical activity (none, up to twice weekly, at least 3 times weekly), baseline SBP (<125, 125-145, ≥145 mmHg), and fasting duration (<8, 8-12, and >12 hours).

All analyses were assessed using SAS version 9.4 (SAS Institute, Cary, NC) and related

plots were created in R version 4.2.1 (The R Foundation for Statistical Computing). Nominal P values are presented and 2-sided $P < 0.05$ indicated significance.

6.3 Results

6.3.1 Relationship of usual levels of lipid-related traits to ASCVD mortality risk

6.3.1.1 ApoB-related factors

Baseline characteristics of 107,504 participants without prior chronic disease except diabetes and aged 35-84 at recruitment by tenths of the LDL-C distribution are shown in **Table 6.1**. **Figures 6.1** and **6.2** show the relative risks of ASCVD death (i.e., the Cox mortality hazard ratios) per 2SD difference in usual levels of ApoB and LDL-C, respectively. Overall, there were 2713 ASCVD deaths at ages 45-85 years. During overall follow-up period, a 2-SD change in usual LDL-C (1.0 mmol/L) and ApoB (0.4 g/L) were associated with an adjusted HR of 1.06 (95% CI 1.03-1.09) and 1.29 (1.25-1.33), respectively. The associations were stronger in women than in men (during full follow-up for LDL-C, 1.10, 1.05-1.15 in women vs 0.98, 0.94-1.03 in men; for ApoB, 1.37, 1.31-1.42 in women vs 1.14, 1.08-1.19 in men). The patterns of associations between sex and age-at-risk groups were broadly similar across analyses of full follow-up period, and those excluding the first 5 or 10 years of follow-up. In women, the strengths of association between LDL-C and ApoB with ASCVD mortality were larger at younger than at older ages but in men this pattern was not observed.

As ApoB and LDL-C are highly correlated, the independent relevance of each lipid factor was then assessed. **Figures 6.3** and **6.4** provide risks of ASCVD death per 2SD difference for usual levels of ApoB *given* LDL-C and LDL-C *given* ApoB, respectively. Given the LDL-C level, higher usual ApoB was positively and log-linearly associated with risk of ASCVD death (HR per 2SD 1.76, 95% CI 1.71-1.81), which was stronger than the risk estimate without accounting for LDL-C (1.29, 1.25-1.33). The associations were observed in both men (1.54, 1.47-1.62) and women (1.88, 1.81-1.96), and were not much

affected by exclusion of early follow-up. In both men and women, the association of ApoB given LDL-C to ASCVD mortality was stronger at younger than at older ages, but even at ages 75-84 a 2-SD higher ApoB given LDL-C level was associated with substantially increased risk (HR 1.48, 95% CI 1.41-1.55). By contrast, given the ApoB level, usual LDL-C (reflecting cholesterol content per LDL particle) was strongly *negatively* associated with the risk (HR per 2SD 0.58, 95% CI 0.56-0.59). Taken together, these results suggest that the small, dense LDL with relatively lower cholesterol content than those in larger ones are particularly hazardous.

By leveraging the advantages of NMR with respect to measuring both concentrations and particle sizes of lipid traits, associations of risk with more direct measures, including LDL particles and cholesterol in LDL particles of varying size were further explored based on the above findings of small dense, LDL particles. **Figure 6.5** shows the age- and sex-specific associations of usual LDL-related subfractions with ASCVD mortality risk during the full period of follow-up period. Similar effect estimates were observed for the concentration (number) of LDL particles (LDL-P: 1.26, 1.22-1.30), and its subfractions with varying size (large LDL-P: 1.28, 1.24-1.32; small LDL-P: 1.33, 1.29-1.37). A slightly stronger effect was observed for VLDL particle (1.51, 1.47-1.56). These findings are biologically sensible as each LDL/VLDL particle contains exactly one ApoB molecule, supporting the close relationship of ApoB and risk of ASCVD observed earlier.

Figure 6.6 shows the risk of ASCVD death per 2SD difference in usual TG levels. During the overall follow-up period, a 2-SD higher usual TG was associated with an adjusted HR of 1.36 (95% CI 1.32-1.41), which was similar to the effect estimates of ApoB-containing particles. This association was stronger in women than men. As TG are mainly carried by precursors of LDL particles (e.g., VLDL and IDL), ApoB was further accounted for to assess the independent relevance of TG. **Figure 6.7** shows risks of ASCVD death per 2SD difference in usual levels of TG *given* ApoB. During the full follow-up period, the associations attenuated to null after accounting for ApoB and LDL-C (for all age groups:

1.01, 0.98-1.04) and an inverse association was observed in older participants (75-84 age-at-risk group). After excluding the first 5/10 years of follow-up, a positive association was observed, albeit with smaller effect size (for analyses excluding the first 10 years of follow-up: with/without accounting for ApoB and LDL-C: 1.35, 1.30-1.40 vs 1.69, 1.62-1.77). These results suggest potential reverse causality where latent or undiagnosed diseases could affect the TG level and its association with the risk of ASCVD, and that the ASCVD risk related to TG was partly dependent on the level of ApoB-containing particles as proxied by ApoB and LDL-C.

6.3.1.2 ApoA1-related factors

Figures 6.8 and **6.9** show the risk of ASCVD death per 2SD difference in usual levels of ApoA1 and HDL-C, respectively. Overall, there were modest inverse associations with ASCVD risk for both ApoA1 (0.95, 0.92-0.98) and HDL-C (0.94, 0.91-0.97) during the full follow-up period. However, the associations were in opposite directions for men and women across age groups for both lipid indices. For example, in the 35-59 age-at risk group, the HRs for HDL-C were 1.46 (1.30-1.65) in men but 0.68 (0.59-0.79) in women. ApoB and TG can affect HDL-C levels (e.g., TG-rich precursors of LDL particles can transfer TG to HDL in exchange for cholesterol through CETP). Therefore, the independent relevance of ApoA1 and HDL-C was assessed by accounting for the effect of ApoB on ASCVD. **Figures 6.10** and **6.11** show risks of ASCVD death per 2SD higher usual levels of ApoA1 *given* ApoB and of HDL-C *given* ApoB, respectively. After accounting for the effect of ApoB, both indices were associated with higher relative risks of ASCVD in men but lower relative risks in women.

Similar to LDL particles, HDL particles also vary in size, cholesterol content, and biological activity. **Figure 6.12** shows the age- and sex-specific associations of usual HDL-related subfractions with the risk of ASCVD mortality during the full follow-up period. HDL-P was inversely associated with the risk of ASCVD mainly in women. However,

larger HDL diameter was related to increased risk: XL HDL-P (per 2-SD change: 1.58, 1.53-1.62) and XL HDL-C (1.48, 1.44-1.52) associated with higher risk, whereas S HDL-P (0.79, 0.76-0.81) and S HDL-C (0.76, 0.74-0.78) were associated with lower risk. These results suggested specific type(s) of HDL particle may be more protective than others, whereas some type(s) may be hazardous especially in men.

6.3.2 Relationship of usual levels of other composite lipid traits to ASCVD mortality risk

Other composite lipid traits of clinical relevance included clinical LDL-C, estimated by the Friedewald equation (LDL-C = TC - HDL-C - TG/5; Figure 6.S1), remnant-C (TC - LDL-C – HDL-C; Figure 6.S2), non-HDL-C (TC – HDL-C; Figure 6.S3), non-HDL-C/HDL-C ratio (Figure 6.S4), and ApoB/ApoA1 ratio (Figure 6.S5). Overall, the strength and direction of the associations with non-ratio lipid indices were generally consistent with corresponding ApoB-related and ApoA1-related traits, as they are broadly equivalent in chemical composition (albeit estimated in different ways). For example, LDL-C and clinical LDL-C, remnant-C and VLDL-C, non-HDL-C and ApoB-containing particles exhibited similar associations with ASCVD risk. The ApoB/ApoA1 ratio was a stronger predictor of ASCVD mortality than the non-HDL-C/HDL-C ratio (again, particularly in women).

6.3.3 Secondary analyses and sensitivity analyses

The main analyses for composite lipid traits were stratified by age-at-risk groups, sex, and follow-up period. Subgroup analyses for ApoB *given* LDL-C further stratified by education level (university, high school, elementary, others), smoking (never, former, current), alcohol consumption (never, former, current), physical activity (none, up to twice weekly, at least 3 times weekly), SBP (<125, 125-145, ≥145 mmHg), BMI (<25, 25-30, ≥30 kg/m^2), and fasting duration (<8, 8-12, and >12 hours) showed broadly consistent results (Figure 6.S6). Sensitivity analyses adjusted for additional confounders including smoking, alcohol consumption, physical activity, SBP, BMI, and fasting duration, for ApoB

given LDL-C showed similar results (full adjustment in sensitivity analysis: HR per 2SD 1.71, 95% CI 1.66-1.76; basic adjustment in the main analysis: HR per 2SD 1.76, 95% CI 1.71-1.81). Sensitivity analyses excluding participants with diabetes (and other chronic diseases) at baseline for all lipid-related traits assessed showed consistent trends of association across age-at-risk groups and sex (Figure 6.S7-6.S16). Sensitivity analyses using transformations of composite lipid measures derived similar trend of association compared with the relevant lipid traits in the main analysis (e.g., direct measures of LDL-C vs clinical LDL-C estimated by Friedewald equation). The regression dilution ratios between baseline and resurvey (~15 years later) used to calculate usual levels of blood lipids are shown in Figure 6.S17.

6.4 Discussion

6.4.1 Main findings in context

In this prospective cohort study of a large Latin American population the association between a range of NMR-measured lipid-related traits and risk of ASCVD death was assessed. For clinically relevant, composite lipid traits (LDL-C, HDL-C, TG, ApoB, and ApoA1), when assessed individually, usual levels of ApoB and TG, but *not* LDL-C, were positively associated with risk, while usual levels of ApoA1 and HDL-C were inversely associated with risk in women but positively associated with risk in men. After accounting for the effect of other lipids, the key findings are that, given LDL-C, usual ApoB was strongly positively associated with risk; and given ApoB, usual LDL-C was negatively associated with risk. Both of these results implicate smaller denser LDL particles as being particularly atherogenic, a finding supported by examination of LDL diameter directly. The strong, log-linear association with TG with ASCVD mortality was largely decreased after accounting for ApoB and LDL-C during overall follow-up. Further investigation of lipid-related subfractions confirmed that cholesterol in small LDL particles (S-LDL-C) was more closely related to risk than L-LDL-C, with each ApoB-containing particle (VLDL-P,

LDL-P, L-LDL-P, and S-LDL-P) that has exactly one ApoB molecule related to similar relative risk of ASCVD. By contrast, smaller HDL particles (S-HDL-P) and their cholesterol contents (S-HDL-C) were negatively associated with risk, but larger ones (XL-HDL-P and XL-HDL-C) were positively associated with risk. Overall, the findings suggest that: 1) plasma ApoB, reflecting the concentration/number of ApoB containing particles, is positively (and potentially causally) related to the risk of ASCVD; 2) for specific types of lipoprotein particles (LDL and HDL), their subfractions with varying sizes differ in their effects on risk; 3) Direct measures of the number of different lipoproteins and their subtypes may be more relevant to the risk than the total amounts of cholesterol within these particles.

6.4.2 Relevance of ApoB-containing lipoproteins to ASCVD

6.4.2.1 Variability in ApoB-containing lipoproteins

As LDL particles greatly outnumber their TG-rich precursors (e.g., VLDL and IDL) in plasma, ApoB level is closely related to the concentration of LDL particles. In most cases, LDL-C concentration is also correlated with LDL particle number and ApoB level, thereby providing similar information on the risk of cardiovascular diseases.[6,13,14] Therefore, LDL-C is widely used as a surrogate for LDL particles and ApoB in most studies to evaluate the clinical effectiveness of lipid-lowering drugs and to predict cardiovascular risk. Increasing evidence from genetic studies, prospective cohort studies, Mendelian randomization studies, and randomized clinical trials of LDL-lowering therapies have established that low-density lipoproteins have a causal effect on ASCVD.[6] When assessed individually in the current study, however, LDL-C was not significantly related to risk, whereas LDL particle concentration and ApoB were strongly associated with risk. In addition, after accounting for the effect of ApoB, LDL-C was inversely related to risk, suggesting LDL particles with lower cholesterol content per particle (i.e., small, dense LDL particles) are more harmful. The underlying biological mechanisms underlying small,

dense LDL-P include CETP-mediated transfer of TG from VLDL and their remnants for cholesterol in LDL decreased LDL-C.[15] Subsequent hydrolysis of TG within LDL particles by HL activity shrinks LDL to smaller, denser and cholesterol-depleted LDL particles. Previous studies also suggested that under certain conditions where TG levels exceeded 150 mg/dL (e.g., the metabolic syndrome, Type 2 diabetes, and hypertriglyceridemia), smaller, denser, and cholesterol-depleted LDL particles predominated in LDL particles leading to discordance between LDL-C and ApoB.[16] Considering the study population from MCPS has extremely high prevalence of diabetes and obesity, such discordance could explain the observed lack of association with LDL-C without accounting for ApoB. Additional studies using discordance analysis (e.g., in populations with an absolute difference of >10 percentage points between their ApoB and LDL-C percentiles), further demonstrate that ApoB was more strongly associated with the risk than LDL-C or non-HDL-C.[13,17,18] Discordant groups defined by multiple methods suggested that about 20% to 60% of the population have discordant levels of ApoB and LDL-C, which are sufficiently large to be clinically relevant.[14] LDL-C, given ApoB, was *inversely* related to the risk of ASCVD mortality. For the direct measures of LDL subtypes and their cholesterol contents, it was further observed that cholesterol in small but not large LDL was related to increased risk. These findings again suggest that different LDL particles with varying sizes differ, perhaps substantially when TG level is extremely high, in their adverse effects. Direct measurements of LDL particles and their subfractions may be more informative about risk than the total LDL-C. The extensive NMR measures of lipoprotein particle subfractions could be more informative than conventional clinical lipid panel to predict ASCVD risk and better understand the comparative role of different lipoprotein subfractions with varying size and biological functions.[19]

6.4.2.1 ApoB as a more accurate and informative marker for ASCVD

In line with previous evidence,[20-23] the current findings collectively suggest the superiority of ApoB over cholesterol as a marker for ASCVD based on the variability in ApoB-

containing particles, the discordant changes in LDL-C and ApoB, and the observed differences in their association with risk. After accounting for LDL-C, ApoB was strongly and log-linearly associated with increased risk of ASCVD in both men and women. Each of ApoB-containing particle was related to similar risk (e.g., VLDL, LDL, and also both S-LDL and L-LDL), which can be explained by the fact that cholesterol content per ApoB particle varied but each particle has one ApoB, and its atherogenic effect first depends on the entry of ApoB particles into the arterial wall, which mostly depends on the ApoB particles in the arterial lumen. The traditional model of atherosclerosis presumes that the cholesterol content of VLDL, LDL, and Lp(a) determine the mass of cholesterol deposited within the arterial wall and will drive atherogenesis. However, the current findings, in conjunction with previous substantial evidence,[20,21] indicate that ApoB more accurately measures the risk related to the ApoB particles than does LDL-C or non-HDL-C. S-LDL-C was more strongly related to ASCVD mortality than L-LDL-C, whereas S-LDL-P and L-LDL-P had similar associations with risk. This may be explained by the following atherosclerosis model of ApoB-containing particles. First, among all ApoB particles, smaller, cholesterol-depleted particles enter the arterial wall more easily, bind more avidly to the glycosaminoglycans, and thus are more easily trapped than larger, cholesterol-rich particles. By contrast, cholesterol-rich particles could deposit more cholesterol despite fewer particles entering the arterial wall. The net result is that all ApoB-containing particles are equally atherogenic.[14]

Evidence from MR studies further support circulating number of ApoB particles as the most direct and accurate measure of atherogenic injury attributable to the ApoB-containing particles, which will be thoroughly assessed in **Chapter 7**. MR studies combining variants in the *CETP* and *HMGCR* genes to mimic the effects of CETP inhibitors and statins have found that combination use was associated with discordant reductions in LDL-C and ApoB concentration (i.e., the same reduction in LDL-C but attenuated reductions in ApoB level).[21]

Taken together, the evidence suggests that ApoB-containing particles are atherogenic and the risk of subintimal retention and accumulation of these particles leading to atherosclerosis is more closely associated with the number (concentration) of circulating ApoB particles than with the variable mass of cholesterol within them. Direct measures of ApoB integrate the information from traditional lipid panels, and may be more accurately and closely associate with cardiovascular risk (especially in populations with high prevalence of obesity and diabetes).

6.4.3 Relevance of TG to ASCVD

When assessed individually, TG was strongly and log-linearly associated with the ASCVD mortality risk, with even stronger associations after excluding the first 5 or 10 years of follow-up. However, high TG level is related to a low HDL-C and high prevalence of small dense cholesterol-depleted LDL particles, referred to as atherogenic dyslipidemia.[24] These interrelated atherogenic changes in lipoproteins may jointly increase the risk of ASCVD, making it difficult to disentangle the independent effect of TG,[25,26] especially in the setting of traditional epidemiological studies. After accounting for ApoB and LDL-C, the overall association between TG and ASCVD mortality attenuated to the null, with an apparently inverse association in older participants. These results suggest potential reverse causality where latent or undiagnosed diseases could affect the TG level and its association with the risk of ASCVD, and that the risk related to TG was partly dependent on the level of ApoB-containing particles. The current findings are in line with previous evidence. A meta-analysis of about 300,000 individuals without CVD at baseline reported that the positive association between TG and ASCVD risk (per 1-SD change, HR, 1.37, 1.31-1.42) was not significant after adjusting for HDL-C and non-HDL-C (0.99, 0.94-1.05).[7] In randomized trials of statins, on-treatment TG levels have been found to be significantly related to CVD risk in statin-treated patients,[27,28] although the association was attenuated after accounting for HDL-C and ApoB-ApoA1 ratio.[28] Taken together, although TG levels appear to provide information about vascular

risk, the independence of TG as a causal factor in the development of ASCVD remains uncertain based on the evidence from the current study and previous observational epidemiological studies.

Evidence from randomized trials also do not provide consistent evidence that lowering TG level reduce the risk of CVD events.[8] The addition of TG-lowering treatments such as fibrates and niacin to statin therapy did not lead to significant clinical benefit, although the TG level was reduced by 30% to 50%.[29-31] However, in those with a subgroup of patients with hypertriglyceridemia (baseline TG > 200 mg/dL) and low HDL-C (< 35 mg/dL) TG-lowering drugs were associated with lower risk of ASCVD events, as the corresponding reduction in apoB-containing particles would only reach clinical significance in those with hypertriglyceridemia,[26,29,32,33] which appears to explain why these trials evaluating fibrates and niacin failed to consistently suggest a reduced risk. Similar to fibrates and niacin, studies of OM3FAs have reported inconsistent results and the clinical benefit was significant in patients with TG levels > 150 mg/dL.[34-36]

MR studies have further demonstrated that TG-lowering *LPL* variants and LDL-C-lowering *LDLR* variants are related to lower risk of ASCVD per unit change in ApoB-containing lipoproteins, suggesting the potential clinical benefit of lowering TG levels is similar to the benefit of lowering LDL-C per unit change in ApoB.[37] Again, this supports the view that ApoB is more accurately and closely associated with cardiovascular risk than the TG or cholesterol content within ApoB-containing particles. If so, any clinical benefit of lowering whether TG or cholesterol should be proportional to the absolute change in ApoB.

6.4.4 Relevance of HDL-related traits to ASCVD

Overall, inverse associations of HDL-C and ApoA1 with ASCVD risk were observed in several population subgroups, which were stronger after accounting for ApoB and LDL-C. However, the overall associations were inconsistent. Notably, HDL-related

subfractions were differentially associated with risk, with larger HDL particles and their cholesterol content strongly positively associated with risk, whereas smaller counterparts were negatively associated with risk. Current evidence on the relevance of HDL particles and HDL-C to the development of ASCVD has changed dramatically over recent years and remains inconsistent across previous epidemiological and genetic studies. Several observational studies have reported an inverse association between HDL-C and ASCVD independent of LDL-C,[7,38-40] although with a low range of HDL-C levels studied. Studies investigating larger ranges have indicated a U-shaped association of HDL-C with ASCVD and all-cause mortality, with an increased risk of CVD at high level of HDL-C.[41-44] The current findings add to this previous evidence. Although HDL-P and HDL-C were associated with a lower risk of ASCVD overall, larger HDL particles (which make up most of the HDL cholesterol in plasma) were associated with increased risk, suggesting that different HDL particles with different particle sizes may have a heterogeneous role in ASCVD, possibly due to different function in reverse cholesterol transport. Macrophage cholesterol efflux (MCE) is considered to be a key determinant of atherogenesis and concomitant inflammation in the model of HDL-related reverse cholesterol transport,[45] with inconsistent associations of MCE with ASCVD risk reported.[46-50] Notably, the reversibility of MCE has indicated that high HDL-C may reversely transfer to macrophages and cause atherosclerosis,[50] which may partially explain the increased risk related to the large but not small HDL particles. Similarly, higher mean HDL particle size and HDL-C levels were paradoxically associated with increased risk of CVD in EPIC-Norfolk and IDEAL study.[51] A previous MR study of lipid subfractions also found that only medium and small HDL particles and their lipid contents were causally related to decreased risk and larger HDL particle size increased the risk of CVD outcomes.[52] It should be noted, however, that several studies also found that large HDL particles, but not small HDL particles, or both, were associated with a lower risk and some reported a lack of association, as reviewed in Chapter 2. [17,20,51,53-59] Collectively, the current findings,

in combination with previous evidence, suggest that HDL subfractions may have particular functionality, reflecting different potential role in pathophysiological functions.

Evidence from randomized trials also do not provide evidence that increasing HDL-C level reduce the risk of CVD events. Niacin and fibrates that can lower TG level and increase HDL-C level, and CETP inhibitors that can profoundly increase HDL-C level and slightly decrease LDL-C level. The addition of niacin to a background statin therapy in patients with CVD did not reduce the risk of major vascular events or death. [30,31,50] The only clinical benefit observed in a subgroup of patients with hypertriglyceridemia was largely explained by the reduction in ApoB-containing particles, as discussed in the above TG section. In a meta-analysis investigating the effect of raising HDL-C on CVD risk including 108 randomized clinical trials and about 300,000 individuals on statins, fibrates, niacin, and combinations of lipid-modifying drugs, no significant association was found between elevation in HDL-C and risk of non-fatal and fatal ASCVD, and all-cause mortality after adjusting for change in LDL-C.[60] CETP inhibitors also failed to show clinical benefit although they greatly increased the level of HDL-C by about 100%. Their effects on ASCVD events were either non-significant, negligible,[61-64] or small and likely explained by a reduction in non-HDL-C.[65,66] Evidence from genetic studies also showed that genetically reduced HDL-C is not associated with an increased risk of atherosclerotic cardiovascular events after adjusting for apoB.[21,67-70] The association of HDL-P and HDL-C with ASCVD risk warrants further investigation, especially using the direct of measures of more specific HDL subtypes, to elucidate their biological functions and causality in the development of ASCVD.

6.4.5 Heterogeneity of association with ASCVD between sex

The strength of associations of apoB-containing particles and lipids (e.g., LDL-P, LDL-C, TG) and apoB with risk of ASCVD death were somewhat stronger in women than in men especially at younger ages (35-59 years). There was also unexpected qualitative

heterogeneity in associations with HDL-related traits between sex where HDL-C and ApoA1 were positively related to the risk in men whereas negatively in women as expected. Although there is no consensus that atherosclerotic lipids such as LDL-C and TG are better predictors of ASCVD in women than in men,[71] the current findings from MCPS are largely in line with insights from other large prospective studies. The Copenhagen City Heart Study reported the risks of both MI and all-cause mortality with non-fasting TG are constantly higher in women than in men after adjusting for both traditional confounders such as age, TC, BMI, hypertension, diabetes, smoking, alcohol consumption, physical activity, lipid-lowering therapy, as well as sex-related covariables such as post-menopausal status and hormone replacement therapy.[72] Data from the Copenhagen City Heart Study and the Copenhagen General Population Study reported that for LDL-C, the associations with CVD death were only observed in women (per 1 mmol/L increase in LDL-C: HR 5.31, 3.37-8.37) but not men (0.94, 0.54-1.62), and the associations with all-cause mortality were stronger in women (4.19, 1.72-10.2) than men (1.28, 1.02-1.60).[71]

These findings of sex heterogeneity may be partly explained by both the different impacts of traditional CVD risk factors on women vs. men, and the impacts of sex-specific risk factors. For example, although more prevalent in men, diabetes was associated with a great *relative* (not necessarily absoulte) CVD risk in women than in men across age groups, and women may have greater adiposity and more CVD risk factors at the time of diagnosis. Sex-specific 'risk-enhancing' factors (e.g., polycystic ovary syndrome, gestational diabetes, and early menopause) are also considered in the current guidelines for CVD risk in women, which may underlie the observed sex heterogeneity.[73,74] Women are also disproportionately affected by chronic kidney disease and auto-immune inflammatory disease that are risk factors for ASCVD.[71] Gender-related sociocultural contributors may also contribute to the observed heterogeneity. Emerging experimental evidence from basic research suggests that sex (e.g., sex hormones and X-chromosome)

plays a significant role in the development of atherosclerosis. In addition, HDL particles of different sizes were found to have heterogeneous associations with ASCVD between the sex in the current study. It's likely that the composition of lipid or lipid subfraction profiles may differ by sex, reflecting the impact of prevalent diseases as well as broader environmental and lifestyle factors. Further research is required to explore the detailed mechanisms behind these sex differences in ASCVD risk related to blood lipids.

6.4.6 Strength and limitations

The current study systematically evaluates a range of NMR lipid-related traits from a clinical perspective (using composite measures similar to chemical assays) and at a more refined resolution (using lipid subfractions of varying sizes and potential biological functions) in a large, understudied Latin American population with comprehensive and reliable recorded baseline and resurvey data. The current study is one of the largest population-based metabolomic studies (~120,000 participants) especially conducted in the Latin American population, providing more statistical power and more accurate risk estimates. The repetitive measures of NMR data in resurvey allowed us to correct for the regression dilution bias that may underestimate the association of interest. The long follow-up period (~20 years) and large sample size also enable us to account for potential reverse causality by excluding participants with chronic diseases at baseline and conducting analyses excluding the first several years of follow-up. The baseline usage of lipid-lowering drugs in this population was low (albeit comparable with other countries around that time), which means the plasma lipid levels were largely unaffected by those drugs. Previous study of NMR lipid traits suggested that the widespread use of lipid-lowering drugs may distort the association with cardiovascular endpoints.[75] Given multiple lipid traits are inherently correlated, we also accounted for effects of other lipids to assess the relatively independent effect of lipid of interest. Overall, the current study provides robust and comprehensive evidence from a resource-limited setting for the prevention of ASCVD by modifying lipid levels, prioritizing ApoB as the most accurate,

direct, and reliable measure of atherogenic risk attributable to the ApoB-containing lipoproteins, especially in populations with high levels of obesity, diabetes, and TG levels.

However, the findings from this study should be interpreted with caution in the context of its limitations. First, although several previous studies indicated that the association between lipid traits and cardiovascular events varied by vascular domain (e.g., CHD, ischemic stroke, intracerebral hemorrhage, and PAD) and ethnicity, almost all ASCVD deaths were consisted of CHD, which limited the ability to conduct stratified analyses to explore potential heterogeneity. The exclusive use of Latin American participants might also limit the generalizability of our findings to other ethnic groups and populations. Also, there should be caution in extrapolating current findings to other settings in Latin America. Second, although the study outcomes were ASCVD events most relevant to the atherosclerotic plaque pathogenesis, the only available endpoint was cause-specific mortality. The lack of incident CVD events may limit the statistical power and the ability to conduct the above mentioned domain-specific analyses. Third, although we adjusted a set of covariates to account for potential confounding, residual confounding cannot be ruled out in this observational study. Mendelian randomization that is less susceptible to potential biases will be used in next chapter to explore the causality of identified significant associations. Finally, the attributed causes of death were ascertained by death certificates only, however, this tend to be accurate and reliable in Mexico. Finally, it should be noted that the associations between lipids and ASCVD in men aged 35–59 years showed a somewhat opposite direction compared to the overall trend or the direction expected based on prior studies. Potential underlying reasons may include that this group, recruited at home mostly during working hours, may not be fully representative of their peers. Excluding the first 10 years of follow-up, which partially reversed some of the opposite associations, may suggest that the younger men recruited to MCPS could have a higher risk of disease outcomes and frailty.

6.5 Conclusion

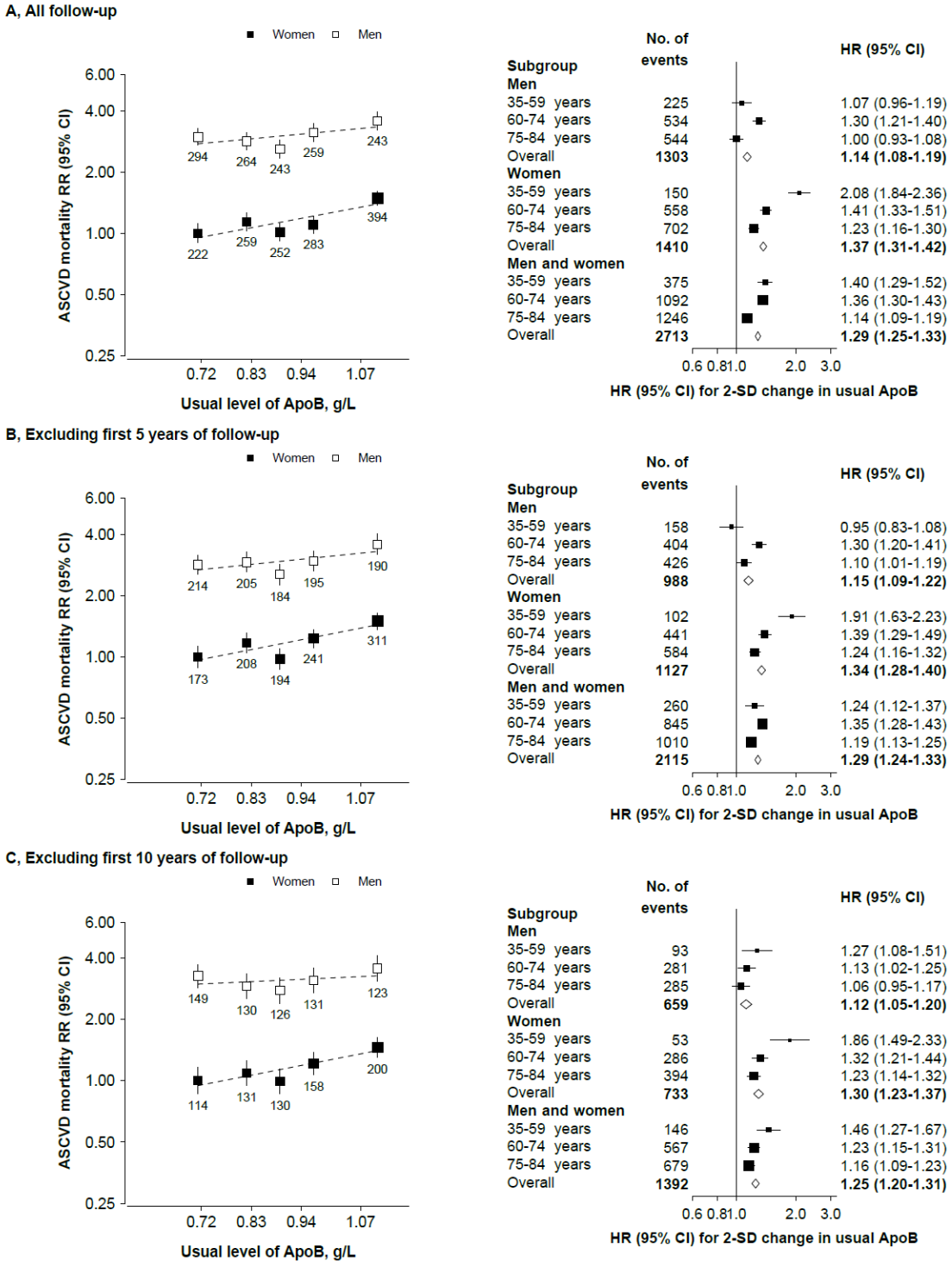
Based on a large Latin American population from Mexico City, the current findings highlight plasma ApoB (reflecting the concentration/number of ApoB containing particles) as being closely correlated with the risk of ASCVD. For specific types of lipoprotein particles (both LDL and HDL), their subfractions with varying sizes seemed to differ in their effects on risk. Direct measures of the number of different lipoproteins and their subtypes may be more relevant to risk than the total amounts of cholesterol within these particles, especially in populations with high levels of obesity and TG levels. These findings will facilitate targeted research on lipid subfractions and inform prevention and care strategies for CVD.

Table 6.1. Baseline characteristics of 107 504 participants without prior chronic disease (other than diabetes) and aged 35-84 at recruitment by groups of LDL-C

	Baseline lipid groups (tenths of distribution)										Overall (n=107 504)
	I (n=10 698)	II (n=10 698)	III (n=10 699)	IV (n=10 698)	V (n=10 698)	VI (n=10 699)	VII (n=10 698)	VIII (n=10 699)	IX (n=10 698)	X (n=10 698)	
Blood lipid and apolipoprotein											
LDL-C, mmol/L	0.9 (0.2)	1.2 (0.1)	1.4 (0.0)	1.5 (0.0)	1.7 (0.0)	1.8 (0.0)	1.9 (0.0)	2.0 (0.0)	2.2 (0.1)	2.6 (0.3)	1.7 (0.5)
HDL-C, mmol/L	0.8 (0.2)	0.9 (0.2)	1.0 (0.2)	1.0 (0.2)	1.0 (0.2)	1.0 (0.2)	1.1 (0.2)	1.1 (0.2)	1.1 (0.2)	1.2 (0.2)	1.0 (0.2)
TG, mmol/L	1.4 (0.8)	1.4 (0.6)	1.4 (0.6)	1.5 (0.6)	1.5 (0.6)	1.6 (0.6)	1.6 (0.6)	1.7 (0.6)	1.8 (0.7)	2.0 (0.7)	1.6 (0.7)
ApoB, g/L	0.6 (0.1)	0.7 (0.1)	0.8 (0.1)	0.8 (0.1)	0.9 (0.1)	0.9 (0.1)	1.0 (0.1)	1.0 (0.1)	1.1 (0.1)	1.3 (0.2)	0.9 (0.2)
ApoA1, g/L	1.1 (0.2)	1.1 (0.2)	1.2 (0.2)	1.2 (0.2)	1.2 (0.2)	1.2 (0.2)	1.3 (0.2)	1.3 (0.2)	1.3 (0.2)	1.4 (0.2)	1.2 (0.2)
Age and sex											
Age, years	52 (13)	51 (13)	51 (13)	51 (13)	51 (12)	51 (12)	51 (12)	51 (12)	52 (12)	53 (11)	51 (12)
Male	3983 (37%)	3667 (34%)	3634 (34%)	3495 (33%)	3518 (33%)	3492 (33%)	3554 (33%)	3447 (32%)	3345 (31%)	3230 (30%)	35 518 (33%)
Blood pressure											
Systolic blood pressure, mmHg	128 (17)	127 (16)	127 (17)	127 (16)	127 (17)	127 (16)	127 (16)	127 (17)	128 (16)	128 (17)	127 (17)
Diastolic blood pressure, mmHg	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	84 (10)	83 (10)
Taking blood pressure lowering medication	1495 (14%)	1524 (14%)	1401 (13%)	1500 (14%)	1419 (13%)	1530 (14%)	1513 (14%)	1557 (15%)	1642 (15%)	1769 (17%)	15 441 (14%)
Socio-economic status and lifestyle behaviours											
Resident of Coyoacán	4685 (44%)	4538 (42%)	4576 (43%)	4636 (43%)	4698 (44%)	4771 (45%)	4931 (46%)	4959 (46%)	4897 (46%)	4869 (46%)	47 683 (44%)
University/college educated	1307 (12%)	1505 (14%)	1541 (14%)	1590 (15%)	1703 (16%)	1767 (17%)	1788 (17%)	1819 (17%)	1885 (18%)	1932 (18%)	16 897 (16%)
Current smoker	3308 (31%)	3375 (32%)	3414 (32%)	3468 (32%)	3431 (32%)	3454 (32%)	3510 (33%)	3503 (33%)	3502 (33%)	3430 (32%)	34 552 (32%)
Current drinker	8048 (75%)	7966 (74%)	7968 (74%)	7958 (74%)	8047 (75%)	7929 (74%)	8009 (75%)	8059 (75%)	8095 (76%)	8041 (75%)	80 512 (75%)
Any regular leisure-time physical activity	1996 (19%)	2110 (20%)	2194 (21%)	2240 (21%)	2402 (22%)	2484 (23%)	2573 (24%)	2576 (24%)	2666 (25%)	2752 (26%)	24 089 (22%)
Physical measurements											
BMI, kg/m^2	28.9 (4.9)	29.0 (5.1)	29.0 (5.1)	29.1 (5.0)	28.9 (4.9)	29.0 (5.0)	29.0 (4.9)	28.9 (4.8)	29.0 (4.7)	28.8 (4.6)	29.0 (4.9)
Waist circumference, cm	95 (12)	94 (12)	94 (12)	94 (12)	94 (12)	94 (12)	94 (11)	94 (11)	94 (11)	94 (11)	94 (11)
Hip circumference, cm	104 (10)	105 (11)	105 (11)	105 (11)	104 (10)	105 (11)	105 (10)	105 (10)	105 (10)	104 (10)	105 (10)
Diabetes and glycosylated hemoglobin											
Diabetes status											
No diabetes	8523 (80%)	8733 (82%)	8814 (82%)	8892 (83%)	8973 (84%)	8938 (84%)	8914 (83%)	8858 (83%)	8766 (82%)	8509 (80%)	88 314 (82%)
Undiagnosed diabetes	612 (6%)	588 (5%)	556 (5%)	541 (5%)	510 (5%)	470 (4%)	495 (5%)	479 (4%)	475 (4%)	463 (4%)	5227 (5%)
Diagnosed diabetes with HbA1c <9%	659 (6%)	563 (5%)	534 (5%)	521 (5%)	505 (5%)	498 (5%)	470 (4%)	518 (5%)	531 (5%)	533 (5%)	5375 (5%)
Diagnosed diabetes with HbA1c <11%	904 (8%)	814 (8%)	795 (7%)	744 (7%)	710 (7%)	793 (7%)	819 (8%)	844 (8%)	926 (9%)	1193 (11%)	8588 (8%)
Medication use											
Any anti-thrombotic	263 (2%)	267 (2%)	235 (2%)	246 (2%)	245 (2%)	277 (3%)	271 (3%)	258 (2%)	268 (3%)	281 (3%)	2622 (2%)
Any antidiabetic	1272 (12%)	1138 (11%)	1106 (10%)	1016 (9%)	979 (9%)	1014 (9%)	1007 (9%)	1037 (10%)	1098 (10%)	1302 (12%)	11 051 (10%)

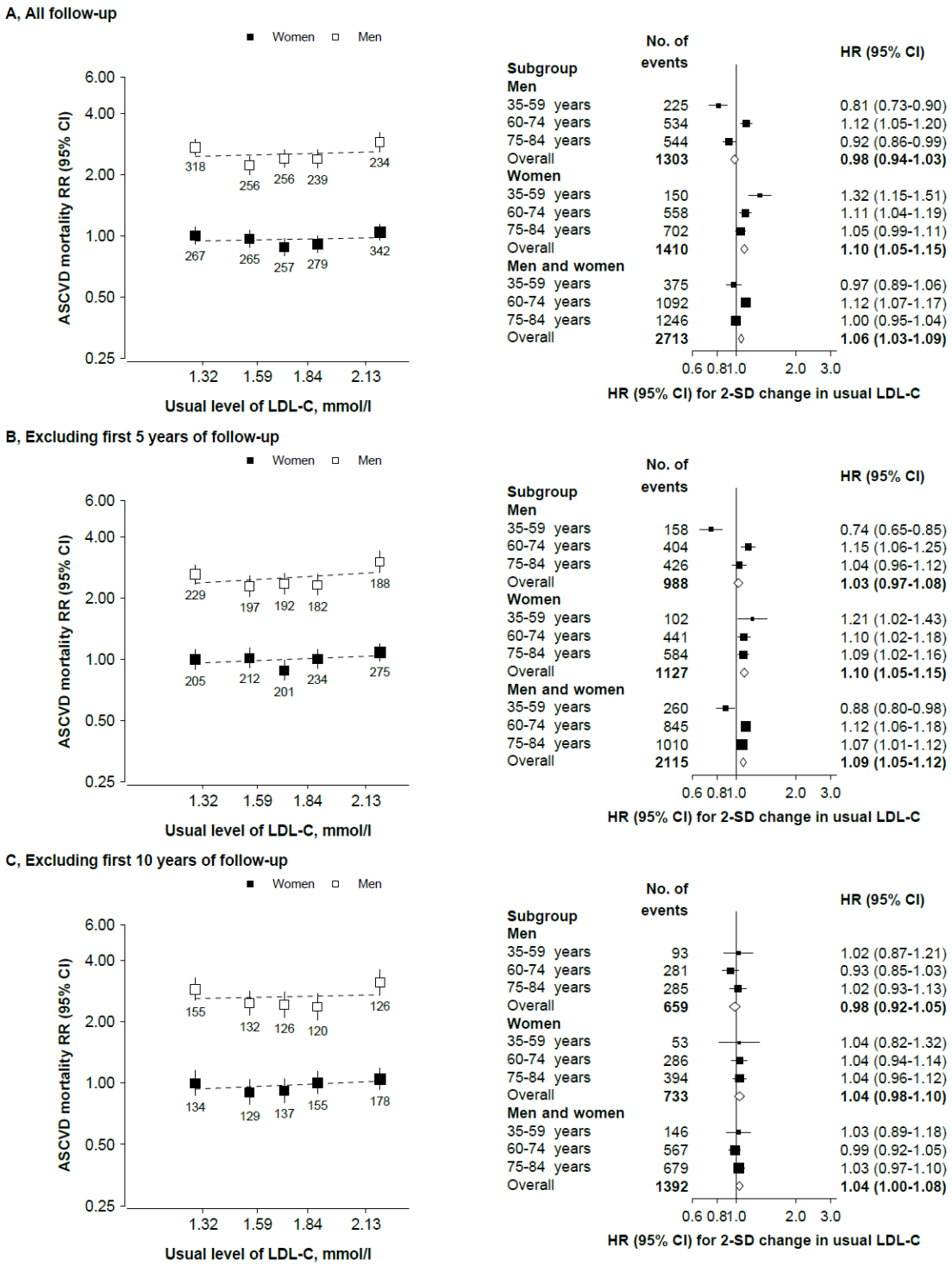
Mean (SD) or n (column %) shown. BMI=Body mass index. Table excludes those with prior chronic disease at recruitment (ischemic heart disease, stroke, chronic kidney disease, cirrhosis, cancer, emphysema), missing or out-of-range data on any analysis covariate (sex, district of residence, educational level attained, anthropometry, smoking status, alcohol intake, leisure time physical activity, diabetes status), uncertain follow-up, or missing or out-of-range blood lipid.

Figure 6.1. The association between ApoB and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period



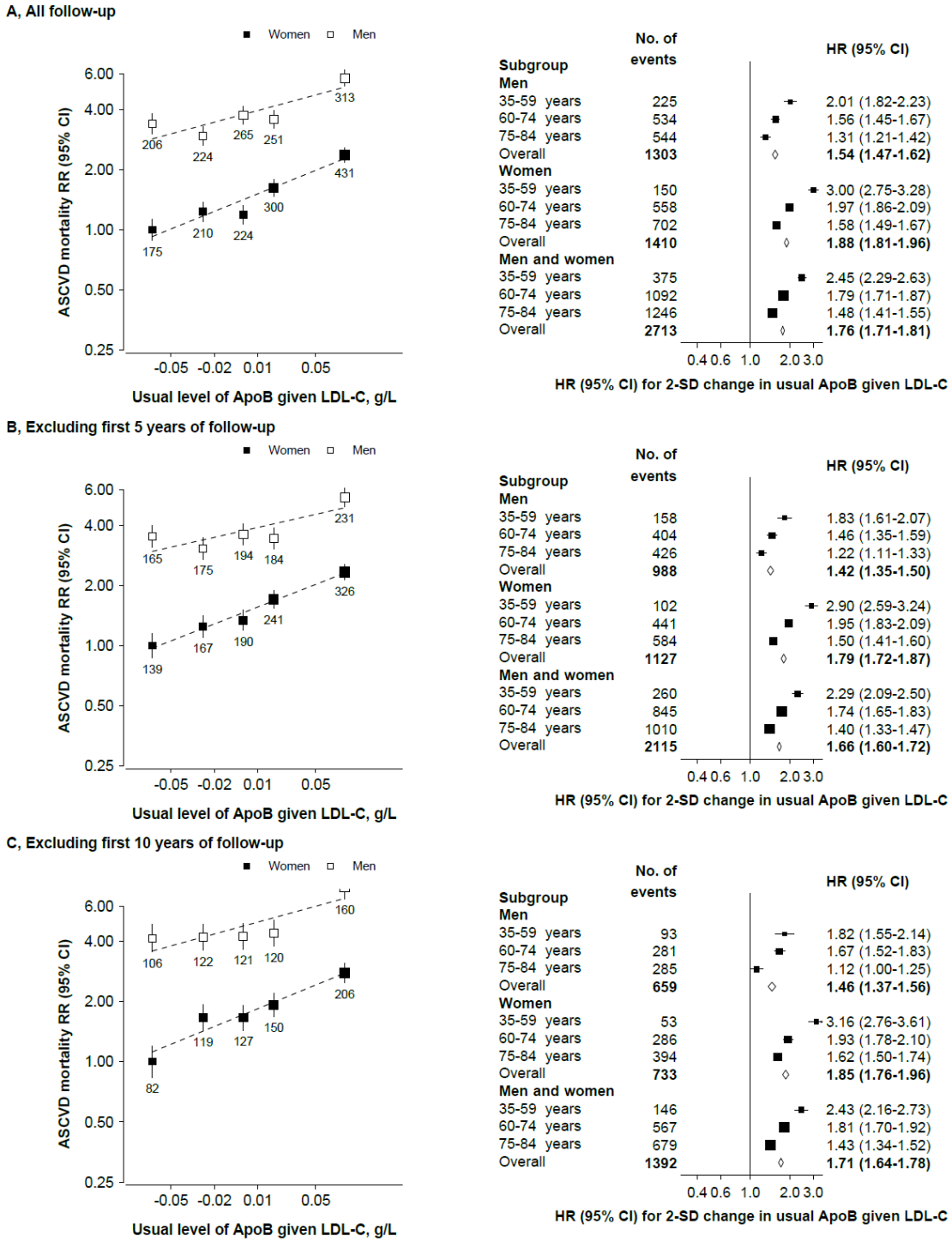
Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. 2-SD change in ApoB correspond to 0.4 g/l. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.2. The association between LDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period



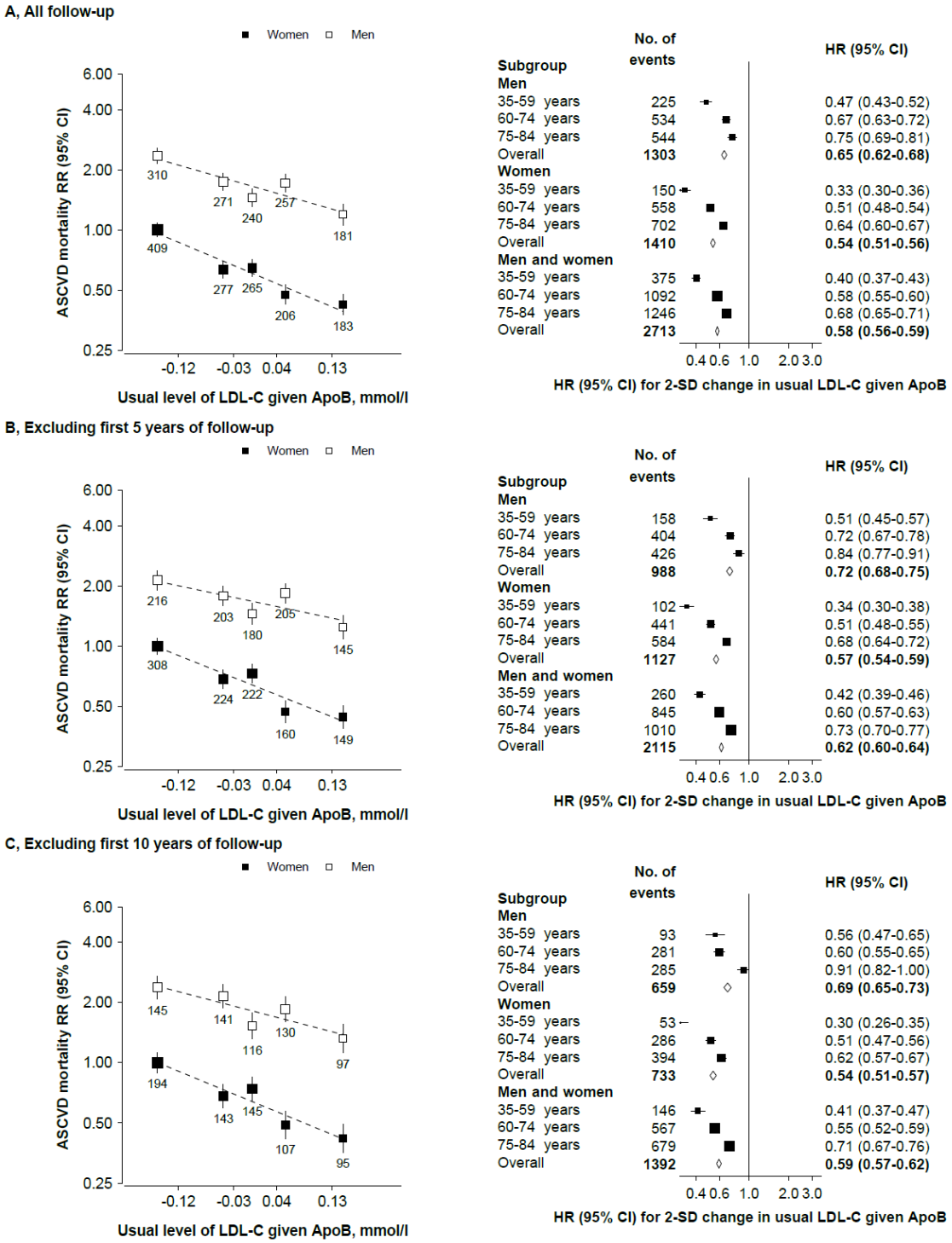
Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. 2-SD change in LDL-C correspond to 1 mmol/l. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.3. The association between ApoB given LDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period



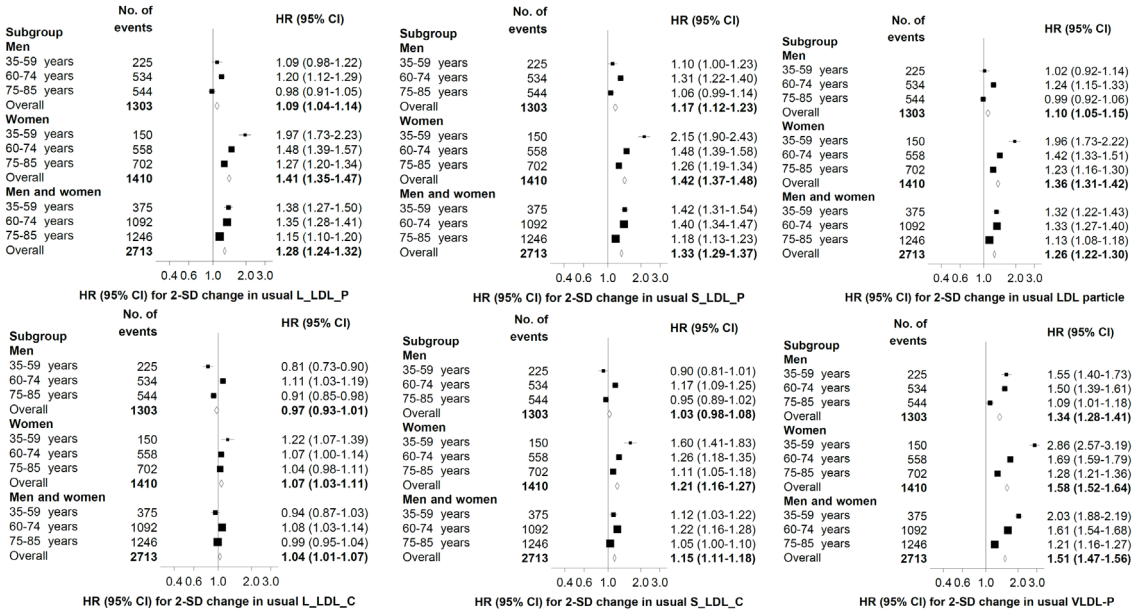
Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. 2-SD change in ApoB given LDL-C correspond to 0.14 g/l. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.4. The association between LDL-C given ApoB and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period



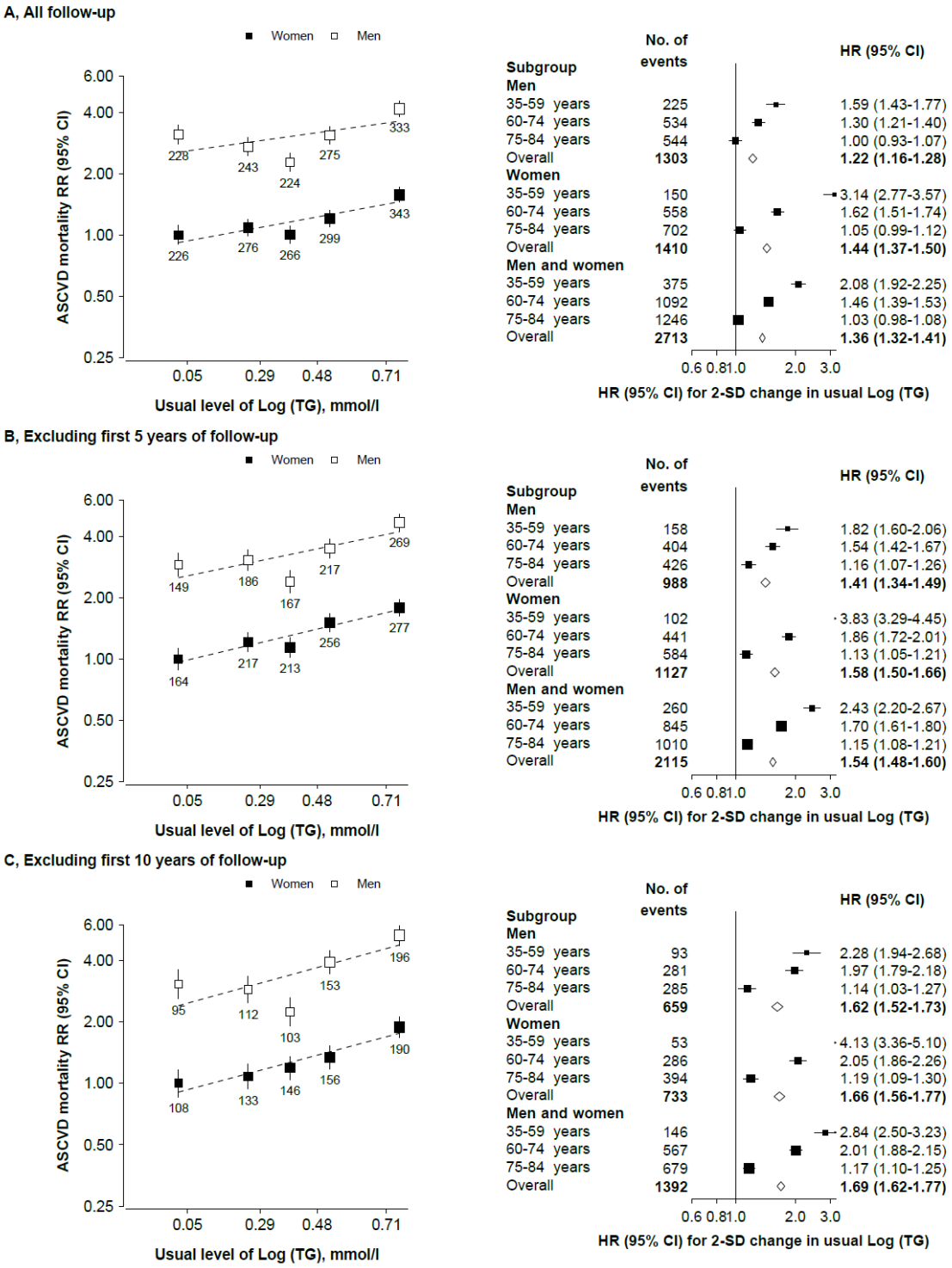
Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. 2-SD change in LDL-C given ApoB correspond to 0.32 mmol/l. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.5. Age- and sex-specific association of usual LDL-related subfractions and ASCVD mortality during full follow-up period



Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. Top-left, L-LDL-P; Top-mid, S-LDL-P; Top-right, LDL-P; Bottom-left, L-LDL-C; Bottom-mid, S-LDL-C; Bottom-right, VLDL-P. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

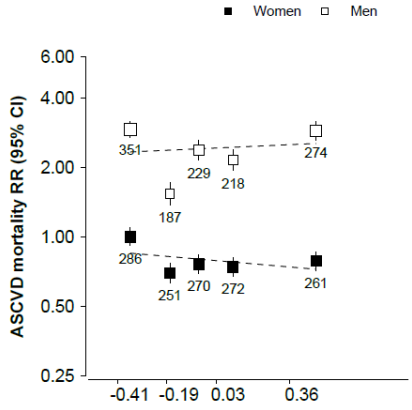
Figure 6.6. The association between TG and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period

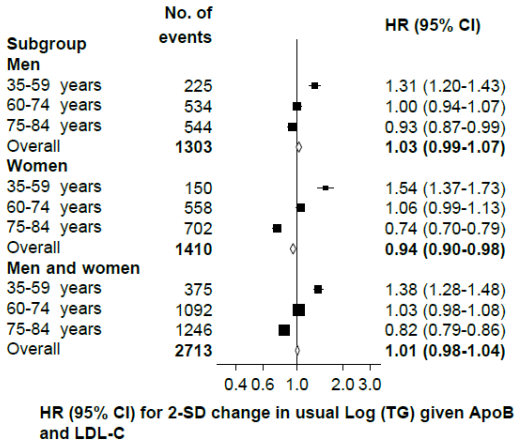


Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. 2-SD change in log transformed TG correspond to 0.80 mmol/l. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

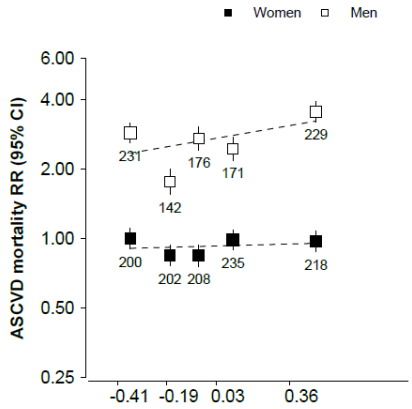
Figure 6.7. The association between TG given ApoB and LDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period

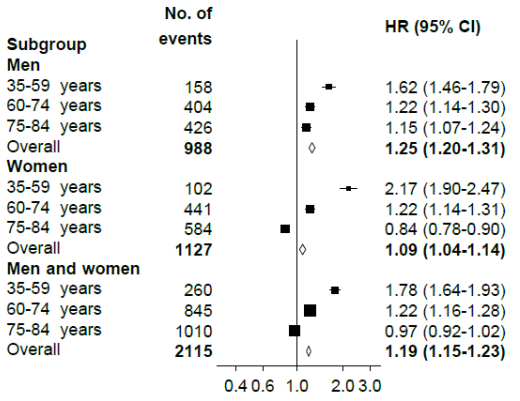
A, All follow-up



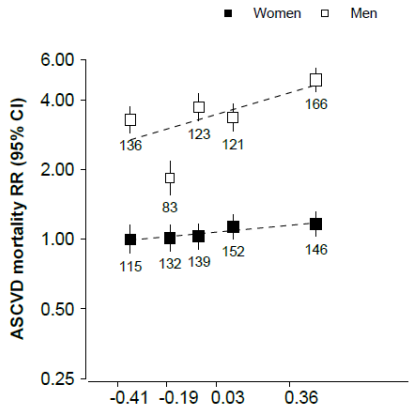


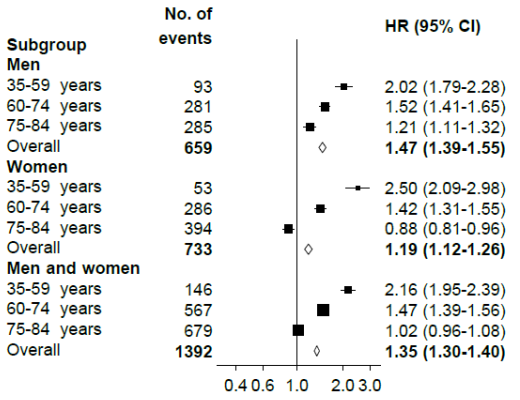
B, Excluding first 5 years of follow-up





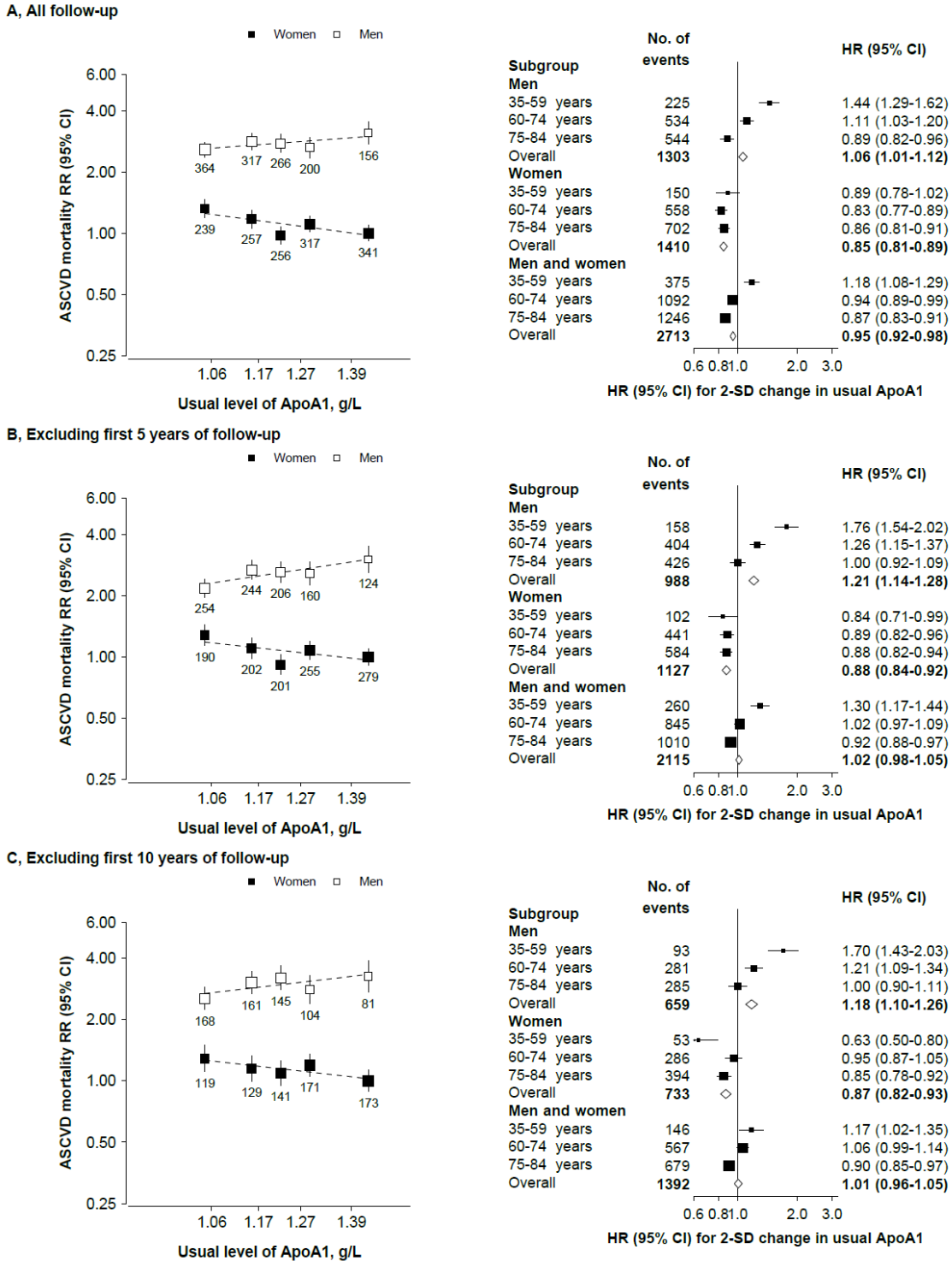
C, Excluding first 10 years of follow-up





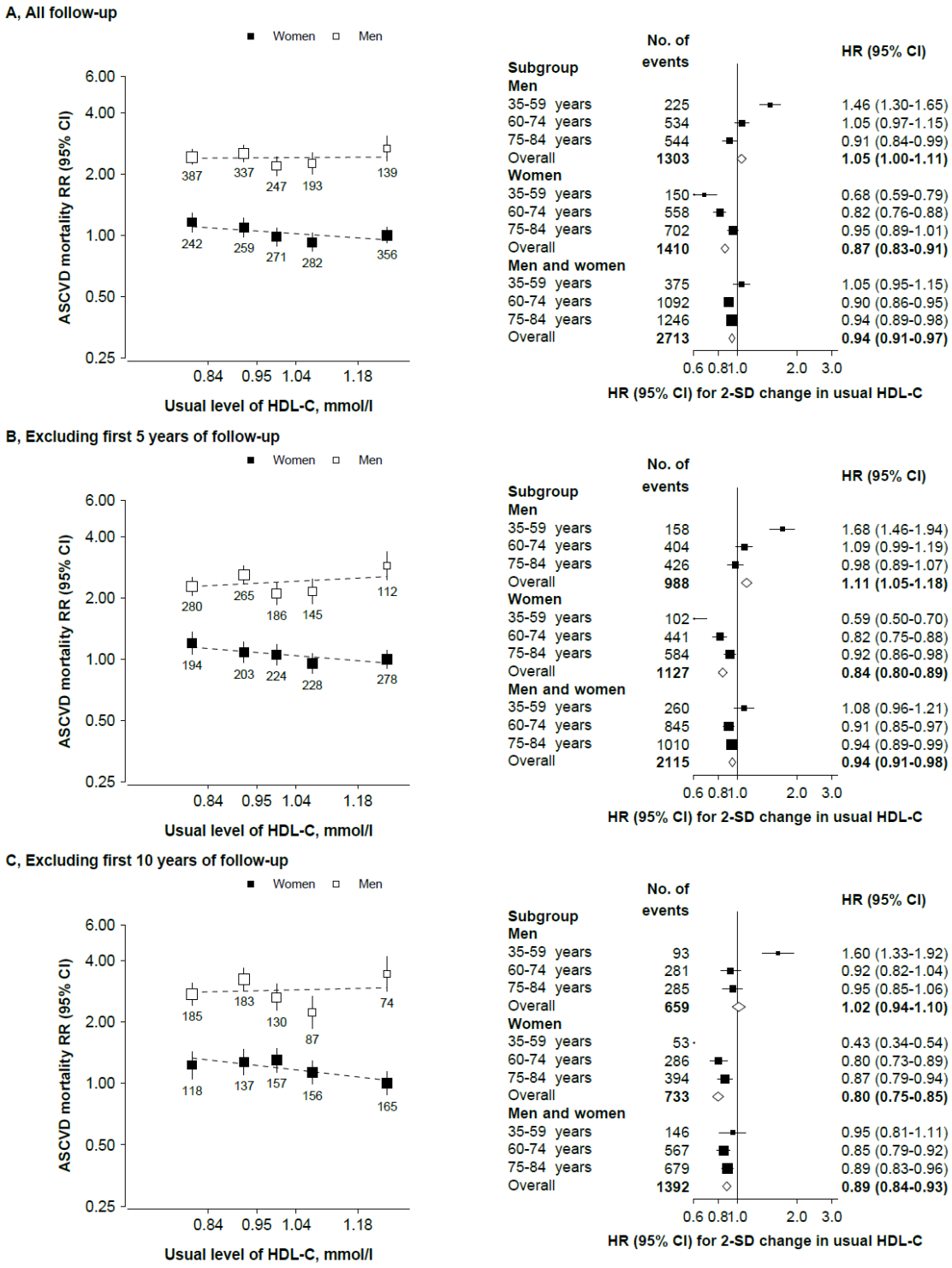
Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. 2-SD change in TG given ApoB and LDL-C correspond to 1.06 mmol/l. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.8. The association between ApoA1 and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period



Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. 2-SD change in ApoA1 correspond to 0.4 g/l. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

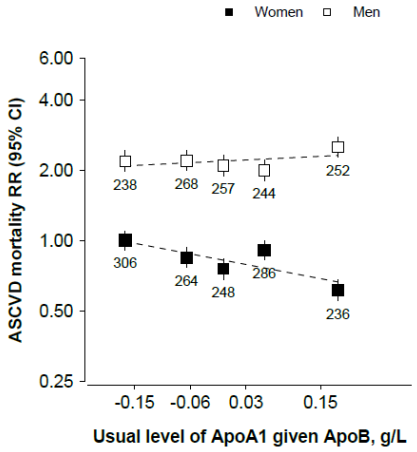
Figure 6.9. The association between HDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period

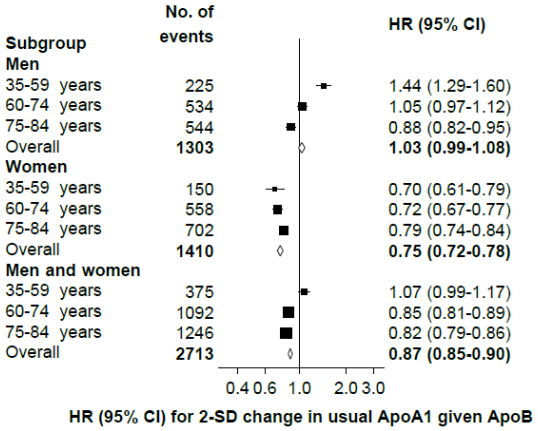


Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. 2-SD change in HDL-C correspond to 0.44 mmol/l. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

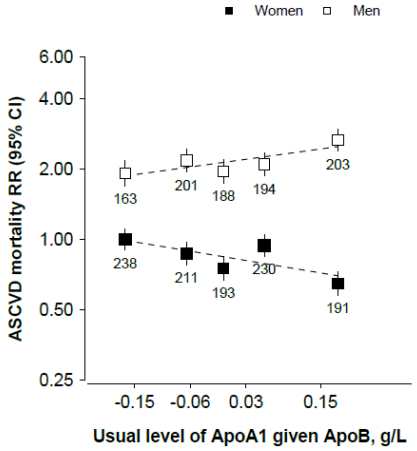
Figure 6.10. The association between ApoA1 given ApoB and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period

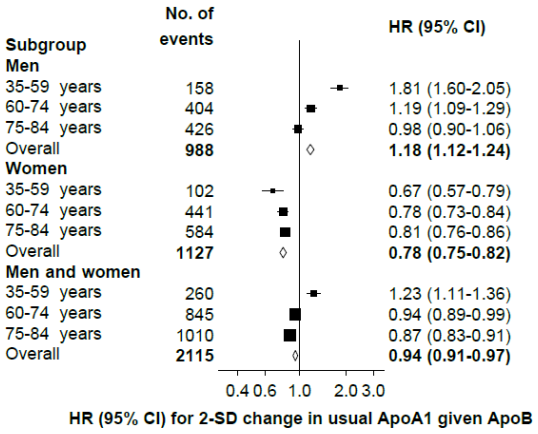
A, All follow-up



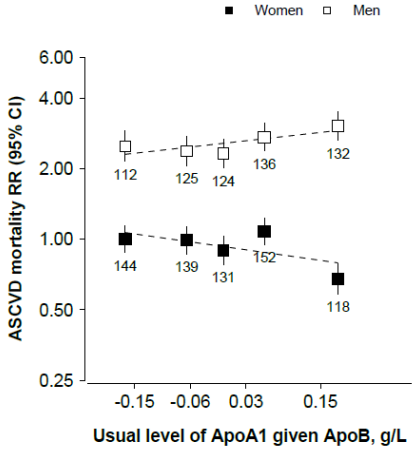


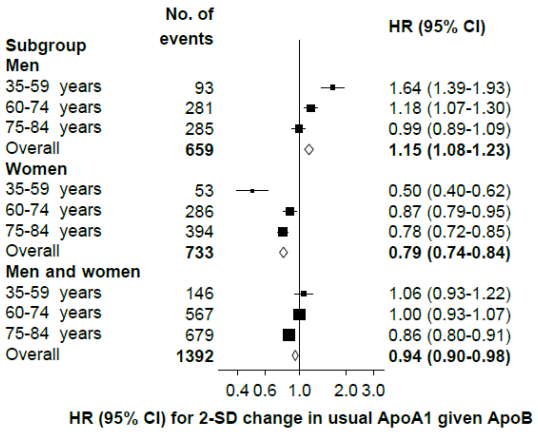
B, Excluding first 5 years of follow-up





C, Excluding first 10 years of follow-up

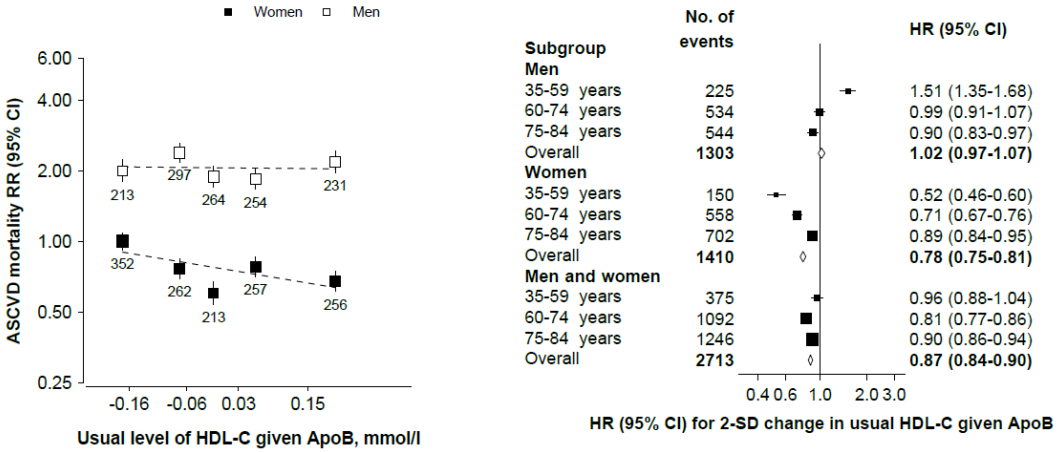




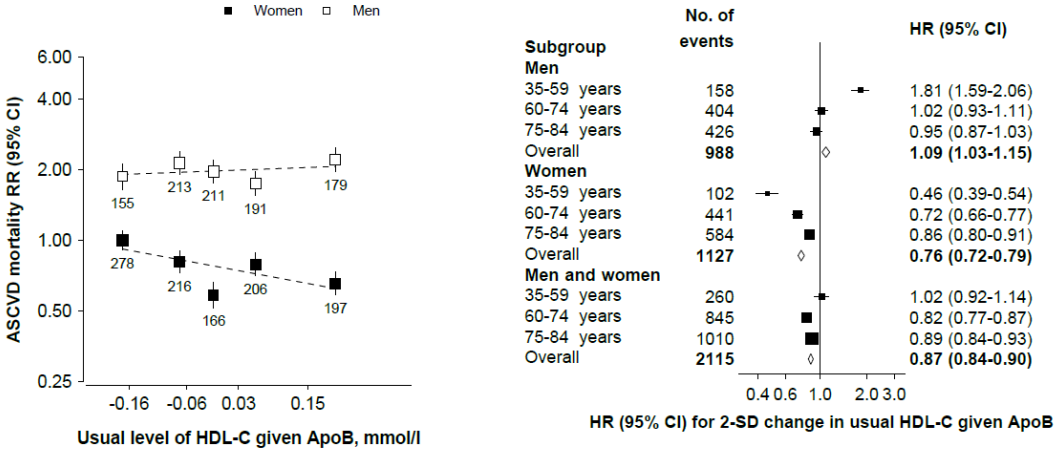
Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. 2-SD change in ApoA1 given ApoB correspond to 0.40 g/l. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.11. The association between HDL-C given ApoB and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period

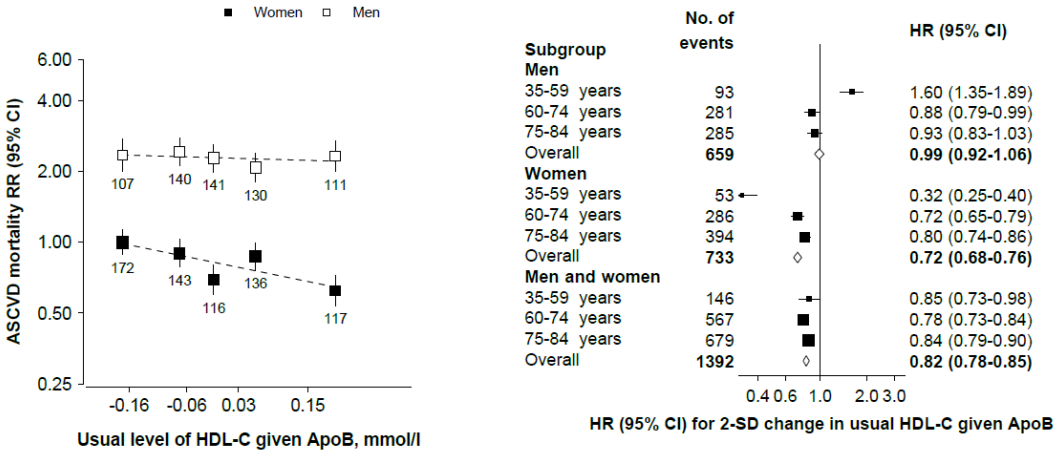
A, All follow-up



B, Excluding first 5 years of follow-up

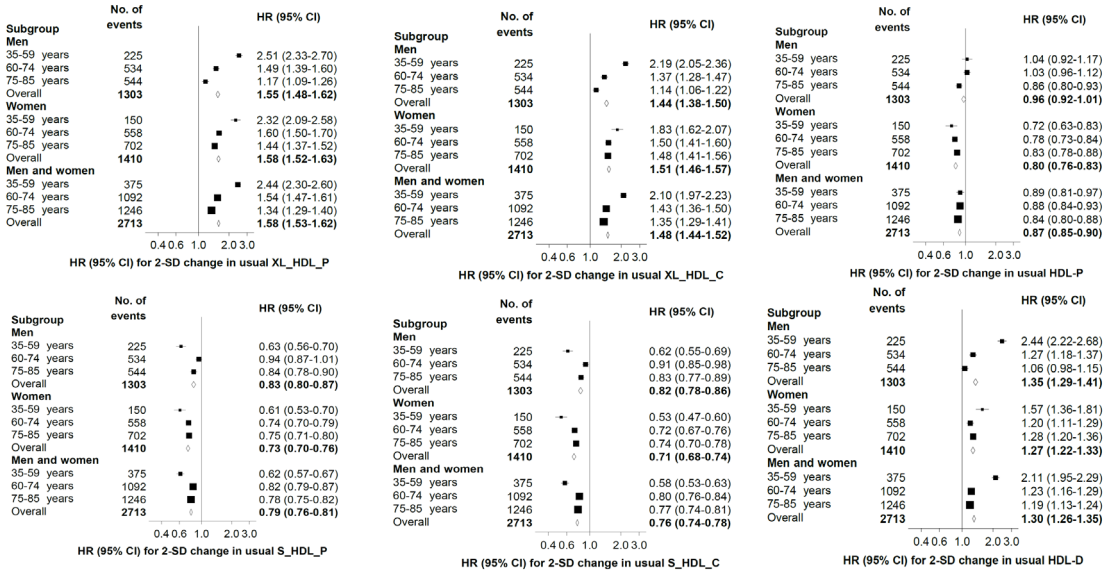


C, Excluding first 10 years of follow-up



Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. 2-SD change in HDL-C given ApoB correspond to 0.42 g/l. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.12. The association between HDL-related subfractions and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period



Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Reference

1 Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet (London, England)* **396**, 1204-1222, doi:10.1016/s0140-6736(20)30925-9 (2020).

2 Roth, G. A. *et al.* Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019: Update From the GBD 2019 Study. *Journal of the American College of Cardiology* **76**, 2982-3021, doi:10.1016/j.jacc.2020.11.010 (2020).

3 Yusuf, S. *et al.* Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet (London, England)* **364**, 937-952, doi:10.1016/s0140-6736(04)17018-9 (2004).

4 McQueen, M. J. *et al.* Lipids, lipoproteins, and apolipoproteins as risk markers of myocardial infarction in 52 countries (the INTERHEART study): a case-control study. *Lancet (London, England)* **372**, 224-233, doi:10.1016/s0140-6736(08)61076-4 (2008).

5 Borén, J. *et al.* Low-density lipoproteins cause atherosclerotic cardiovascular disease: pathophysiological, genetic, and therapeutic insights: a consensus statement from the European Atherosclerosis Society Consensus Panel. *European heart journal* **41**, 2313-2330, doi:10.1093/eurheartj/ehz962 (2020).

6 Ference, B. A. *et al.* Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. *European heart journal* **38**, 2459-2472, doi:10.1093/eurheartj/ehx144 (2017).

7 Di Angelantonio, E. *et al.* Major lipids, apolipoproteins, and risk of vascular disease. *Jama* **302**, 1993-2000, doi:10.1001/jama.2009.1619 (2009).

8 Mach, F. *et al.* 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk. *European heart journal* **41**, 111-188, doi:10.1093/eurheartj/ehz455 (2020).

9 Martin, S. S. *et al.* Comparison of a novel method vs the Friedewald equation for estimating low-density lipoprotein cholesterol levels from the standard lipid profile. *Jama* **310**, 2061-2068 (2013).

10 Plummer, M. Improved estimates of floating absolute risk. *Statistics in medicine* **23**, 93-104, doi:10.1002/sim.1485 (2004).

11 Clarke, R. *et al.* Underestimation of risk associations due to regression dilution in long-term follow-up of prospective studies. *American journal of epidemiology* **150**, 341-353, doi:10.1093/oxfordjournals.aje.a010013

(1999).
12 MacMahon, S. *et al.* Blood pressure, stroke, and coronary heart disease. Part 1, Prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *Lancet (London, England)* **335**, 765-774, doi:10.1016/0140-6736(90)90878-9 (1990).
13 Welsh, C. *et al.* Comparison of Conventional Lipoprotein Tests and Apolipoproteins in the Prediction of Cardiovascular Disease. *Circulation* **140**, 542-552, doi:10.1161/circulationaha.119.041149 (2019).
14 Sniderman, A. D. *et al.* Apolipoprotein B Particles and Cardiovascular Disease: A Narrative Review. *JAMA cardiology* **4**, 1287-1295, doi:10.1001/jamacardio.2019.3780 (2019).
15 Adiels, M. *et al.* Overproduction of large VLDL particles is driven by increased liver fat content in man. *Diabetologia* **49**, 755-765, doi:10.1007/s00125-005-0125-z (2006).
16 Griffin, B. A. *et al.* Rapid isolation of low density lipoprotein (LDL) subfractions from plasma by density gradient ultracentrifugation. *Atherosclerosis* **83**, 59-67, doi:10.1016/0021-9150(90)90131-2 (1990).
17 Mora, S. *et al.* Lipoprotein particle profiles by nuclear magnetic resonance compared with standard lipids and apolipoproteins in predicting incident cardiovascular disease in women. *Circulation* **119**, 931-939, doi:10.1161/circulationaha.108.816181 (2009).
18 Lawler, P. R. *et al.* Discordance between Circulating Atherogenic Cholesterol Mass and Lipoprotein Particle Concentration in Relation to Future Coronary Events in Women. *Clinical chemistry* **63**, 870-879, doi:10.1373/clinchem.2016.264515 (2017).
19 Parish, S. *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. *European heart journal* **30**, 2137-2146, doi:10.1093/eurheartj/ehp221 (2009).
20 Akbaraly, T. *et al.* Association of circulating metabolites with healthy diet and risk of cardiovascular disease: analysis of two cohort studies. *Scientific reports* **8(1)**, 8620 (2018).
21 Ference, B. A. *et al.* Association of Genetic Variants Related to CETP Inhibitors and Statins With Lipoprotein Levels and Cardiovascular Risk. *Jama* **318**, 947-956, doi:10.1001/jama.2017.11467 (2017).
22 Marston, N. A. *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 cardiology* **7**, 250-256, doi:10.1001/jamacardio.2021.5083 (2022).
23 Sniderman, A. D., Navar, A. M. & Thanassoulis, G. Apolipoprotein B vs

Low-Density Lipoprotein Cholesterol and Non–High-Density Lipoprotein Cholesterol as the Primary Measure of Apolipoprotein B Lipoprotein-Related Risk: The Debate Is Over. *JAMA cardiology* **7**, 257-258, doi:10.1001/jamacardio.2021.5080 (2022).

24 Miller, M. *et al.* Triglycerides and cardiovascular disease: a scientific statement from the American Heart Association. *Circulation* **123**, 2292-2333, doi:10.1161/CIR.0b013e3182160726 (2011).

25 Ballantyne, C. M. *et al.* Influence of low high-density lipoprotein cholesterol and elevated triglyceride on coronary heart disease events and response to simvastatin therapy in 4S. *Circulation* **104**, 3046-3051, doi:10.1161/hc5001.100624 (2001).

26 Manninen, V. *et al.* Joint effects of serum triglyceride and LDL cholesterol and HDL cholesterol concentrations on coronary heart disease risk in the Helsinki Heart Study. Implications for treatment. *Circulation* **85**, 37-45, doi:10.1161/01.cir.85.1.37 (1992).

27 Miller, M. *et al.* Impact of triglyceride levels beyond low-density lipoprotein cholesterol after acute coronary syndrome in the PROVE IT-TIMI 22 trial. *Journal of the American College of Cardiology* **51**, 724-730, doi:10.1016/j.jacc.2007.10.038 (2008).

28 Faergeman, O. *et al.* Plasma triglycerides and cardiovascular events in the Treating to New Targets and Incremental Decrease in End-Points through Aggressive Lipid Lowering trials of statins in patients with coronary artery disease. *The American journal of cardiology* **104**, 459-463, doi:10.1016/j.amjcard.2009.04.008 (2009).

29 Ginsberg, H. N. *et al.* Effects of combination lipid therapy in type 2 diabetes mellitus. *The New England journal of medicine* **362**, 1563-1574, doi:10.1056/NEJMoa1001282 (2010).

30 Boden, W. E. *et al.* Niacin in patients with low HDL cholesterol levels receiving intensive statin therapy. *The New England journal of medicine* **365**, 2255-2267, doi:10.1056/NEJMoa1107579 (2011).

31 Landray, M. J. *et al.* Effects of extended-release niacin with laropiprant in high-risk patients. *The New England journal of medicine* **371**, 203-212, doi:10.1056/NEJMoa1300955 (2014).

32 Robins, S. J. *et al.* Relation of gemfibrozil treatment and lipid levels with major coronary events: VA-HIT: a randomized controlled trial. *Jama* **285**, 1585-1591, doi:10.1001/jama.285.12.1585 (2001).

33 Guyton, J. R. *et al.* Relationship of lipoproteins to cardiovascular events: the AIM-HIGH Trial (Atherothrombosis Intervention in Metabolic Syndrome With Low HDL/High Triglycerides and Impact on Global Health Outcomes). *Journal of the American College of Cardiology* **62**, 1580-1584, doi:10.1016/j.jacc.2013.07.023 (2013).

34 Bucher, H. C., Hengstler, P., Schindler, C. & Meier, G. N-3 polyunsaturated fatty acids in coronary heart disease: a meta-analysis of randomized controlled trials. *The American journal of medicine* **112**, 298-304, doi:10.1016/s0002-9343(01)01114-7 (2002).

35 Rizos, E. C., Ntzani, E. E., Bika, E., Kostapanos, M. S. & Elisaf, M. S. Association between omega-3 fatty acid supplementation and risk of major cardiovascular disease events: a systematic review and meta-analysis. *Jama* **308**, 1024-1033, doi:10.1001/2012.jama.11374 (2012).

36 Yokoyama, M. *et al.* Effects of eicosapentaenoic acid on major coronary events in hypercholesterolaemic patients (JELIS): a randomised open-label, blinded endpoint analysis. *Lancet (London, England)* **369**, 1090-1098, doi:10.1016/s0140-6736(07)60527-3 (2007).

37 Ference, B. A. *et al.* Association of Triglyceride-Lowering LPL Variants and LDL-C-Lowering LDLR Variants With Risk of Coronary Heart Disease. *Jama* **321**, 364-373, doi:10.1001/jama.2018.20045 (2019).

38 Gofman, J. W., Young, W. & Tandy, R. Ischemic heart disease, atherosclerosis, and longevity. *Circulation* **34**, 679-697, doi:10.1161/01.cir.34.4.679 (1966).

39 Gordon, T., Castelli, W. P., Hjortland, M. C., Kannel, W. B. & Dawber, T. R. High density lipoprotein as a protective factor against coronary heart disease. The Framingham Study. *The American journal of medicine* **62**, 707-714, doi:10.1016/0002-9343(77)90874-9 (1977).

40 Castelli, W. P. *et al.* Incidence of coronary heart disease and lipoprotein cholesterol levels. The Framingham Study. *Jama* **256**, 2835-2838 (1986).

41 Madsen, C. M., Varbo, A. & Nordestgaard, B. G. Extreme high high-density lipoprotein cholesterol is paradoxically associated with high mortality in men and women: two prospective cohort studies. *European heart journal* **38**, 2478-2486, doi:10.1093/eurheartj/ehx163 (2017).

42 Ko, D. T. *et al.* High-Density Lipoprotein Cholesterol and Cause-Specific Mortality in Individuals Without Previous Cardiovascular Conditions: The CANHEART Study. *Journal of the American College of Cardiology* **68**, 2073-2083, doi:10.1016/j.jacc.2016.08.038 (2016).

43 Wilkins, J. T. *et al.* Coronary heart disease risks associated with high levels of HDL cholesterol. *Journal of the American Heart Association* **3**, e000519, doi:10.1161/jaha.113.000519 (2014).

44 Mazidi, M., Mikhailidis, D. P. & Banach, M. Associations between risk of overall mortality, cause-specific mortality and level of inflammatory factors with extremely low and high high-density lipoprotein cholesterol levels among American adults. *International journal of cardiology* **276**, 242-247, doi:10.1016/j.ijcard.2018.11.095 (2019).

45 Rosenson, R. S. *et al.* Cholesterol efflux and atheroprotection: advancing

the concept of reverse cholesterol transport. *Circulation* **125**, 1905-1919, doi:10.1161/circulationaha.111.066589 (2012).

46 Khera, A. V. *et al.* Cholesterol efflux capacity, high-density lipoprotein function, and atherosclerosis. *The New England journal of medicine* **364**, 127-135, doi:10.1056/NEJMoa1001689 (2011).

47 Guerin, M. *et al.* Association of Serum Cholesterol Efflux Capacity With Mortality in Patients With ST-Segment Elevation Myocardial Infarction. *Journal of the American College of Cardiology* **72**, 3259-3269, doi:10.1016/j.jacc.2018.09.080 (2018).

48 Soria-Florido, M. T. *et al.* Dysfunctional High-Density Lipoproteins Are Associated With a Greater Incidence of Acute Coronary Syndrome in a Population at High Cardiovascular Risk: A Nested Case-Control Study. *Circulation* **141**, 444-453, doi:10.1161/circulationaha.119.041658 (2020).

49 Li, X. M. *et al.* Paradoxical association of enhanced cholesterol efflux with increased incident cardiovascular risks. *Arteriosclerosis, thrombosis, and vascular biology* **33**, 1696-1705, doi:10.1161/atvbaha.113.301373 (2013).

50 Pownall, H. J., Rosales, C., Gillard, B. K. & Gotto, A. M., Jr. High-density lipoproteins, reverse cholesterol transport and atherogenesis. *Nature reviews. Cardiology* **18**, 712-723, doi:10.1038/s41569-021-00538-z (2021).

51 van der Steeg, W. A. *et al.* High-density lipoprotein cholesterol, high-density lipoprotein particle size, and apolipoprotein A-I: significance for cardiovascular risk: the IDEAL and EPIC-Norfolk studies. *Journal of the American College of Cardiology* **51**, 634-642, doi:10.1016/j.jacc.2007.09.060 (2008).

52 Zhao, Q. *et al.* A Mendelian randomization study of the role of lipoprotein subfractions in coronary artery disease. *eLife* **10**, doi:10.7554/eLife.58361 (2021).

53 Holmes, M. V. *et al.* Lipids, Lipoproteins, and Metabolites and Risk of Myocardial Infarction and Stroke. *Journal of the American College of Cardiology* **71**, 620-632, doi:10.1016/j.jacc.2017.12.006 (2018).

54 Aday, A. W. *et al.* Lipoprotein Particle Profiles, Standard Lipids, and Peripheral Artery Disease Incidence. *Circulation* **138**, 2330-2341, doi:10.1161/circulationaha.118.035432 (2018).

55 Würtz, P. *et al.* Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts. *Circulation* **131**, 774-785, doi:10.1161/circulationaha.114.013116 (2015).

56 Vojinovic, D. *et al.* Association of circulating metabolites in plasma or serum and risk of stroke: Meta-analysis from seven prospective cohorts. *Neurology* **96**, e1110-1123, doi:10.1212/wnl.0000000000011236 (2020).

57 El Harchaoui, K. *et al.* High-density lipoprotein particle size and

concentration and coronary risk. *Annals of internal medicine* **150**, 84-93, doi:10.7326/0003-4819-150-2-200901200-00006 (2009).

58 Mora, S., Glynn, R. J. & Ridker, P. M. High-density lipoprotein cholesterol, size, particle number, and residual vascular risk after potent statin therapy. *Circulation* **128**, 1189-1197, doi:10.1161/circulationaha.113.002671 (2013).

59 Singh, K. *et al.* Associations Between High-Density Lipoprotein Particles and Ischemic Events by Vascular Domain, Sex, and Ethnicity: A Pooled Cohort Analysis. *Circulation* **142**, 657-669, doi:10.1161/circulationaha.120.045713 (2020).

60 Briel, M. *et al.* Association between change in high density lipoprotein cholesterol and cardiovascular disease morbidity and mortality: systematic review and meta-regression analysis. *BMJ (Clinical research ed.)* **338**, b92, doi:10.1136/bmj.b92 (2009).

61 Barter, P. J. *et al.* Effects of torcetrapib in patients at high risk for coronary events. *The New England journal of medicine* **357**, 2109-2122, doi:10.1056/NEJMoa0706628 (2007).

62 Williams, S. A. *et al.* Improving Assessment of Drug Safety Through Proteomics: Early Detection and Mechanistic Characterization of the Unforeseen Harmful Effects of Torcetrapib. *Circulation* **137**, 999-1010, doi:10.1161/circulationaha.117.028213 (2018).

63 Schwartz, G. G. *et al.* Effects of dalcetrapib in patients with a recent acute coronary syndrome. *The New England journal of medicine* **367**, 2089-2099, doi:10.1056/NEJMoa1206797 (2012).

64 Lincoff, A. M. *et al.* Evacetrapib and Cardiovascular Outcomes in High-Risk Vascular Disease. *The New England journal of medicine* **376**, 1933-1942, doi:10.1056/NEJMoa1609581 (2017).

65 Bowman, L. *et al.* Effects of Anacetrapib in Patients with Atherosclerotic Vascular Disease. *The New England journal of medicine* **377**, 1217-1227, doi:10.1056/NEJMoa1706444 (2017).

66 Tall, A. R. & Rader, D. J. Trials and Tribulations of CETP Inhibitors. *Circulation research* **122**, 106-112, doi:10.1161/circresaha.117.311978 (2018).

67 Frikke-Schmidt, R. *et al.* Association of loss-of-function mutations in the ABCA1 gene with high-density lipoprotein cholesterol levels and risk of ischemic heart disease. *Jama* **299**, 2524-2532, doi:10.1001/jama.299.21.2524 (2008).

68 Voight, B. F. *et al.* Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. *Lancet (London, England)* **380**, 572-580, doi:10.1016/s0140-6736(12)60312-2 (2012).

69 Johannsen, T. H., Frikke-Schmidt, R., Schou, J., Nordestgaard, B. G. & Tybjærg-Hansen, A. Genetic inhibition of CETP, ischemic vascular disease

and mortality, and possible adverse effects. *Journal of the American College of Cardiology* **60**, 2041-2048, doi:10.1016/j.jacc.2012.07.045 (2012).

70 Haase, C. L. *et al.* LCAT, HDL cholesterol and ischemic cardiovascular disease: a Mendelian randomization study of HDL cholesterol in 54,500 individuals. *The Journal of clinical endocrinology and metabolism* **97**, E248-256, doi:10.1210/jc.2011-1846 (2012).

71 Roeters van Lennep, J. E. *et al.* Women, lipids, and atherosclerotic cardiovascular disease: a call to action from the European Atherosclerosis Society. *European heart journal* **44**, 4157-4173, doi:10.1093/eurheartj/ehad472 (2023).

72 Nordestgaard, B. G., Benn, M., Schnohr, P. & Tybjærg-Hansen, A. Nonfasting Triglycerides and Risk of Myocardial Infarction, Ischemic Heart Disease, and Death in Men and Women. *Jama* **298**, 299-308, doi:10.1001/jama.298.3.299 (2007).

73 Visseren, F. L. J. *et al.* 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice: Developed by the Task Force for cardiovascular disease prevention in clinical practice with representatives of the European Society of Cardiology and 12 medical societies With the special contribution of the European Association of Preventive Cardiology (EAPC). *European heart journal* **42**, 3227-3337, doi:10.1093/eurheartj/ehab484 (2021).

74 Arnett, D. K. *et al.* 2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Circulation* **140**, e596-e646, doi:doi:10.1161/CIR.0000000000000678 (2019).

75 Julkunen, H. *et al.* Atlas of plasma NMR biomarkers for health and disease in 118,461 individuals from the UK Biobank. *Nature communications* **14**, 604, doi:10.1038/s41467-023-36231-7 (2023).

CHAPTER 7: Genetic associations between major lipid traits and atherosclerotic cardiovascular mortality: Mendelian randomization study

Chapter summary

The current chapter aims to assess whether major lipid traits have a causal effect on ASCVD using Mendelian Randomization and then compare the results with those from observational analyses in Chapter 6. In univariable MR for the assessment of individual lipids or apolipoproteins, both LDL-C and ApoB were positively associated with the risk of ASCVD (HR per 2SD change: 1.31, 95% CI 1.03-1.67 and 1.71, 1.25-2.43, respectively), with no significant effect observed for any other lipid traits. In multivariable MR, after accounting for measured pleiotropy to estimate the direct effect of lipid traits, ApoB remained strongly predictive after adjustment for LDL-C (HR per 2SD change: 1.71, 1.13-2.59) whereas LDL-C became non-significant (1.06, 0.78-1.42) after adjustment for ApoB. After adjustment for ApoA1 and ApoB, HDL-C was *positively* related to risk (1.35, 1.01-1.79). Taken together with the observational findings, these results indicate that ApoB, rather than the cholesterol or TG contained in the ApoB particles *per se*, is the strongest causal lipid factor for ASCVD risk. The more cholesterol the HDL particles carry the higher the risk (e.g., smaller, denser HDL particles may be protective). The current findings do not support TG as a causal factor in the development of ASCVD. Overall, reducing the number of ApoB-containing particles (e.g., as achieved through wider use of statins) should remain the focus of efforts to reduce ASCVD risk in Mexican adults.

7.1 Background

As described in **Chapter 1**, evidence from genetic studies, prospective cohort studies, Mendelian randomization studies, and randomized clinical trials of lipid-lowering therapies in predominantly Western populations have suggested LDL and other apoB-containing lipoproteins (VLDL remnants, IDL, and Lp(a)) have a causal effect on ASCVD. However, the relationships of plasma TG and HDL-C levels with the risk of ASCVD has been controversial. Most of these studies have tended to rely on traditional lipid panel (TG, TC, LDL-C, and HDL-C) used in routine clinical practice. Given the complexities and inter-relatedness of lipid traits, a comprehensive NMR approach might be more preferable than traditional chemical assay to quantify the concentration and size of lipoprotein particles and their lipid contents. As described in **Chapter 2**, results of observational studies investigating the association between apoB-containing lipoproteins measured by NMR spectroscopy and the risk of ASCVD are in line with the current evidence based on traditional lipid panels. These associations were attenuated but remained significant in studies further adjusting for chemically measured lipids such as LDL-C and HDL-C. However, previous NMR studies are limited to participants of primarily European ancestry while sample sizes have been modest compared with studies of traditional lipids. Analyses of NMR data from the MCPS have shown positive associations of various NMR lipid traits (e.g. apoB-containing lipoproteins and their proxy, ApoB) with the risk of premature cardiovascular mortality (**Chapter 6**). Notably, specific lipoproteins (LDL and HDL) with varying particle sizes and biological functions, as well as their lipid contents seemed to differ in their effects on risk. Direct measurements of specific subtypes of lipoproteins particles may be more informative about risk than the cholesterol or TG content measured in traditional panels.

Conventional epidemiological associations between blood lipids and cardiovascular outcomes are vulnerable to confounding by lifestyle, socio-economic, and other associated metabolic factors. In addition, as many aspects of lipid metabolism are

affected by comorbidities including cardiovascular diseases, diabetes, and other chronic diseases, reverse causality, or a bi-directional relationship, may explain part of these associations. Furthermore, most previous studies have tended to rely on traditional lipid panels, meaning that the observed associations could be biased by measurement error (e.g. when LDL-C was not directly measured but was estimated by the Friedewald equation). Despite efforts to control for these limitations, risk estimates from conventional multivariable regression in previous observational studies of both traditional and NMR lipid panels are inevitably susceptible to residual or unmeasured confounding, measurement error, and reverse causality, making causal inference challenging. Mendelian randomization (MR) can be used as a useful approach to investigating the likely causal nature of the associations between NMR lipid traits and ASCVD.

MR analyses use genetic variants robustly associated with potentially modifiable risk factors (e.g. lipid traits) as proxies to explore their causal effects on diseases (e.g. ASCVD). As genotypes (or alleles) are randomly allocated at conception, they are largely free from confounding that could bias the typical observational findings. This, as well as the invariant nature of genetic variants, provides an analogy to randomized clinical trials, in which exposure is allocated randomly and is not affected by subsequent disease. The MR method reduces the potential bias that are common in conventional observational studies such as confounding, reverse causation and measurement error, *provided certain assumptions are met*. If the exposure causally affects the outcome, then genetic variants associated with the exposure should be expected to affect the outcome to a proportional degree, provided the variants do not influence the outcome through separate pathways other than the exposure of interest, a phenomenon known as horizontal pleiotropy. Under these conditions, genetic variants robustly associated with the exposure of interest and not associated with confounders of the exposure-outcome association may serve as instrumental variables to estimate the causal association with outcome. Conventional one-sample MR uses a single sample in which exposure,

outcome, and genetic instrument are measured within the same population, which may have power issues due to inadequate sample size. In a two-sample MR design, instruments can be taken from the summary level data of two independent large-scale, nonoverlapping genome-wide associations studies (GWAS), which greatly increase the scope and statistical power of MR.

Previous MR studies investigating the association between NMR lipid traits and ASCVD (**Chapter 2**) have used summary-level genetic data from individuals of mostly European ancestry (using methods of two-sample MR, three-sample MR, or MR-BMA), with a lack of study using individual participant data (one-sample MR). Three existing MR studies used genetic estimates from the current largest GWAS of NMR metabolism (N= 24,925), but even this sample size is still modest compared with GWAS from large consortia of traditional lipids, potentially leading to weak instrument bias (ie, where only a small proportion of the exposure variance is explained by genetic variation, potentially causing a bias towards the null in two-sample MR) and may have inadequate power to identify robust genetic instruments. These limitations of current MR studies on NMR lipids might result in null results and hinder causal inference. Furthermore, the genetic associations of NMR lipids with ASCVD reported in previous MR studies using summary statistics may not be directly comparable with observational associations reported in traditional epidemiological studies due to the use of different study populations and statistical methods (logistic regression for genetic associations and Cox regression for observational associations). Cox regression analysis can also be applied to calculate MR-derived risk estimates using genetic risk scores in the one-sample MR setting, from which estimates can be compared with observational estimates in the same study.

Overall, the potential causal relationship of NMR lipid traits with ASCVD remains largely unknown, especially in the Latin American population. With the availability of both genetic and NMR lipid data in a large prospective cohort of about 150,000 Latin American individuals, the MCPS will facilitate MR analyses on these associations and allow a direct

comparison between observational and genetic estimates derived from the same study population. The ongoing GWAS of NMR lipids in MCPS (**Chapter 5**) will also help to overcome the challenges of sample size as well as fill the current evidence gap and facilitate the identification of genetic variants influencing the levels of lipoprotein subfractions.

The current chapter aims to assess the causal nature of the relationship between blood lipid levels and ASCVD using genetic variants robustly associated with lipid traits as instrumental variables in an MR analysis of ASCVD risk in the MCPS. The objectives of this chapter are to: a) Select and validate SNPs associated with lipid levels from external large-scale lipid GWAS; b) Develop the GRS for different lipid traits and examine the association of the GRS with corresponding lipid levels in the MCPS; c) Use an MR approach to assess whether these lipid traits have a causal effect on ASCVD risk and assess the extent to which findings are robust to potential pleiotropy and bias from other sources; d) Compare the observational versus MR estimates of the association between NMR lipid traits and ASCVD derived from MCPS; and e) Based on the prior key observational findings of ApoB-containing particles and ApoB in Chapter 6, assess the relatively independent effects of lipid traits given ApoB using multivariable MR.

7.2 Methods

7.2.1 Genotyping

Of the 159,755 participants recruited into MCPS, a blood sample was successfully taken, processed and stored for 155,453 (97.3%), as detailed in **Chapter 3**. Of these, DNA was extracted for 146,068 (94.0%) and sent for genotyping and exome sequencing, of which 145,266 (99.5%) were successfully processed. DNA extracted at the UK Biocentre with a 260:280nm ratio of ≥1.8 and a 260:230nm ratio of 2.0-2.2 on UV-VIS spectroscopy was deemed uncontaminated. DNA yield was also checked with absorbance at 260nm. Pure DNA was suspended in TE buffer to generate samples containing >1.5μg at

concentrations of 10-200ng/μL (with 98% of samples meeting the more stringent criteria of >2μg at concentrations of 20-200ng/μL). Genomic DNA samples were transferred to the Regeneron Genetics Center (RGC) from the UK Biocenter, where about 650,000 variants were genotyped on the second-generation Illumina Global Screening Array beadchip. Chip variants provide whole-genome coverage of cardiometabolic trait-influencing variants and facilitated high quality imputation of variants with MAF>1% in the 1000 Genomes Project (which included Latin American ancestry individuals from Los Angeles, Medellin and Lima and Puerto Rica, and Spanish ancestry individuals from Iberia). The average genotype call rate per sample was 98.4% and 98.4% of samples had a call rate above 90%. Over 9,430,000 autosomal variants were imputed using diverse TOPMed reference panel including 47,159 European, 24,267 African, and 17,085 admixed American genomes, with imputation info score >0.3. The MCPS derived genomic principle components (PCs) and the first 7 PCs were used in the analyses to limit population stratification.

7.2.2 Genetic quality control and exclusions

Multiple quality control (QC) steps on individuals and SNPs were performed prior to inclusion in the present study. Samples failed initial QC steps if they met one or more of the following criteria: disagreement between genetically-determined and reported sex (n=1,827), low quality samples with call rates below 90% (n=2,276), genotyping chip variants discordant with exome data (n=44), genetically-identified duplicated sample (n=1,063), samples flagged by MCPS for exclusion (n=1,122), and other reasons (n=223). In total, 4,435 unique samples were excluded and data on 650,380 variants in 140,831 individuals remained. Following additional QC steps established in conjunction with the initial QC by RGC, data on 580,437 variants in the remaining 138,511 individuals were used for downstream analysis. Additional QC steps implemented in PLINK 1.9 included: checks for consistency of genotypes in sex chromosomes (steps 1-4), sample- and variant-level missingness filters (step 5-6), the Hardy-Weinberg equilibrium exact

test applied to a set of 81,747 third degree unrelated samples identified from the initial relatedness analysis by PLINK and PRIMUS (step 7), setting genotypes with Mendel errors in nuclear families to missing (step 8), and the second round of steps 5-7 (step 9).

In addition to the participant exclusions applied in the above QC steps, participants were further excluded form the analyses in this chapter if they were aged ≥85 at recruitment, were using lipid-lowering medications, had missing or extreme values of selected covariates, had missing values of NMR lipid measures at baseline, had self-reported chronic diseases (cardiovascular disease, cancer, or chronic renal, hepatic, or pulmonary disease), or had an uncertain mortality match, following the same procedure of participant selection used in the observational analyses. After all exclusions, a total of 107,504 participants were included in the MR analyses (~120k participants had NMR measures at the time of analyses).

7.2.3 Phenotype definitions

Participants were followed through record linkage to Mexican Death Registry, which provided information on medically certified death using ICD-10 codes and personal identifying information (**Chapter 3**). The primary outcome was cardiovascular mortality primarily caused by atherosclerosis, including deaths due to ischaemic heart disease, ischaemic stroke, and peripheral arterial disease. The list of ICD-10 codes contributing to each of these mortality outcomes is shown in Chapter 6 (**Table S6.1**). At the time of analysis, mortality data was available up to December 2020. Although we included an array of 104 NMR measurements related to lipid traits as exposures in the observational analyses in Chapter 6, the relatively limited sample size for GWAS analyses of these NMR metabolites led to a low number of genetic variants reaching genome-wide significance (e.g., <10 for S-LDL-P). Consequently, these explained only a small proportion of variance in lipid subfractions which could lead to weak instrument bias if used to create instrumental variables for use in Mendelian randomization analyses at the

current stage. Therefore, instead, instrumental variables for major composite lipoprotein-related traits (LDL-C, HDL-C, TG, ApoB, and ApoA1) were used as the main exposures in this chapter, using SNPs and SNP weights identified by the vey large-scale GWAS that exist of traditional lipid panels. Statistical QC steps applied for NMR lipid measures including handling of zero values and missing values, and data transformations were performed prior to analysis (**Chapter 3**).

7.2.4 SNP selection and construction of genetic risk scores for lipid traits

Selection of instrumental variables was based on genetic variants associated with blood lipid and apolipoprotein levels identified in the existing major lipid GWASs. Eligible GWAS of blood lipids identified in the systematic review (**Chapter 2**) included three Global Lipids Genetics Consortium GWAS (GLGC 2013, 2017, 2021), GWAS of traditional lipids and apolipoprotein in the UK Biobank (UKB 2020), and GWAS of traditional lipids in the Million Veteran Program with replication using GLGC data (MVP 2018). All of these GWAS were based on participants of primarily European ancestry (70%-100%). Of these, summary statistics for MVP 2018 were not publicly available and UKB 2020 was limited to participants of European ancestry in UK and not standard reporting of GWAS (results of genetic associations were embedded in the MR study). Thus, summary statistics from GLGC were used to generate instrumental variables associated with three major lipid traits (LDL-C, HDL-C, or TG). Although UKB 2020 was limited to participants mainly of European ancestry, and therefore suboptimal for the selection of genetic instruments for major lipids in other ancestries, it was the largest available GWAS of apolipoproteins and was therefore used to identify instrumental variables associated with ApoB and ApoA1.

GLGC 2013 comprised 60 individual studies including 188,578 participants and identified 248 independent SNPs (r^2<0.001 in the 1000 Genomes European-ancestry Reference Panel) associated with any major lipid trait (TC, HDL-C, LDL-C, or TG) at genome-wide

significance level ($P<5\times10^{-8}$). GLGC 2017 comprised 73 studies including >300,000 participants and identified 406 independent SNPs significantly associated with any major lipid trait (444 SNPs at $P<2.1\times10^{-7}$ and 406 SNPs at $P<5\times10^{-8}$). The most recent and largest GLGC 2021 included about 1.65 million individuals (including 350,000 of non-European ancestries) and identified 1,765 independent SNPs ($P<5\times10^{-8}$, ±500 kb) significantly associated with any major lipid trait (TC, HDL-C, LDL-C, nonHDL-C, or TG). Based on the instrumental variables from GLGC 2013, 2017, and 2021 (significantly associated with LDL-C, HDL-C, or TG at $P<5\times10^{-8}$, within 10 MB and at $r^2<0.001$), three sets of weighted GRS for major lipids (LDL-C, HDL-C, or TG) were generated by multiplying the number of effect allele by the effect sizes as reported in the relevant GLGC GWAS for each SNP (this assumes that SNPs have additive effects on lipid traits).

For each GRS, participants were split into ten equally-sized groups. Linear regression was then used to estimate the mean (SE) levels of blood lipids across the tenths after adjustment for age, sex, study region, and ancestry score. Similarly, associations per 1-SD higher GRS with blood lipid levels at baseline was assessed using linear regression adjusted for the same factors to select the optimum GRS for each blood lipid marker. The percentage of variance of each lipid explained by the GRS was estimated by linear regression model. GRS using SNPs from GLGC 2021 explained more of the variance in blood lipids and was therefore used for the primary instrumental variables in the one-sample MR analyses. Associations of GRS (GLGC 2021) with major lipid traits (TG, LDL-C, HDL-C) and lipoprotein subfractions (i.e. concentration of LDL particles) were also examined in subgroups based on age at recruitment (<45, 45-59, ≥60 years), sex, study region (Coyoacán and Iztapalapa), and ancestry score (tertile). To investigate potential pleiotropic effects of the instrumental variables (effects which may influence ASCVD risk other than through lipid traits), the prevalence and mean values of baseline characteristics including age, sex, blood pressure, socio-economic status and lifestyle factors, physical measures, diabetes status, and medication use, were calculated across

GRS tenths for each lipid trait. To evaluate the degree to which genetic instruments for one lipid or apolipoprotein trait also associated with other lipid related traits, associations per 1-SD higher GRS for each trait with levels of lipids or apolipoprotein traits at baseline was assessed using general linear regression.

A 2020 GWAS of UKB included 439,214 participants of European ancestry and identified 176 and 273 independent (within 10 MB and at r^2<0.001) SNPs associated with ApoB and ApoA1 at P<5×10^{-8}, respectively. Similarly, weighted GRS for these apolipoproteins was generated, their association with apolipoprotein and lipoprotein subfractions in MCPS were examined by subgroups, and the difference in baseline characteristics between GRS tenths for each apolipoprotein was assessed.

Analyses were done using SAS version 9.4 (SAS Institute, Cary, NC) and plots were created in R version 4.2.1 (The R Foundation for Statistical Computing).

7.2.5 MR analyses

Mendelian randomization is a method that uses genetic variants robustly associated with exposure as instrumental variable to infer the potential causal effect of the exposure and outcome, providing key assumptions for validity are met. As variant alleles are randomly allocated at conception and not dependent on population-level variation, MR could be considered analogous to a natural RCT where genetic instruments rather than interventions are randomized (**Figure 7.1**, panel **A**). Notably, the RCT estimates the effect of the intervention over the timeframe of the study, whereas MR estimates the lifetime effects of the genetic variants, leading to substantial differences in the effect estimates. Interpretation of MR results relies on several key assumptions that the genetic variants are associated with the exposure of interest, not associated with confounders, and only associated with the outcome through the exposure (**Figure 7.1**, panel **B**). The detailed methods of estimation and applicable sensitivity analyses are outlined in the following sections.

7.2.5.1 One sample MR analysis

First, a univariable one-sample MR analysis of each lipid-related trait was conducted. A two-stage least-squares (2SLS) method was adopted. The first stage model consisted of a linear regression of each lipid related trait (TG, LDL-C, HDL-C, ApoB, and ApoA1) on the corresponding weighted GRS and the second stage model consisted of a Cox regression of ASCVD mortality on the fitted values from the first stage model, with adjustment for age, sex, region, and the first 7 PCs of population structure in both stages. The first stage can be written as:

$$X = \pi_0 + \pi G + v_x$$

where G is the weighted allele score and π is a single coefficient for the association of the GRS with the exposure. The outcome is then regressed on the predicted value of the exposure, $\hat{X}$:

$$Y = \alpha + \beta\hat{X} + u$$

where Y is the outcome, α is a constant, β is the effect of the exposure on the outcome and u is a random error term assumed to be unrelated to v_x if the assumptions of instrument variables is met (no or balanced pleiotropy are assumed). The estimated value of β ($\hat{\beta}$) is an estimator for the effect of X on Y. 2SLS estimation was conducted in R using the *ivreg* function from the AER package.

The risk of ASCVD mortality across the fifths of weighted GRS for each lipid-related traits were also estimated by Cox regression model using the floating absolute risk method[1] with adjustment for age-at-risk, sex, region, and the first 7 PCs.

7.2.5.2 Multivariable MR analysis

Regardless of the pleiotropy-robust statistical method being used in the sensitivity analyses (***7.2.6***), potential biological or correlated pleiotropy and violation of MR assumptions may still exist in the univariable MR. Multivariable MR (MVMR) analysis[2]

that includes multiple correlated exposures and a set of associated genetic instruments was therefore used to correct for measured pleiotropy and estimate the *direct* effect of each lipid trait on the cardiovascular outcome (ie, the causal effect that is not mediated by any other trait included in the model if the MVMR assumptions are met, whereas univariable MR estimates *total* effect of the exposure). In the multivariable MR model, the direct effect of a lipid related trait on ASCVD can be considered as an 'adjusted' effect after accounting for the mediation or confounding effects of other traits included (**Figure 7.1**, panel **C**). An MVMR model including ApoB, LDL-C, and triglycerides was fitted to identify the atherogenic effect of ApoB-containing lipoproteins and their lipid contents on the risk of ASCVD mortality. Any trait(s) remaining significant in the MVMR model were then further adjusted for HDL-C and ApoA1 to investigate the potential causal effect of HDL-related traits on ASCVD. In the corresponding MVMR model, SNPs associated with all traits included in the model were used as genetic instruments. MVMR allows genetic variants to be associated with all exposures in the model provided that they are not associated with the outcome through other pathways (the third assumption: exclusion restriction). Extensions of the 2SLS or random-effects IVW meta-analysis approaches were used to estimate the causal effect of each trait included in the model. The MVMR model can be written as:

$$X_1 = \pi_{01} + \pi_1 G + v_{x1}$$

$$X_2 = \pi_{02} + \pi_2 G + v_{x2}$$

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$$

where Y is the outcome, α is a constant, and u is a random error term, and β_1 and β_2 are the direct effect of the exposure X_1 and X_2 on the outcome, respectively. G is the genetic instrument (weighted allele score in 2SLS or individual SNP in IVW) and π is a single coefficient for the association of the genetic instruments with the exposure. MVMR analyses using IVW or other pleiotropy-robust methods were conducted in R using the *MendelianRandomization* package and analyses using 2SLS regression were

conducted in R using the *ivreg* function from the AER package.

7.2.6 Sensitivity analyses

The instrumental variable assumptions are required to hold for the results from MR analyses to provide robust causal estimates. Only the first assumption (the robust association between instruments and exposure of interest) can be formally tested and the other two assumptions can be disproved by a range of sensitivity analyses (but cannot be tested to be true). Firstly, to test whether genetic instruments (weighted GRS) are robustly associated with the exposure, the first-stage *F* statistic was calculated (an *F* statistic <10 indicates a weak instrument and therefore potential for weak instrument bias). For the multivariable MR, Sanderson-Windmeijer conditional F-statistics for the first stage model were calculated.[2] Weak instrument bias results from small sample size and/or a small proportion of variance explained by genetic variants and will bias estimates toward the observational estimates in one-sample MR. Secondly, to test the second assumption that genetic instruments should not be associated with confounders of the exposure-outcome association, the prevalence and mean values of baseline characteristics were calculated across GRS tenths for each lipid trait. Any potential confounders found to be associated with the GRS were further adjusted for in the 2SLS regression. Finally, to assess potential horizontal pleiotropy — a violation of the third assumption, an array of methods under different assumptions of the pleiotropy were employed. These included weighted median,[3] weighted mode,[4] and MR-Egger regression[5] were developed for two-sample MR where gene-exposure and gene-outcome estimates derived from separate non-overlapping samples were independent.[6] Recent study have suggested that two-sample MR methods including inverse-variance weighting (IVW) meta-analysis[7] can be safely used for one-sample MR setting (univariable or multivariable) where the two correlated estimates are obtained from the same individuals, although MR-Egger regression is not recommended and will bias the estimates toward the observational estimates when applied to one-sample MR.[8] Thus,

results from 2SLS method were compared with IVW meta-analysis, and weighted median (this approach assumes that at least 50% of the SNPs are valid and the median MR estimate was selected as causal estimate). A consistent effect across multiple methods strengthens causal evidence. To assess robustness of the causal estimates, further tests were conducted, including leave-one-out analysis and Cochran's Q statistic to detect heterogeneous outcomes (genetic instruments exaggerated in magnitude), and the MR Egger intercept test of deviation from the null to detect horizontal pleiotropy. Sensitivity analyses for univariable or multivariable MR using pleiotropy-robust methods were conducted in R using the *TwoSampleMR* package.

7.2.7 Power calculations

Large sample sizes are often required in MR analyses to avoid power issues that may lead to biased results and hinder important discovery.[9] An online power calculator[10] was used to estimate statistical power for 2SLS MR analyses using a genetic instrument (e.g. weighted GRS or individual SNP), a continuous exposure variable (e.g. concentration of LDL-C) and a binary outcome variable (e.g. ASCVD mortality). Using events of ASCVD mortality (~5% cases) and number of participants with NMR lipid measures (~110,000) in the MCPS, power was calculated based on a range of HRs for ASCVD per 1 SD higher lipid-related traits, the proportion of variance explained for the association between weighted GRS and lipid traits (R^2), and a type-I error rate of 0.05 (**Table 7.1**). Previous MR studies based on GWAS of blood lipids from GLGC reported ORs for CHD of 1.53, 0.80, and 1.49 per 1 SD higher LDL-C, HDL-C, and TG, respectively. MR studies based on lipid GWAS from UKB reported similar ORs for CHD of 1.66, 0.80, and 1.34 per 1 SD higher LDL-C, HDL-C, and TG, respectively, and ORs of 1.73 and 0.83 per 1 SD higher ApoB and ApoA1, respectively. If the true relative risk estimates for CVD per SD change in lipid traits lie in the range of [0.80-1) to (1-1.73], the power based on a R^2 of 5% (variance in lipid traits explained by GRS) were estimated to be 88% to 100%, which is large enough to test the null hypothesis that the exposure has no causal effect on the

outcome.

7.3 Results

7.3.1 Associations of genetic risk scores with baseline lipid related traits

Three sets of externally weighted GRS for major lipid traits (LDL-C, HDL-C, and TG) were generated based on summary statistics of GLGC 2013, 2017, and 2021 GWAS. In the GLGC 2013, 78, 83, and 54 independent SNPs were reported to be significantly associated with levels of LDL-C, HDL-C, and triglycerides ($P<5\times10^{-8}$ and $r^2<0.001$), respectively, and 72, 81, and 54 SNPs exist in the MCPS and were used to construct the weighted GRS for three major lipid traits. Similarly, 168, 217, and 170 SNPs were reported in the GLGC 2017 GWAS, and 154, 200, and 166 SNPs exist in the MCPS and were used to construct the weighted GRS for three major lipid traits. 442, 562, and 480 SNPs were reported in the GLGC 2021 GWAS, and 411, 526, and 458 SNPs exist in the MCPS and were used to construct the weighted GRS for three major lipid traits. Three sets of weighted GRS for major lipid traits were robustly associated with linear increases in their corresponding lipid traits after adjustment for age, sex, ancestry score, and region. For each of LDL-C, HDL-C and triglycerides, the GRS constructed using SNPs from GLGC 2021 was demonstrably more strongly related to the lipid than the GRS constructed from either earlier GLGC report (as expected) (**Figure 7.2-7.4**). Per 1-SD increase in GRS from GLGC 2021, lipid traits increased by 0.20±0.01 SD (0.08±0.004 mmol/L) unit, 0.21±0.01 SD (0.04±0.002 mmol/L) unit, and 0.25±0.01 SD (0.09±0.004 mmol/L) unit for GRS-LDL-C, GRS-HDL-C, and GRS-TG, respectively. The linear associations of GRS and major lipids were consistent in subgroups based on sex (**Figure 7.S1**, **S5**, and **S9**), age (**Figure 7.S2**, **S6**, and **S10**), region (**Figure 7.S3**, **S7**, and **S11**), and ancestry score (**Figure 7.S4**, **S8**, and **S12**). The GRS for major lipid traits explained 3.59%, 4.26%, and 5.60% of variation in LDL-C, HDL-C, and TG across the study population, respectively. The effect sizes of genetic variants in the MCPS and GLGC

2021 were highly correlated for LDL-C (r=0.89), HDL-C (r=0.83), and TG (r=0.86) (**Figure 7.7**-**7.9**).

Weighted GRS for major apolipoproteins were generated based on results from the UKB 2020 GWAS. 176 and 273 independent SNPs were reported to be significantly associated with levels of ApoB and ApoA1 ($P<5\times10^{-8}$ and $r^2<0.001$), respectively, of which 167 and 262 SNPs existed in the MCPS and were used to construct the weighted GRS for ApoB and ApoA1 (**Figure 7.5** and **7.6**). Per 1-SD increase in GRS from UKB 2020 was associated with an increase in ApoB of 0.18±0.01 SD (0.03±0.002 g/L) unit and an increase in ApoA1 of 0.17±0.01 SD (0.02±0.001 g/L) unit. The GRS for major apolipoproteins explained 1.88% and 2.65% of variation in ApoB and ApoA1 across the study population, respectively. The effect sizes of genetic variants in the MCPS and UKB 2020 were modestly correlated for ApoB (r=0.66) and ApoA1 (r=0.42) (**Figure 7.10** and **7.11**).

There were no apparent absolute differences in other baseline characteristics across GRS tenths for each lipid trait (**Table 7.S1**-**7.S5**). The mean or percentage difference in characteristics between top to bottom tenths of GRS were generally lower (between 0 and 5%) or the five lipid-related traits.

Notably, many SNPs identified for one lipid-related trait were also associated with other lipid traits (at $P<5\times10^{-8}$) and widespread associations of weighted GRS with each lipid-related trait were identified, indicating their pleiotropic nature and high degree of inter-relatedness among apolipoproteins, lipoproteins and their lipid contents (**Figure 7.12**). For example, in addition to its association with LDL-C, weighted GRS for LDL-C was positively associated with ApoB (0.22 SD unit per SD change in GRS) and TG (0.10 SD unit per SD change in GRS). Similarly, there were widespread associations of GRS for major lipid-related traits with lipoprotein subfractions and their lipid contents (**Figure 7.13-7.17**), highlighting the intercorrelation of lipid metabolism and the physiological

relatedness of the composition of lipid particles in the circulation. For example, GRS for LDL-C was positively associated with the levels of ApoB, VLDL and LDL particles and their lipid contents (free or esterified cholesterol, triglycerides, phospholipids), and was inversely associated with the levels of HDL particles and their contents. These widespread associations indicate an inherent difficulty in disentangling the exclusive effects of individual lipid components on the risk of cardiovascular diseases and justify the use of multivariable MR to produce more reliable estimates after account for inevitable pleiotropy.

7.3.2 One-sample univariable and multivariable MR between lipid traits and ASCVD

When assessed individually using one-sample univariable MR, we identified LDL-C and ApoB were significantly associated with a higher risk of ASCVD (**Figure 7.19**a). Two SD-higher LDL-C (1 mmol/L) was associated with a 31% (HR 1.31, 1.03-1.67) increased risk of ASCVD and two SD-higher ApoB (0.42 mg/dL) was associated with a 74% (HR 1.74, 1.25-2.43) increased risk. The association with LDL-C was observed in men (HR 1.56, 1.03-2.35) but not in women (1.14, 0.83-1.56). The associations with ApoB were consistently observed in both men (1.92, 1.01-3.66) and women (1.70, 1.13-2.54). Additional subgroup analyses stratified by age groups, regions, ancestry groups, and diabetes shown similar patterns of association with LDL-C and ApoB, although the statistical power was low for several subgroups with small sample sizes (**Figures 7.S21**a-**S23**a). Sensitivity analyses using 2-sample methodological approaches that accounted for potential genetic pleiotropy led to broadly similar results (**Figure 7.20**).

When LDL-C, TG and ApoB were assessed together using one-sample multivariable MR (**Figure 7.19**b), only ApoB retained a robust, potentially causal association with ASCVD (2-SD change HR, 1.71, 1.13-2.59). The association with LDL-C decreased to null after mutual adjustment, with an adjusted HR of 1.06 (0.78-1.27). Similarly, when adjusted for ApoA1 and HDL-C, there was robust evidence for potential causal role of ApoB in the

development of ASCVD (HR, 1.86, 1.32-2.61). The associations with ApoB were broadly consistent across subgroups by sex, age, region, ancestry groups, and diabetes (**Figures 7.S21**b-**S23**b). After adjustment of ApoB and ApoA1, a significant association between HDL-C and ASCVD risk was observed (1.35, 1.01-1.79).

7.3.3 Sensitivity analyses using pleiotropy-robust two-sample MR approaches

Multiple sensitivity analyses using an array of methods under different assumptions of the pleiotropy (**Figure 7.20**) shown consistent results that ApoB was significantly and causally associated with the risk of ASCVD. The MR Egger intercept test did not detect horizontal pleiotropy for ApoB and LDL-C, but detect horizontal pleiotropy for HDL-C, TG, and ApoA1 (**Figure 7.20**).

7.4 Discussion

7.4.1 Main findings in context

In this MR study based on a large understudied Latin American population and externally weighted GRSs, the potential causal associations of major lipids and apolipoproteins with the risk of ASCVD were assessed using both univariable and multivariable analyses. The GRSs for lipids constructed based on the currently largest lipid GWAS meta-analysis of multi-ancestry participants (GLGC 2021[11]) explained more trait variance than those for apolipoproteins based on the GWAS of mainly European participants (UKB 2020[12]), albeit both set of GRSs were validated across population subgroups and closely related to the baseline levels of lipid traits. Given the intrinsic high correlation and relatedness amongst lipoproteins and their lipids and apolipoproteins, expected high pleiotropy were observed where a GRS for one lipid trait was also related to other lipid traits and subfractions, justifying the use of multivariable methods and pleiotropy-robust approaches. By comparing the results from univariable and multivariable results, the key findings are that after accounting for the effects of ApoB, the significant association with LDL-C when assessed individually attenuated markedly toward the null. By contrast, the

association with ApoB with the risk of ASCVD was robust and remained significant after adjusting for other traits (e.g., LDL-C and TG or HDL-C and ApoA1). The MR risk estimates for ApoB were similar to those from observational analyses of single traits (similar to univariable MR) and mutual adjustment (e.g., ApoB *given* LDL-C, similar to multivariable MR). Collectively, these observational and genetic findings provide powerful and consistent evidence that ApoB, as the most accurate and direct proxy of ApoB-containing particles, is ***causally*** associated with the risk of ASCVD. This adds further evidence from a understudied Latin American population to support the emerging consensus that it is the number of atherogenic apoB-containing particles per se rather than their cholesterol or triglycerides content that is a central determinant for the development of ASCVD. The clinical benefit of modifying blood lipids (LDL-C, HDL-C, or TG) should be proportional to the absolute change in ApoB.

7.4.2 Causal relevance of ApoB and LDL-C on ASCVD

In line with the observational findings from the MCPS (Chapter 6) and previous evidence from epidemiological, clinical, and genetic studies, the current MR study supports the view that ApoB, as the proxy of apoB particles, is causally related to the risk of ASCVD, independent of lipid contents such as TG and cholesterol and HDL or ApoA1. As this notion has been discussed in the broad context of traditional epidemiological studies and trials of lipid-modifying medications in **Chapter 6**, the present discussion will mainly focus on comparing current findings with those from MR studies. Individual assessment of blood lipids using univariable MR suggested the adverse effect of both LDL-C and ApoB on ASCVD, which was largely consistent with previous MR studies based on major lipid traits and in line with the common view that atherogenic lipid traits might play a causal role in ASCVD.[12-15] When the direct, independent effect of these lipid traits were assessed using multivariable MR, only ApoB remained related to the increased risk of ASCVD with the effect of LDL-C being diminished, which was consistent with a previous MR study. Taken together, these findings suggest that among the major lipid traits,

including ApoB, ApoA1, LDL-C, HDL-C and TG, ApoB predominates as the underlying cause of ASCVD. Building on the observational findings in MCPS and other cohorts, these MR findings also support the emerging consensus that ApoB unifies and simplifies the information from the conventional lipid panels as to the atherogenic risk attributable to the apoB-containing particles.[16,17] Our findings do not discredit the causal roles of LDL-C/LDL in ASCVD, as ApoB does not occur in physical isolation and is closely related to LDL particles, but suggests that apoB particles per se are the basic unit of injury to the arterial well that allows the atherogenic effects of LDL-C to be expressed.[12,16]

The current findings are further supported by the MR studies using genetic variations in the *CETP* and *HMGCR* genes to mimic the effects of CETP inhibitors and statins. Combination therapy of CETP inhibitor and statin caused large decrease in LDL-C but did not result in significant clinical benefit in ACCELERATE trial.[18] The MR studies found that the combination use was associated with discordant reductions in LDL-C and apoB concentration (with same reduction in LDL-C but attenuated reduction in apoB level), and the risk of CVD was proportional to the *attenuated* reduction in apoB level but significantly less than expected per unit change in LDL-C.[19] These findings collectively indicate that the risk of ASCVD is more closely and causally related to the number of circulating apoB particles than with the variable mass of cholesterol or TG within them.

7.4.3 Relevance of HDL-C and ApoA1 on ASCVD

Although HDL-C and ApoA1 were inversely associated with the risk of ASCVD in observational analyses with/without adjusting for the effect of ApoB, neither trait was significantly related to risk in univariable MR analyses. After adjusting for ApoA1 and ApoB, HDL-C was positively associated with the risk in multivariable analyses. Previous MR studies also challenged the protective role of HDL-C observed in observational studies, showing that genetically reduced HDL-C is not associated with an increased risk of ASCVD.[19-25] Patients with genetic variants in genes including *CETP*, *ABCA1*, *LIPC*,

and *SCARB1* were paradoxically at increased risk of ASCVD despite with very high levels of HDL-C.[26,27] Similarly, in observational analysis (chapter 6), large HDL-P with more cholesterol contents and HDL particle size were positively associated with the risk. Studies exploring larger range of HDL-C concentration have indicated a U-shaped relationship between HDL-C, ASCVD and all-cause mortality, with an increased CVD risk at high level of HDL-C.[28-31] Taken together, the positive association with HDL-C after accounting for ApoA1 and ApoB suggests that HDL particles with very high level of cholesterol (e.g., bigger and less dense HDL-P) may be hazardous and warrants further research.

Evidence from substantial clinical trials of niacin, fibrates, and CETP inhibitors also did not support the protective effects of higher HDL-C. Despite significantly increasing HDL-C level, most of these agents did not reduce all-cause mortality or ASCVD.[32] The current guideline from European Society of Cardiology/European Atherosclerosis Society emphasizes that the ASCVD risk appeared to increase when HDL-C is above 90mg/dL.[33] Taken together, current data suggest that higher HDL-C especially those in larger HDL particles is not necessarily protective against cardiovascular disease and may even be harmful in extremely high levels. The underlying mechanisms by which very high levels of HDL-C could have adverse effects, while lower levels are associated with increased risk, remain unclear. As discussed in chapter 6, HDL-C levels may not be a reliable marker for the complex biological function of HDL particles. Extremely elevated HDL-C levels may represent the dysfunction in HDL metabolism, which may then promote rather than protect the development of ASCVD. An alternative hypothesis is that genetic mutations that result in elevated HDL levels may also contribute to an increased risk of vascular disease through unknown mechanisms.

However, current risk assessment tools of ASCVD commonly do not take into account the emerging evidence that very high levels of HDL-C may not be protective but hazardous. The estimated risk of ASCVD death decreased with increasing HDL-C,

without any threshold level. Overall, the causal effects of HDL-C on ASCVD remain controversial and warrant further investigation. The current observational and MR findings from the MCPS, combined with previous emerging evidence, suggest that HDL particles of varying sizes and biological functions may differ in their effects on ASCVD, and that high levels of cholesterol in larger HDL may be harmful.

7.4.4 Relevance of TG on ASCVD

In observational analyses, TG was positively associated with the risk when assessed individually, whereas the association was largely attenuated after adjusting for the effect of ApoB and LDL-C. In both univariable and multivariable MR analyses, TG was not significantly related to the ASCVD risk. The current results are partly consistent with previous observational and genetic studies suggesting the association of TG within TGRLs with ASCVD often become nonsignificant after adjustment for HDL-C and nonHDL-C/ApoB. As illustrated in previous sections of ApoB-containing particles, the risk reduction of ASCVD related the TG (or cholesterol) within the ApoB-containing particles should be proportional to the relevant reduction in ApoB per se.

As mentioned before, the addition of TG-lowering therapy such as fibrates and niacin to statin therapy did not result in significant overall clinical benefit, although the TG levels were greatly reduced (by 30% to 50%). However, benefits were observed in patient subgroups with hypertriglyceridemia (baseline TG > 200 mg/dL), which explained why these trials failed to consistently suggest a reduced risk of ASCVD. The average decrease in TG levels in these trials was approximately 20 mg/dL to 50 mg/dL for most participants, along with a modest reduction in apoB-containing particles, which was insufficient to achieve clinical significance. Further MR studies using genetic variants in genes involved in TGRLs (*LPL*) and LDL (*LRLR*) metabolism demonstrated that the potential clinical benefit of lowering TG levels is similar to the benefit of lowering LDL-C per unit change in apoB. To reduce the risk of CVD risk by 20% (as can be achieved by

lowering LDL-C by 40 mg/dL)[34] TG levels would need to be reduced by 5-fold this quantity (the concentration of TGRLs can be estimated by dividing TG concentration by 5 on the mg/dL scale), which is about 200 mg/dL, to achieve the same reduction in apoB-containing particles.[35] The absolute reduction in TG levels in the above mentioned TG-lowering trials was only a fraction of that needed to significantly reduce the risk of ASCVD (20 mg/dL to 50 mg/dL), which may explain the observation that the clinical benefit was only shown in subgroups of patients with high baseline TG levels. Taken together, although TG predicts vascular risk, the current findings, combined with previous evidence, do not support TG as a causal factor in the development of ASCVD.

7.4.5 Strength and limitations

The current analysis is one of the largest MR studies conducted in a Latin American population. The most recent data from large lipid multi-ancestry GWAS was used to construct the comprehensive genetic instruments for composite lipid-related traits. Although the variants significantly associated with lipid traits were identified in external populations of mainly European ancestry, they have been validated in MCPS as robust instruments closely related to baseline lipid levels at baseline and not associated other baseline characteristics. Based on the constructed and validated genetic instruments, expected high correlations between GRS for specific traits and other traits were observed, making it challenging to assess the independent, causal effect of each trait. Multivariable MR analyses that could account for the measured horizontal pleiotropy were used and ApoB was identified to have a fundamental causal role in the risk of ASCVD, independent of other lipid or apolipoproteins. Fourth, residual pleiotropy (beyond 'measured' pleiotropy) may still influence the interpretation of multivariable MR results, which should be approached with caution. Finally, a series of two-sample MR approaches more robust to potential unbalanced pleiotropy (by making different assumptions) were used in sensitivity analyses.

There are several limitations, however. First, although the NMR lipid traits included information on lipid subfractions, composite major lipid traits were used as the main exposures. Current available GWAS on NMR lipid subfractions are of limited sample sizes compared to those from large lipid consortium on composite lipids and, consequently, have discovered only a limited number of significant variants for each subfractions (thereby explaining only a small proportion of variance and not lending themselves to the creation of strong instruments). Previous studies have used a two-sample MR-BMA approach by including a set of potentially highly correlated lipid risk factors in a model for variable selection, prioritizing ApoB as the most relevant risk factor and XS-VLDL-P and L-LDL-P from the lipoprotein subfractions. Future studies are needed to investigate the potentially heterogeneous effects of LDL/HDL lipoproteins of different sizes and functions on ASCVD, when larger GWAS on lipid subfractions exist and more developed methods are available, such as the recent published one[36] (GWAS on 233 NMR-measured circulating metabolic traits in 136,016 participants from 33 cohorts). Second, two-sample MR analysis of the same Latin American ancestry was not conducted in this study. The primary reason for this is the limited availability of GWAS data on ASCVD, which have been predominantly conducted in European populations. Ideally, the two study samples should be independent and from same population. Using samples of unrelated individuals from different ancestries can bias the risk estimates due to population stratification and differences in allele frequencies and selection and participation bias.[37] The current findings from one-sample MR should be validated using a two-sample design when more GWAS of ASCVD in Latin American populations are available. Third, although the current findings cast further doubt on the relevance of HDL-C in the development of ASCVD, it is likely that specific HDL subtypes could play an important role as observed for small HDL particles in observational analyses, which facilitate further bespoke MR studies on lipid subfractions. Finally, the MR analyses were focused on the lipid biomarkers rather than specific drug targets. Future MR using

genetic variants in genes that mimic the effects of lipid-modifying drugs of interest in Latin-American population should provide more information on the relevance between blood lipids and ASCVD.

7.5 Conclusion

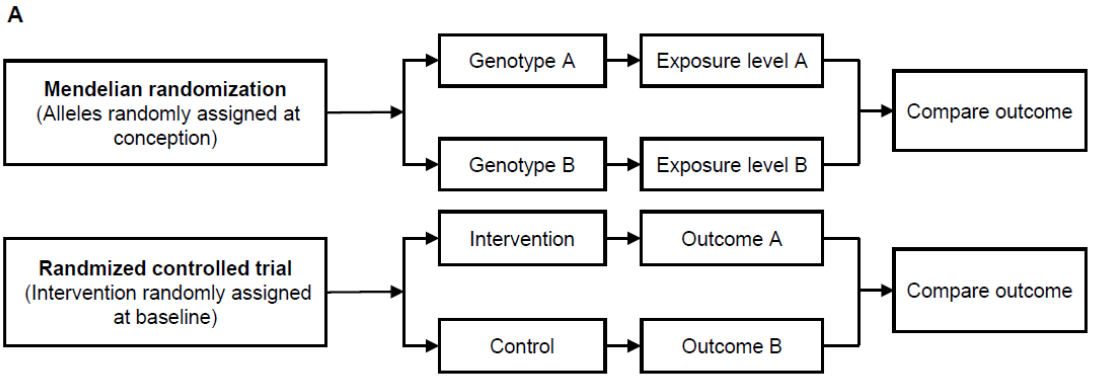
Complementary MR and observational analyses consistently support the view that ApoB (reflecting the number of circulating ApoB-containing particles) is causally associated with the risk of ASCVD. HDL-C, especially that in larger HDL particles, is not necessarily protective against cardiovascular disease and may even be harmful in extremely high levels, which warrants further investigation. Although TG predicted ASCVD risk, the current findings do not support TG as a causal factor in the development of ASCVD. Given the totality of the evidence, ApoB should be the primary measure of atherogenic risk reflecting the apoB-containing lipoproteins and reducing the number of ApoB-containing particles should remain the focus of efforts to reduce ASCVD risk.

Table 7.1 Power calculation for two-stage least squares mendelian randomization of continuous exposure and binary outcome

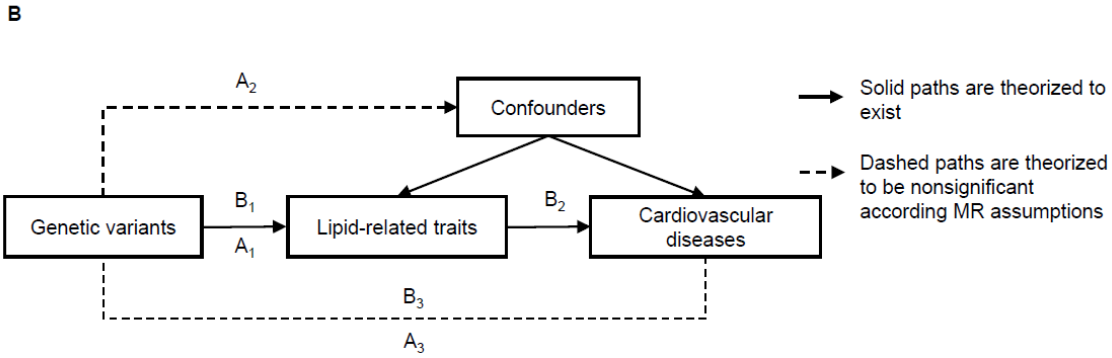
	Statistical power
1% variance explained	
HR 1.01	0.05
HR 1.10	0.13
HR 1.50	0.98
HR 2.00	1.00
5% variance explained	
HR 1.01	0.05
HR 1.10	0.44
HR 1.50	1.00
HR 2.00	1.00

Estimated by an online power calculation for MR study. Results shown are for a binary outcome (ASCVD mortality) based on 5% cases and 140,000 individuals in the MCPS, and a type-I error rate of 0.05.

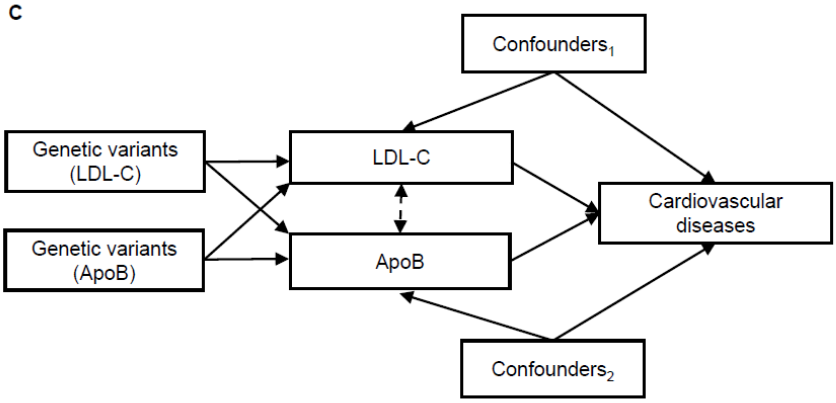
Figure 7.1 Illustration and assumptions of Mendelian randomization



The parallels between MR and RCT. Inference from MR also require that changes genetic variation are equivalent in their effects to changes in the exposure through environmental factors or pharmaceutical manipulation (gene-environment equivalence).



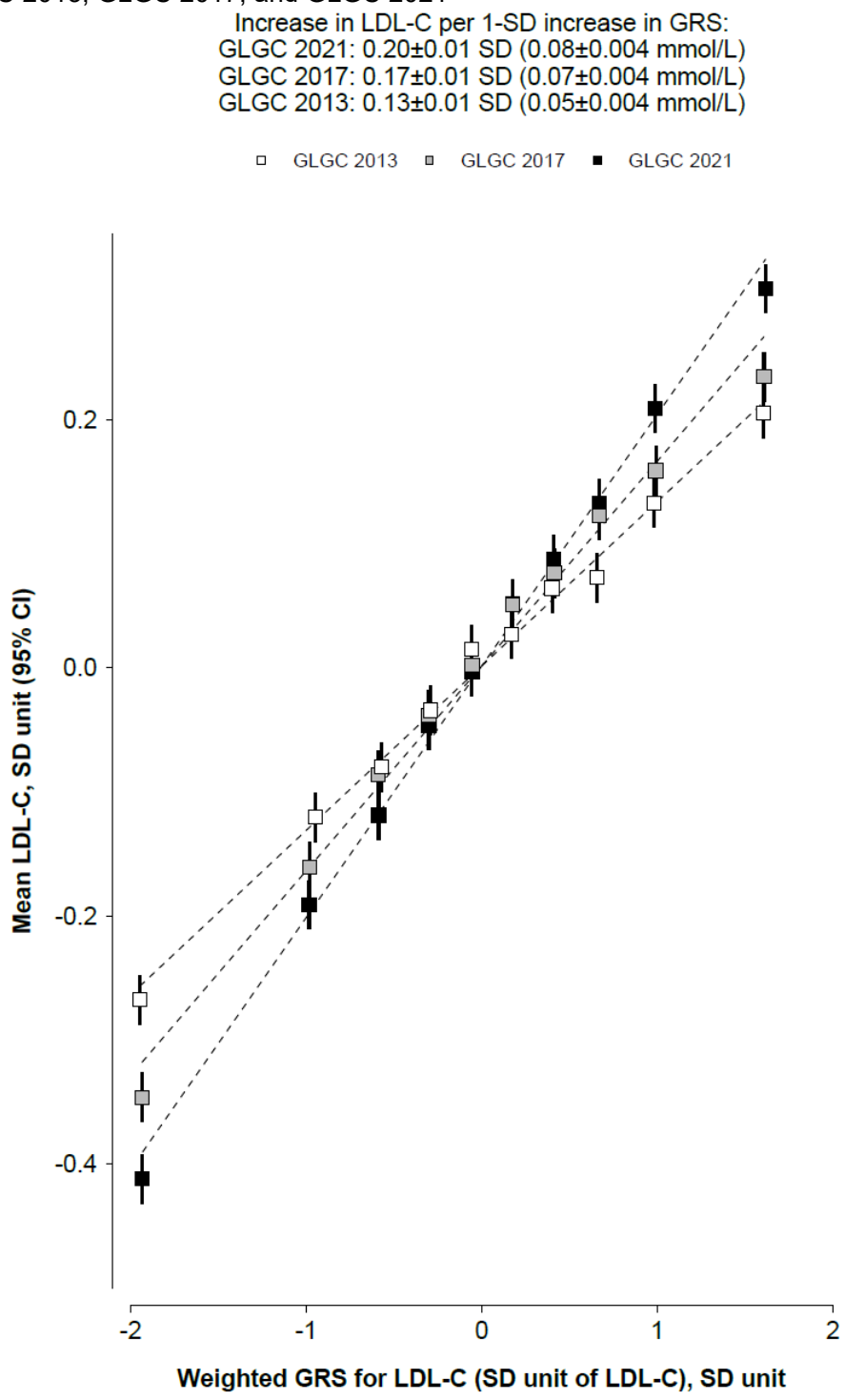
B_2 indicates the potential causal relationship of interest, where B_2=B_3/B_1 and B_3 represent estimated direct effect of genetic variants on the exposure (eg, LDL-C) and outcome (eg, ASCVD), respectively. In MR, genetic variants are valid instrumental variables if the following assumptions are met: A_1 Relevance: The variants are associated with the exposure of interest; A_2 Exchangeability: The variants are independent of confounders of the associations of interest; A_3 Exclusion restriction: The variants are not associated with the outcome through alternative pathways (horizontal pleiotropy).



In conventional MR, genetic variants used as instruments of LDL-C may also associated with ApoB, leading to violation of A_3 and potential confounding of the causal estimate of LDL-C and risk of CVD (total effect of LDL-C). In multivariable MR, genetic variants associated with each exposure of interest are used together with their corresponding genetic associations with each exposure in the analysis, permitting an estimation of the direct effect of each exposure on the outcome (the direct effect of a lipid trait on CVD accounting for either mediation or confounding by other exposures included in the model).

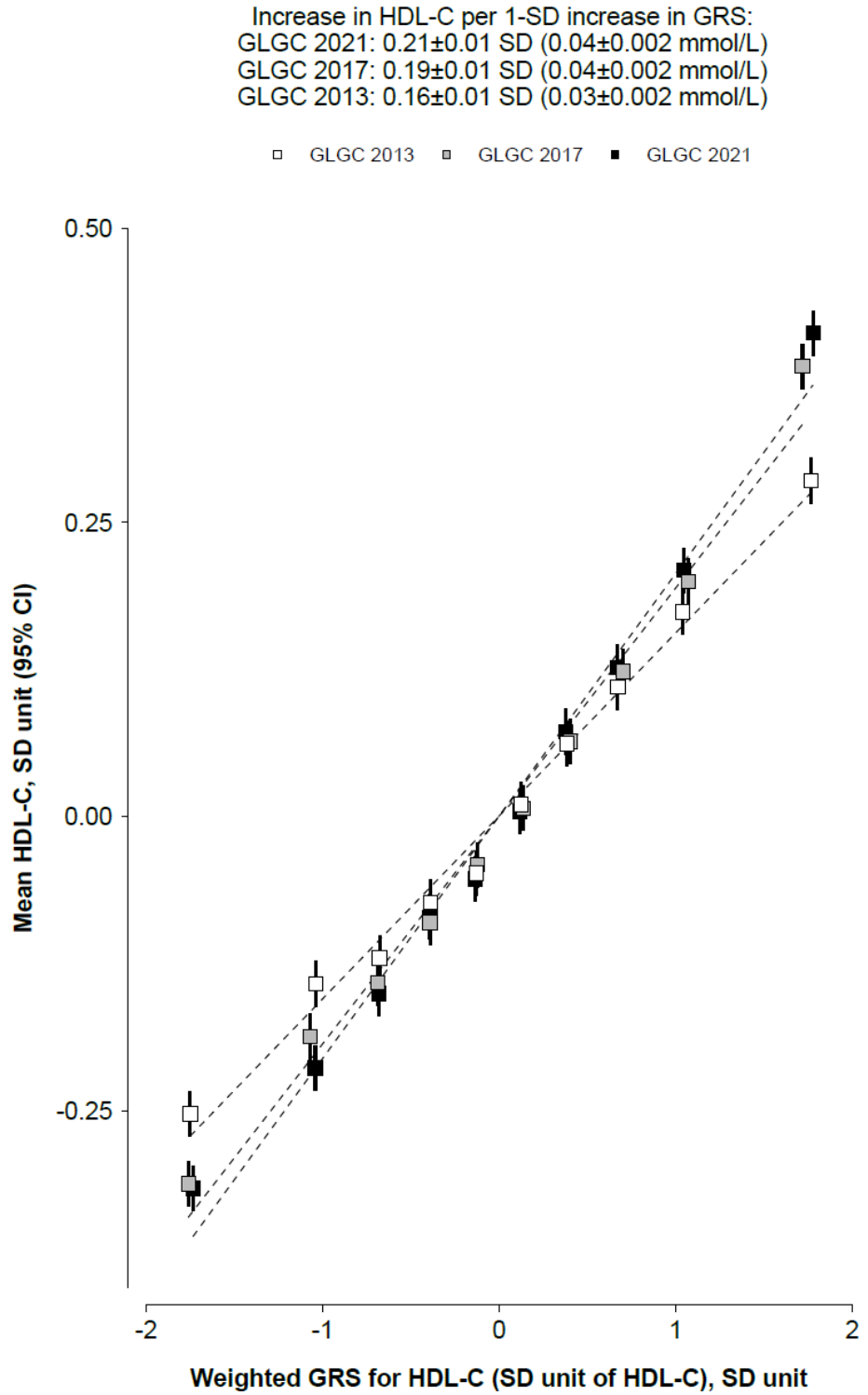
(A) Mendelian randomization as a natural RCT. (B) Principles and assumptions of conventional Mendelian randomization. (C) Multivariable Mendelian randomization.

Figure 7.2 Mean baseline NMR-measured LDL-C by GRS for LDL-C using SNPs from GLGC 2013, GLGC 2017, and GLGC 2021



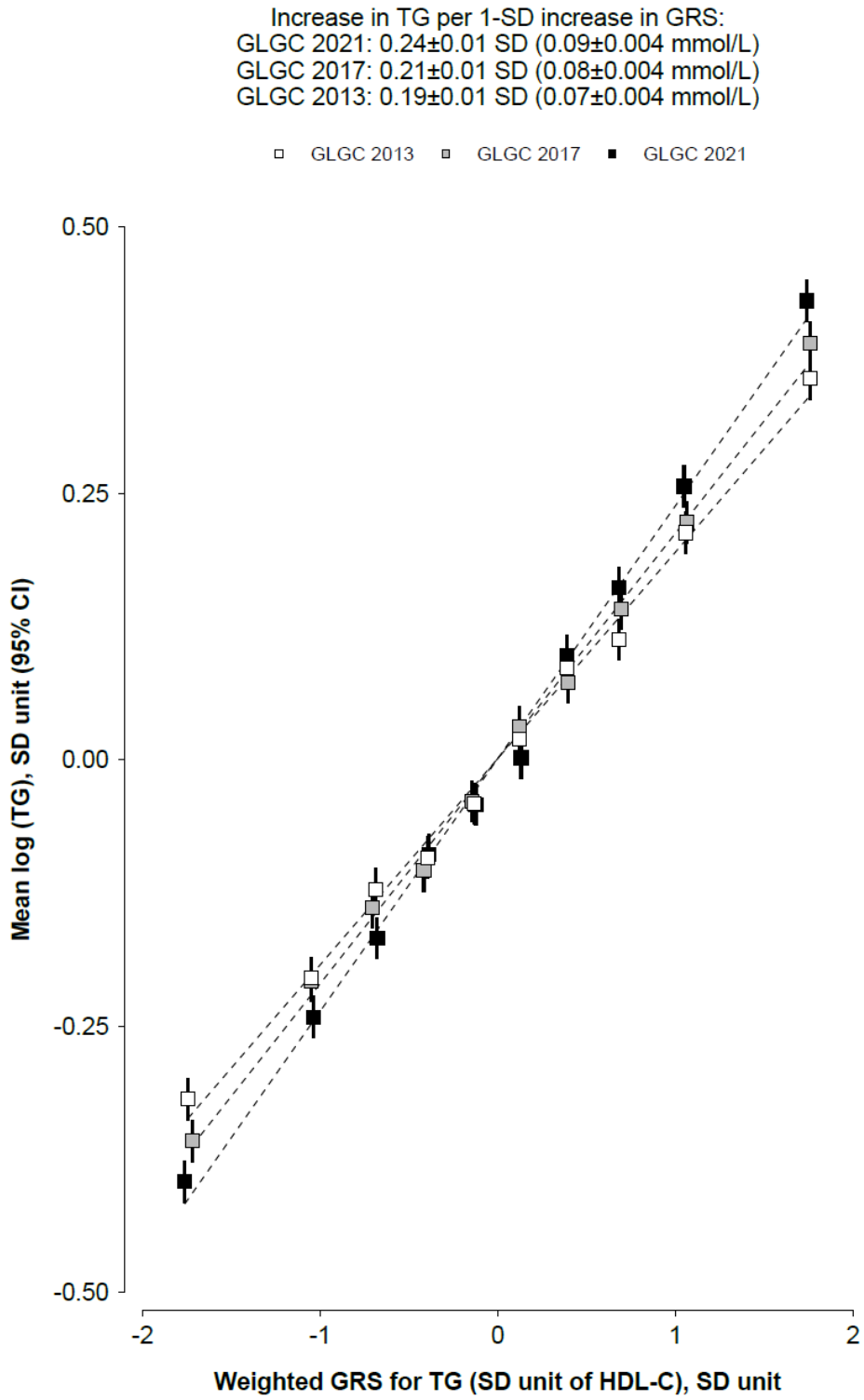
Adjusted for age, sex, study region, and ancestry score.
Squares represent mean LDL-C level with area inversely proportional to the variance.
Vertical lines represent the corresponding 95% confidence intervals.
The GRS consisted of 72 independent SNPs significantly associated with LDL-C levels (P<5e-8 and r2<0.001) from GLGC 2013 (78 SNPs reported), 154 SNPs from GLGC 2017 (168 SNPs reported), and 411 SNPs from GLGC 2021 (442 SNPs reported).

Figure 7.3 Mean baseline NMR-measured HDL-C by GRS for HDL-C using SNPs from GLGC 2013, GLGC 2017, and GLGC 2021



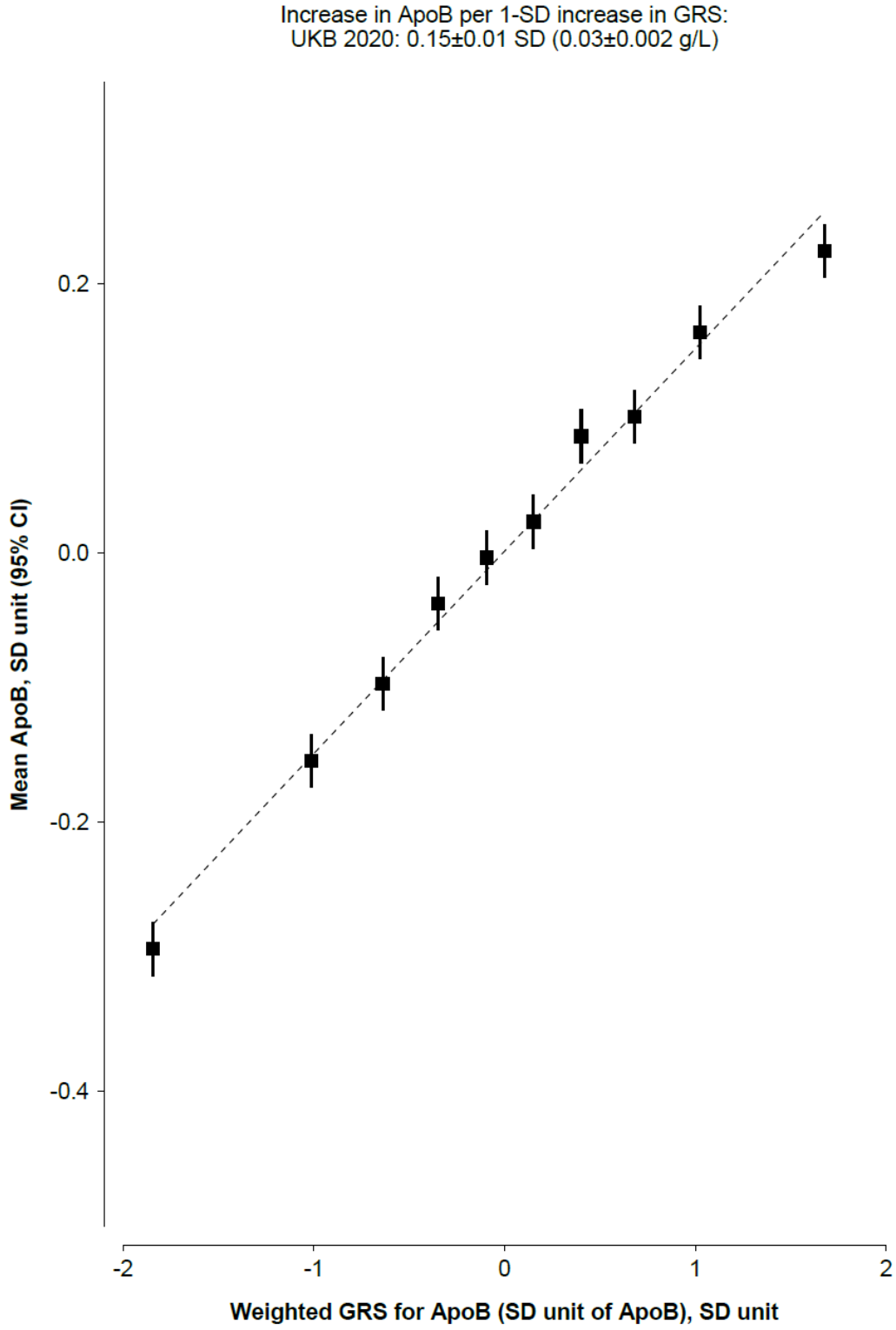
Adjusted for age, sex, study region, and ancestry score.
Squares represent mean HDL-C level with area inversely proportional to the variance.
Vertical lines represent the corresponding 95% confidence intervals.
The GRS consisted of 81 independent SNPs significantly associated with HDL-C levels (P<5e-8 and r2<0.001) from GLGC 2013 (83 SNPs reported), 200 SNPs from GLGC 2017 (217 SNPs reported), and 526 SNPs from GLGC 2021 (562 SNPs reported).

Figure 7.4 Mean baseline NMR-measured TG by GRS for TG using SNPs from GLGC 2013, GLGC 2017, and GLGC 2021



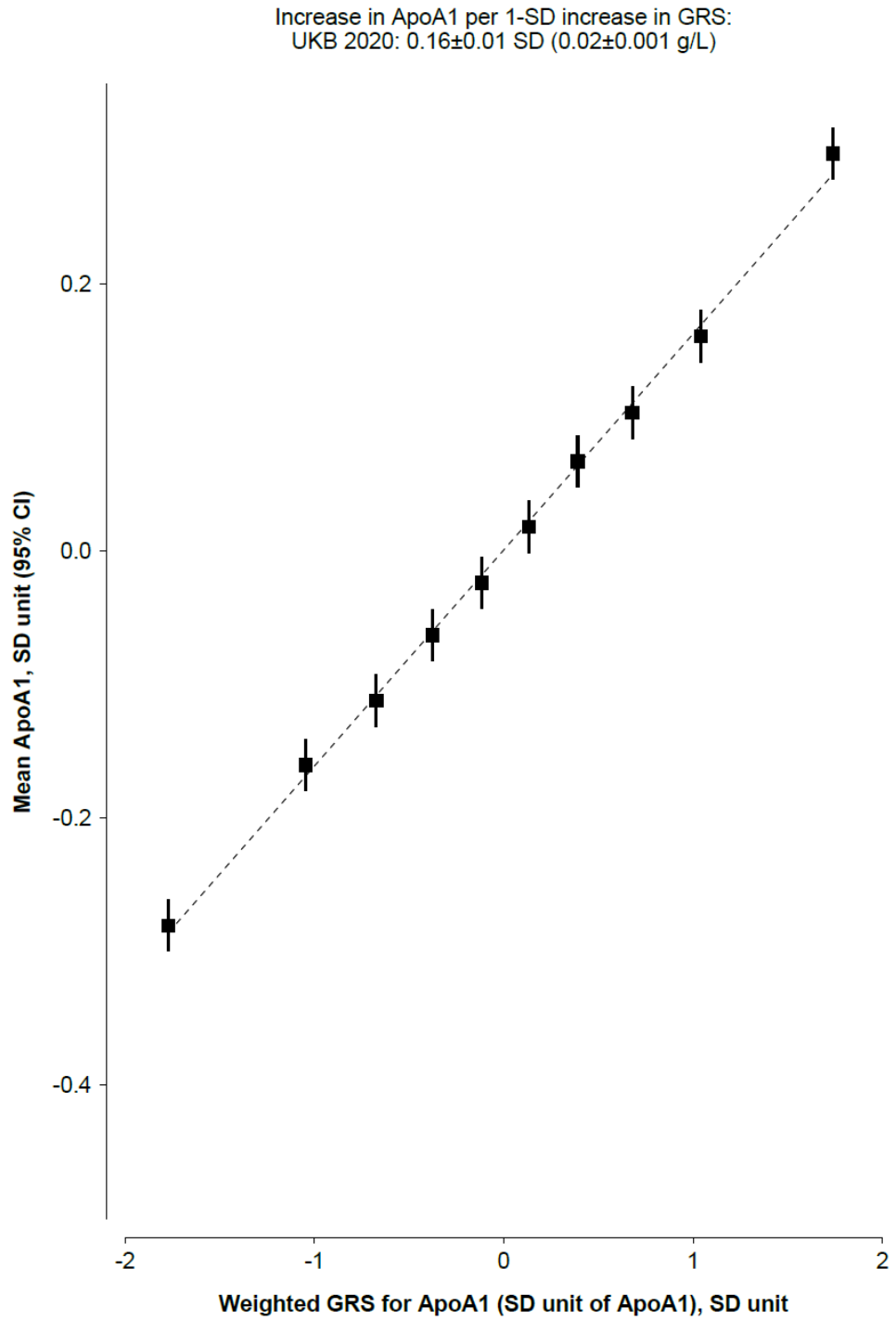
TG levels were natural-log transformed.
Adjusted for age, sex, study region, and ancestry score.
Squares represent mean log (TG) level with area inversely proportional to the variance.
Vertical lines represent the corresponding 95% confidence intervals.
The GRS consisted of 54 independent SNPs significantly associated with TG levels (P<5e-8 and r2<0.001) from GLGC 2013 (54 SNPs reported), 166 SNPs from GLGC 2017 (170 SNPs reported), and 458 SNPs from GLGC 2021 (480 SNPs reported).

Figure 7.5 Mean baseline NMR-measured ApoB by GRS for ApoB using SNPs from UKB 2020



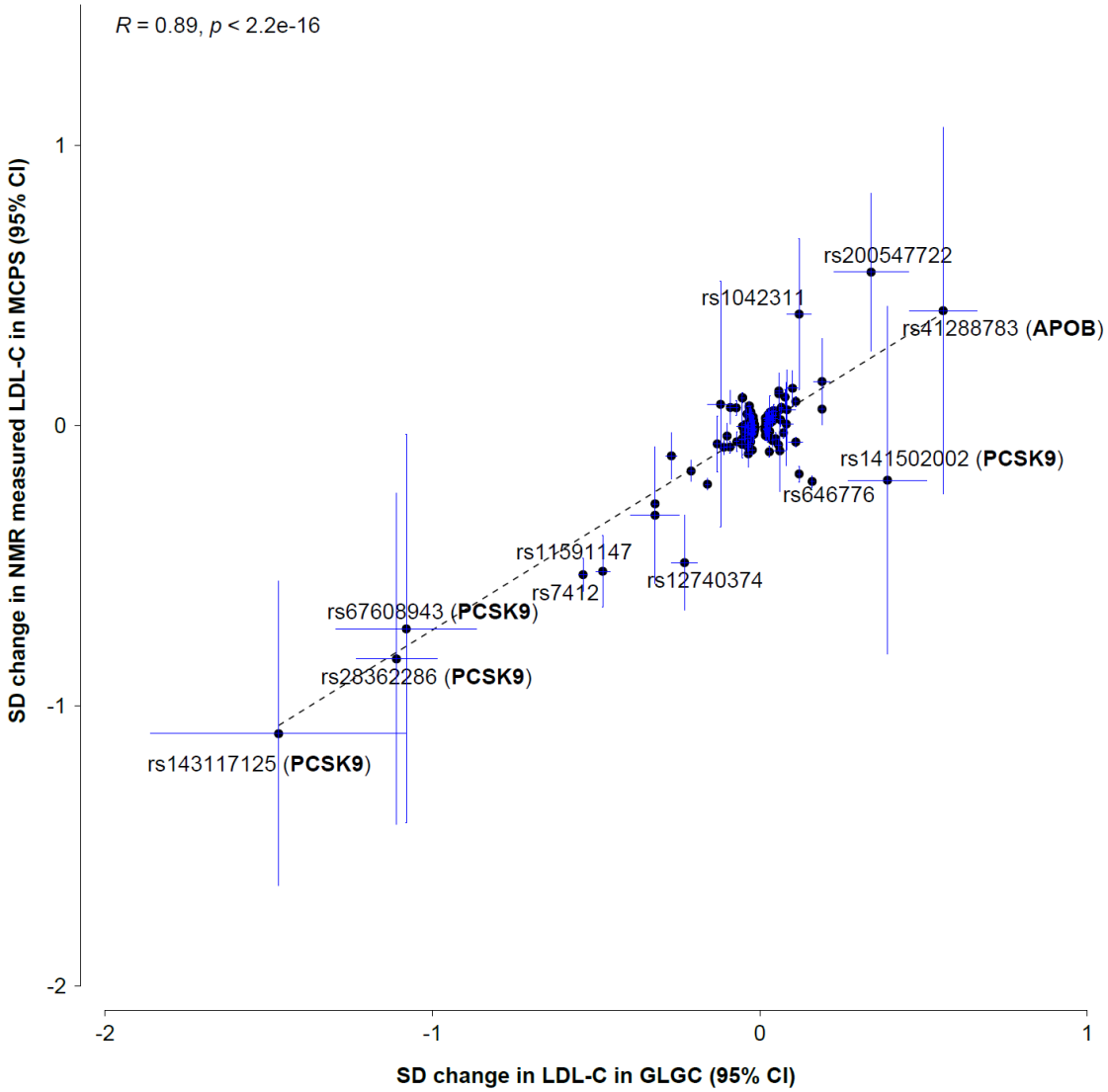
Adjusted for age, sex, and study region.
Squares represent mean ApoB level with area inversely proportional to the variance.
Vertical lines represent the corresponding 95% confidence intervals.
The GRS for ApoB consisted of 167 independent SNPs significantly associated with ApoB levels (P<5e-8 and r2<0.001) from UKB 2020 GWAS (176 SNPs reported).

Figure 7.6 Mean baseline NMR-measured ApoB by GRS for ApoB using SNPs from UKB 2020



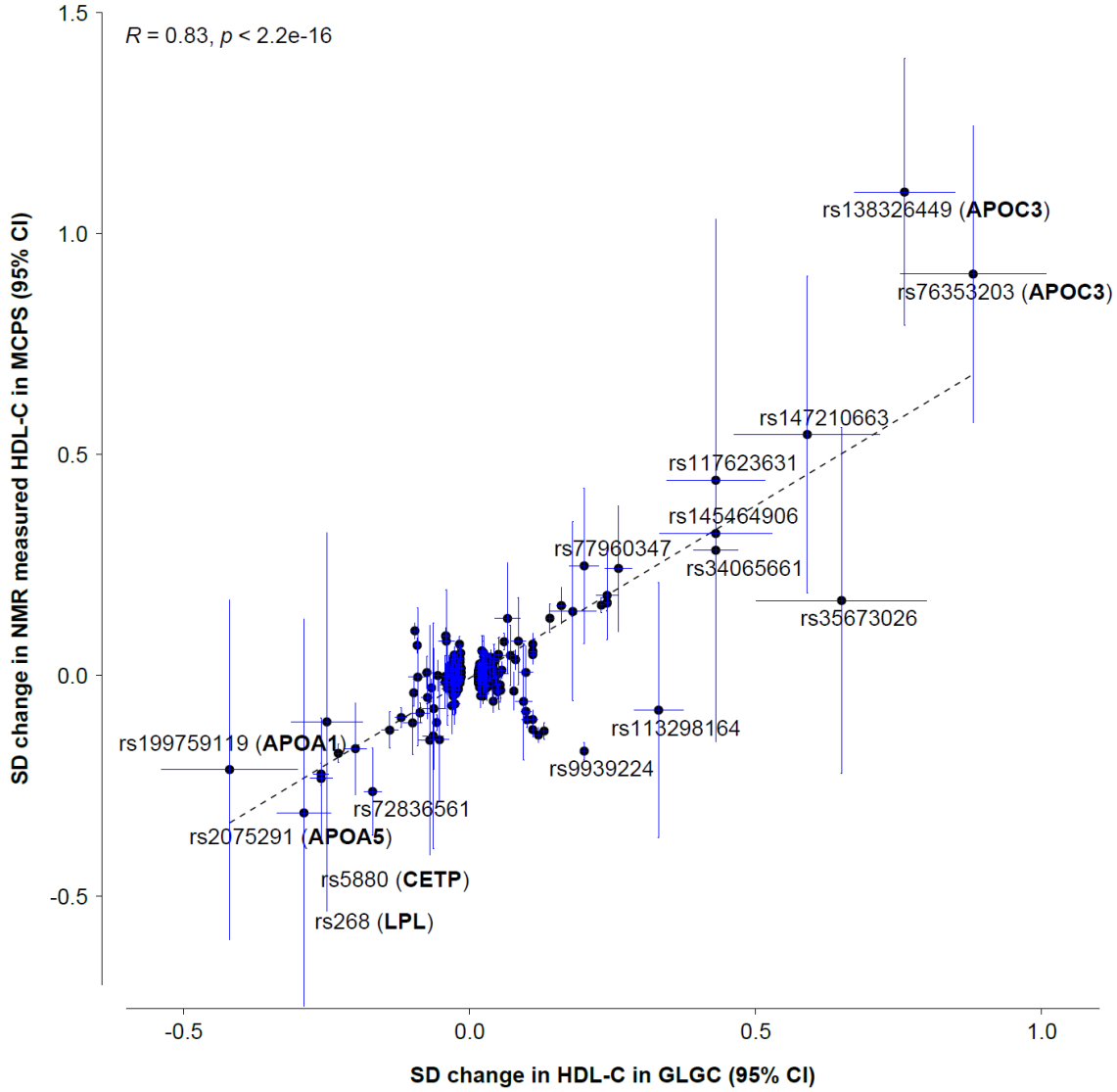
Adjusted for age, sex, ancestry score, and study region.
Squares represent mean ApoA1 level with area inversely proportional to the variance.
Vertical lines represent the corresponding 95% confidence intervals.
The GRS for ApoA1 consisted of 262 independent SNPs significantly associated with ApoA1 levels (P<5e-8 and r2<0.001) from UKB 2020 GWAS (273 SNPs reported).

Figure 7.7 Comparison of effect sizes of genetic variants associated with LDL-C in MCPS and GLGC GWAS



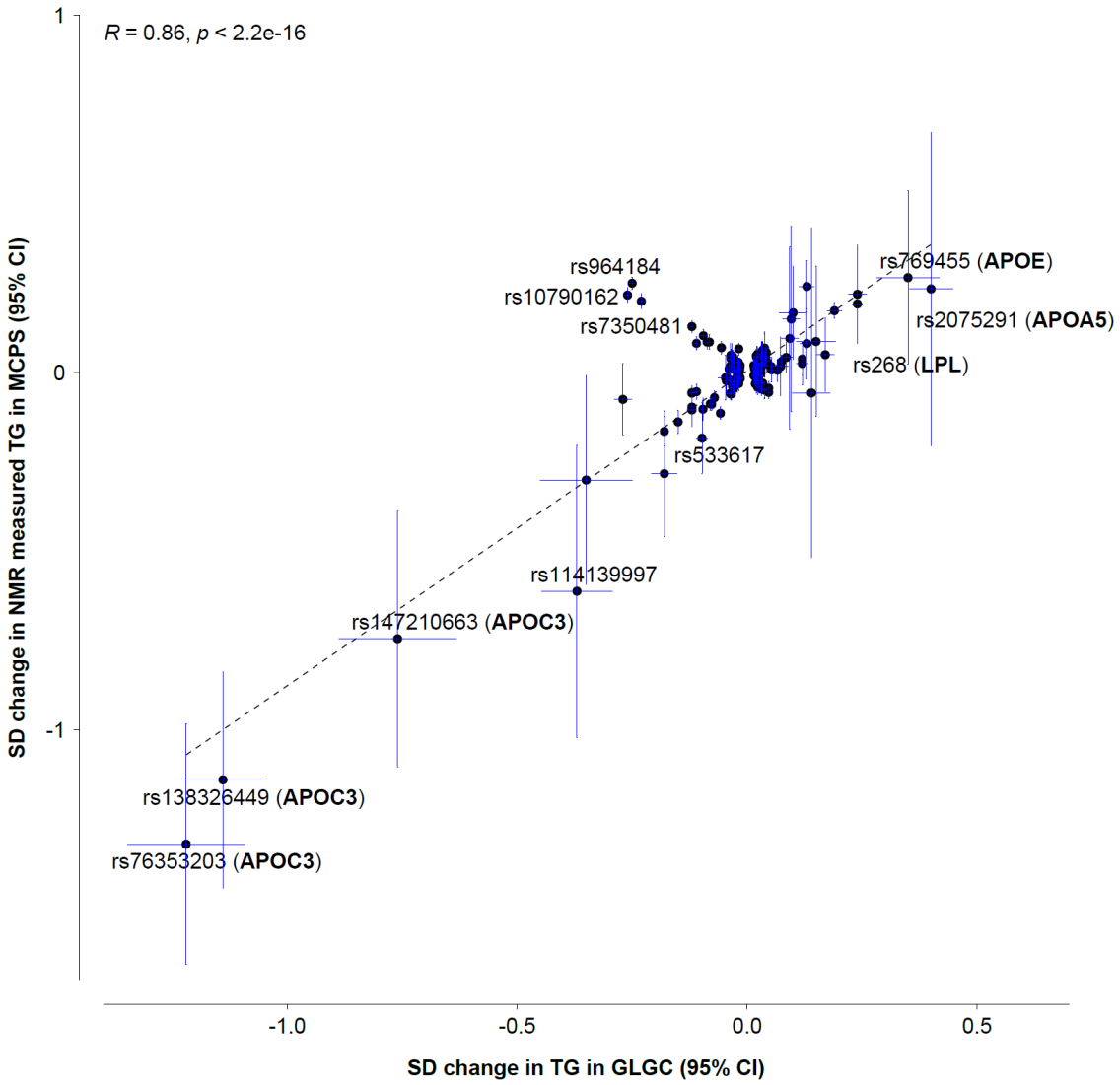
A total of 154 SNPs associated with LDL-C levels were included in the analyses as identified by GLGC GWAS (Liu et al, 2017). Effect sizes represent the standard deviation (SD) increase in LDL-C per additional allele in the GLGC and MCPS. The GLGC analyses (73 studies including >300,000 individuals) were performed by study separately and overall estimates were combined by inverse-variance weighted meta-analysis. The MCPS analyses were conducted in 29,327 individuals with no prior chronic disease and NMR lipid measurements at baseline, and with QC-passed genetic information. Adjustment were made for age, sex, principal components (GLGC: up to four; MCPS: 7) in the GLGC and MCPS. Vertical lines represent the 95% confidence intervals (CIs) of the effect estimates in the MCPS. Vertical lines represent the 95% CIs of the effect estimates in the GLGC study.

Figure 7.8 Comparison of effect sizes of genetic variants associated with HDL-C in MCPS and GLGC GWAS



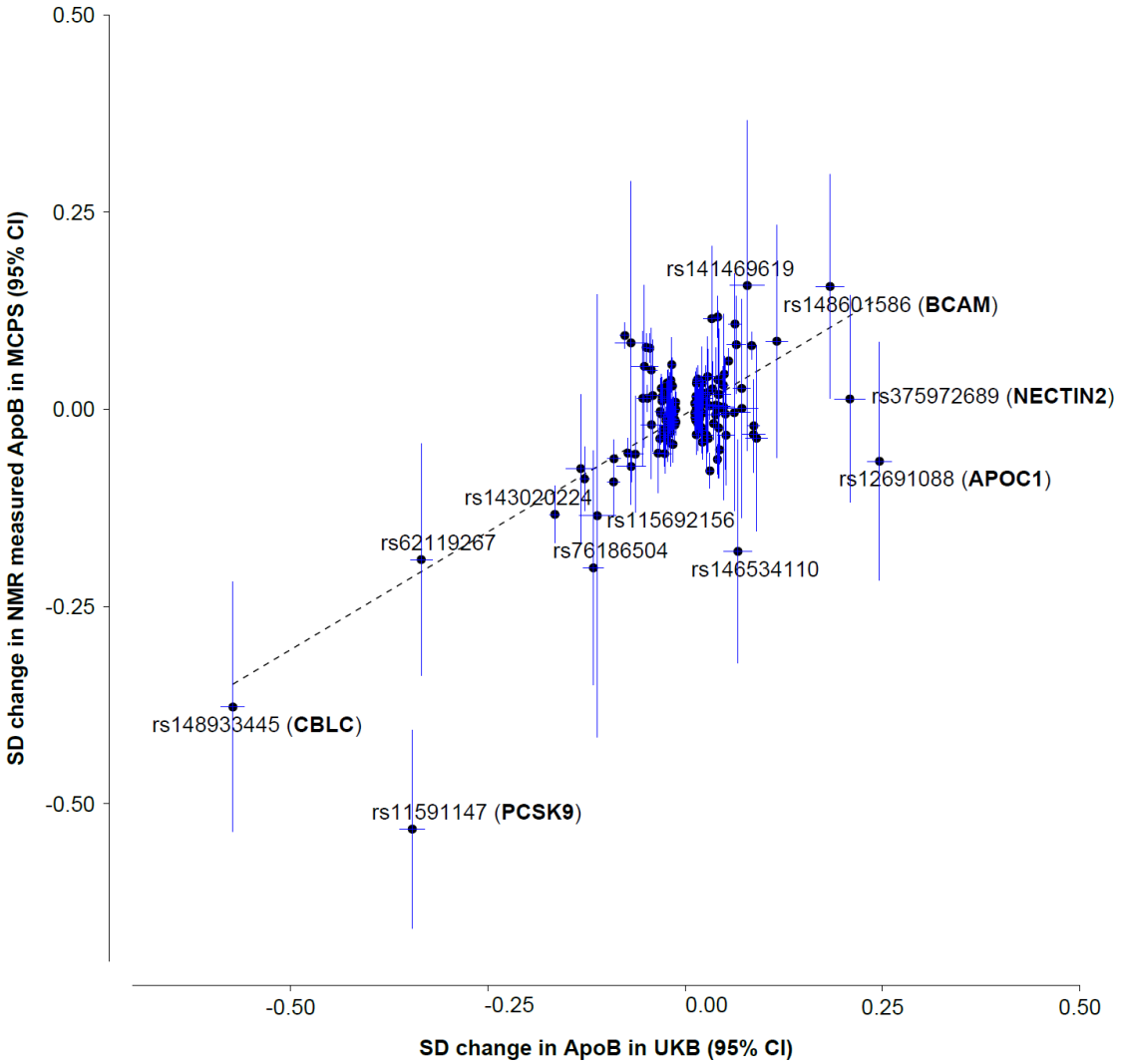
A total of 200 SNPs associated with HDL-C levels were included in the analyses as identified by GLGC GWAS (Liu et al, 2017). Effect sizes represent the standard deviation (SD) increase in HDL-C per additional allele in the GLGC and MCPS. The GLGC analyses (73 studies including >300,000 individuals) were performed by study separately and overall estimates were combined by inverse-variance weighted meta-analysis. The MCPS analyses were conducted in 29,327 individuals with no prior chronic disease and NMR lipid measurements at baseline, and with QC-passed genetic information. Adjustment were made for age, sex, principal components (GLGC: up to four; MCPS: 7) in the GLGC and MCPS. Vertical lines represent the 95% confidence intervals (CIs) of the effect estimates in the MCPS. Vertical lines represent the 95% CIs of the effect estimates in the GLGC study.

Figure 7.9 Comparison of effect sizes of genetic variants associated with TG in MCPS and GLGC GWAS



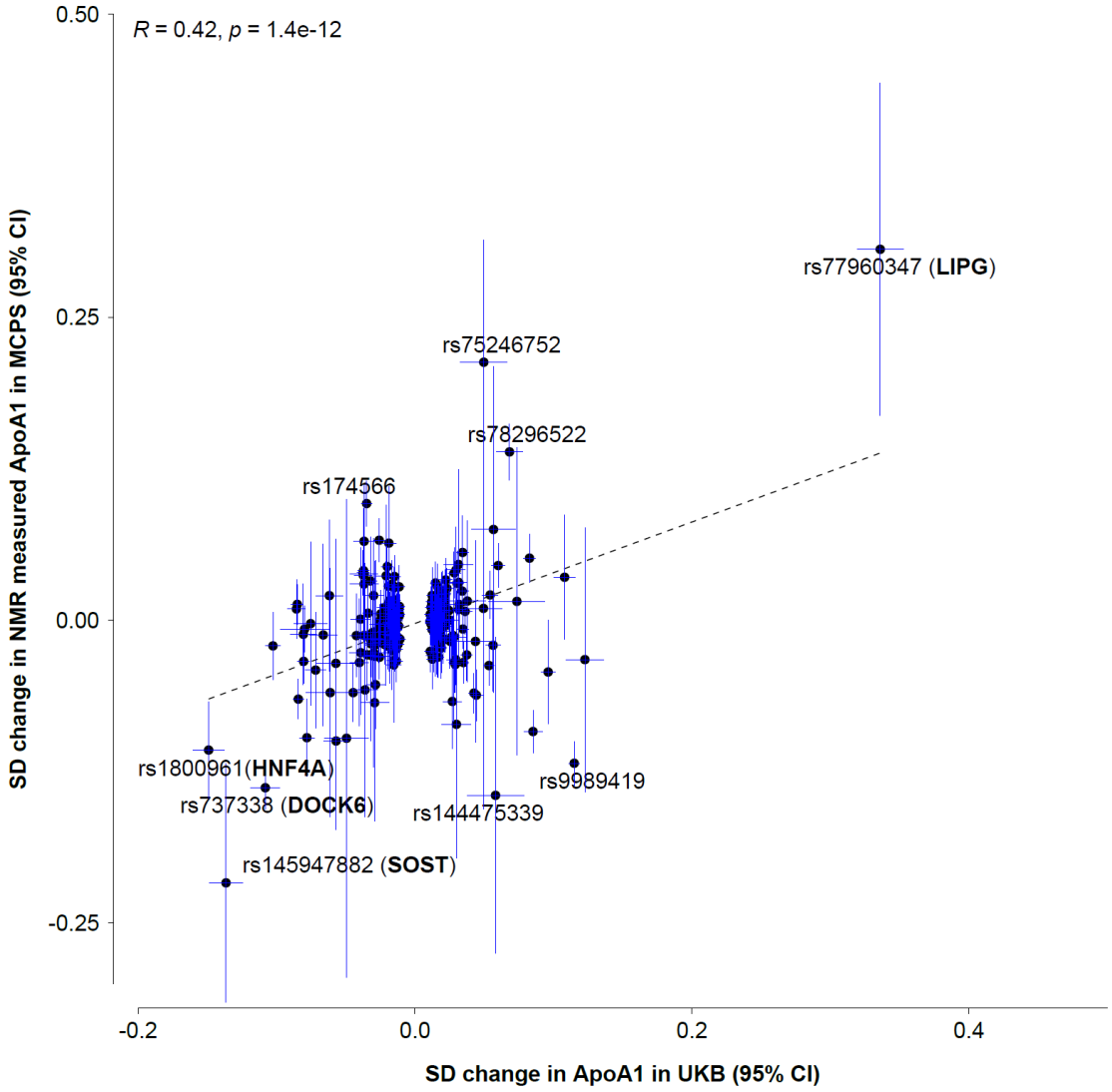
A total of 166 SNPs associated with TG levels were included in the analyses as identified by GLGC GWAS (Liu et al, 2017). Effect sizes represent the standard deviation (SD) increase in HDL-C per additional allele in the GLGC and MCPS. The GLGC analyses (73 studies including >300,000 individuals) were performed by study separately and overall estimates were combined by inverse-variance weighted meta-analysis. The MCPS analyses were conducted in 29,327 individuals with no prior chronic disease and NMR lipid measurements at baseline, and with QC-passed genetic information. Adjustment were made for age, sex, principal components (GLGC: up to four; MCPS: 7) in the GLGC and MCPS. Vertical lines represent the 95% confidence intervals (CIs) of the effect estimates in the MCPS. Vertical lines represent the 95% CIs of the effect estimates in the GLGC study.

Figure 7.10 Comparison of effect sizes of genetic variants associated with ApoB in MCPS and UKB GWAS



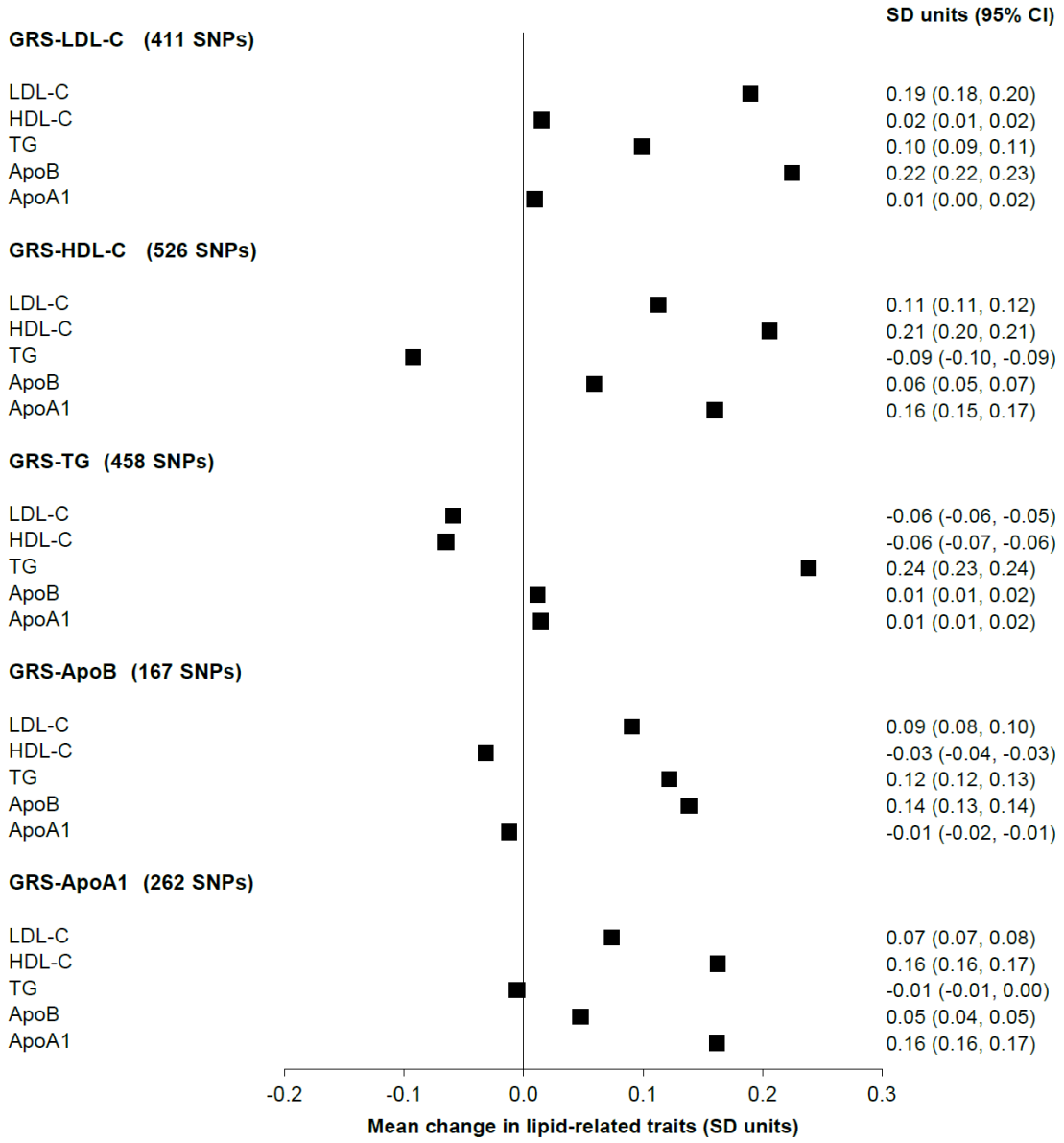
A total of 167 SNPs associated with ApoB levels were included in the analyses as identified by UKB GWAS (Richardson et al, 2020). Effect sizes represent the standard deviation (SD) increase in ApoA1 per additional allele in the UKB and MCPS. The UKB analyses (including >440,000 individuals) were performed using BOLT-LMM (linear mixed model). The MCPS analyses were conducted in 29,327 individuals with no prior chronic disease and NMR lipid measurements at baseline, and with QC-passed genetic information. Adjustment were made for age, sex, principal components (UKB: 40; MCPS: 7) in the UKB and MCPS. Vertical lines represent the 95% confidence intervals (CIs) of the effect estimates in the MCPS. Vertical lines represent the 95% CIs of the effect estimates in the UKB study.

Figure 7.11 Comparison of effect sizes of genetic variants associated with ApoA1 in MCPS and UKB GWAS



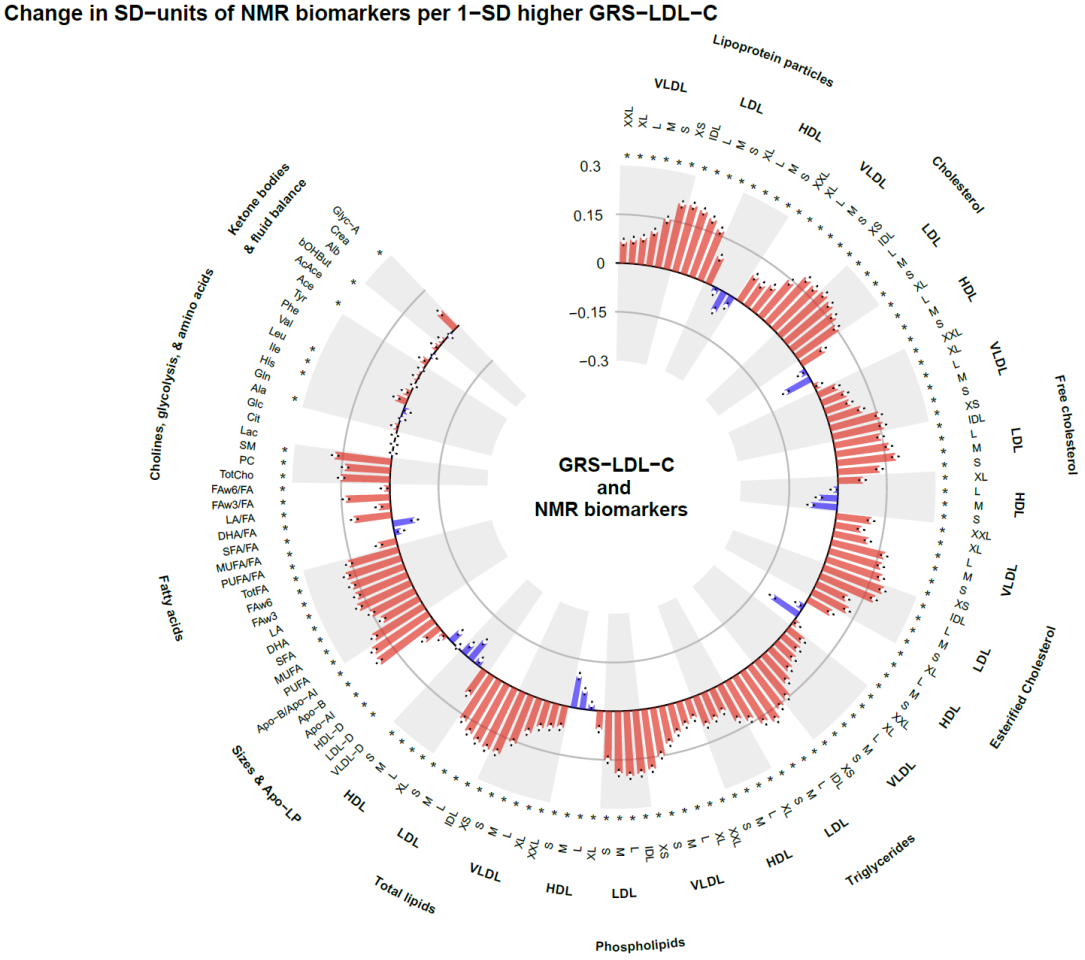
A total of 262 SNPs associated with ApoA1 levels were included in the analyses as identified by UKB GWAS (Richardson et al, 2020). Effect sizes represent the standard deviation (SD) increase in ApoA1 per additional allele in the UKB and MCPS. The UKB analyses (including >440,000 individuals) were performed using BOLT-LMM (linear mixed model). The MCPS analyses were conducted in 29,327 individuals with no prior chronic disease and NMR lipid measurements at baseline, and with QC-passed genetic information. Adjustment were made for age, sex, principal components (UKB: 40; MCPS: 7) in the UKB and MCPS. Vertical lines represent the 95% confidence intervals (CIs) of the effect estimates in the MCPS. Vertical lines represent the 95% CIs of the effect estimates in the UKB study.

Figure 7.12 Mean change in lipid-related traits per 1-SD higher weighted genetic risk scores



The GRS for major lipids included 411, 526, and 458 independent SNPs significantly associated with LDL-C, HDL-C, and TG, respectively, as reported in the GLGC 2021 GWAS. The GRS for major apolipoproteins included 167 and 262 SNPs significantly associated with ApoB and ApoA1, respectively, as reported in the UKB 2020 GWAS. The associations between weighted GRS and lipid-related traits were adjusted for age, sex, and district of residence. Effect estimates represent the SD change in each lipid-related trait per 1-SD increase in weighted genetic risk score. The MCPS analyses were conducted in 97,379 individuals with no prior chronic disease and NMR lipid measurements at baseline, and with QC-passed genetic information.

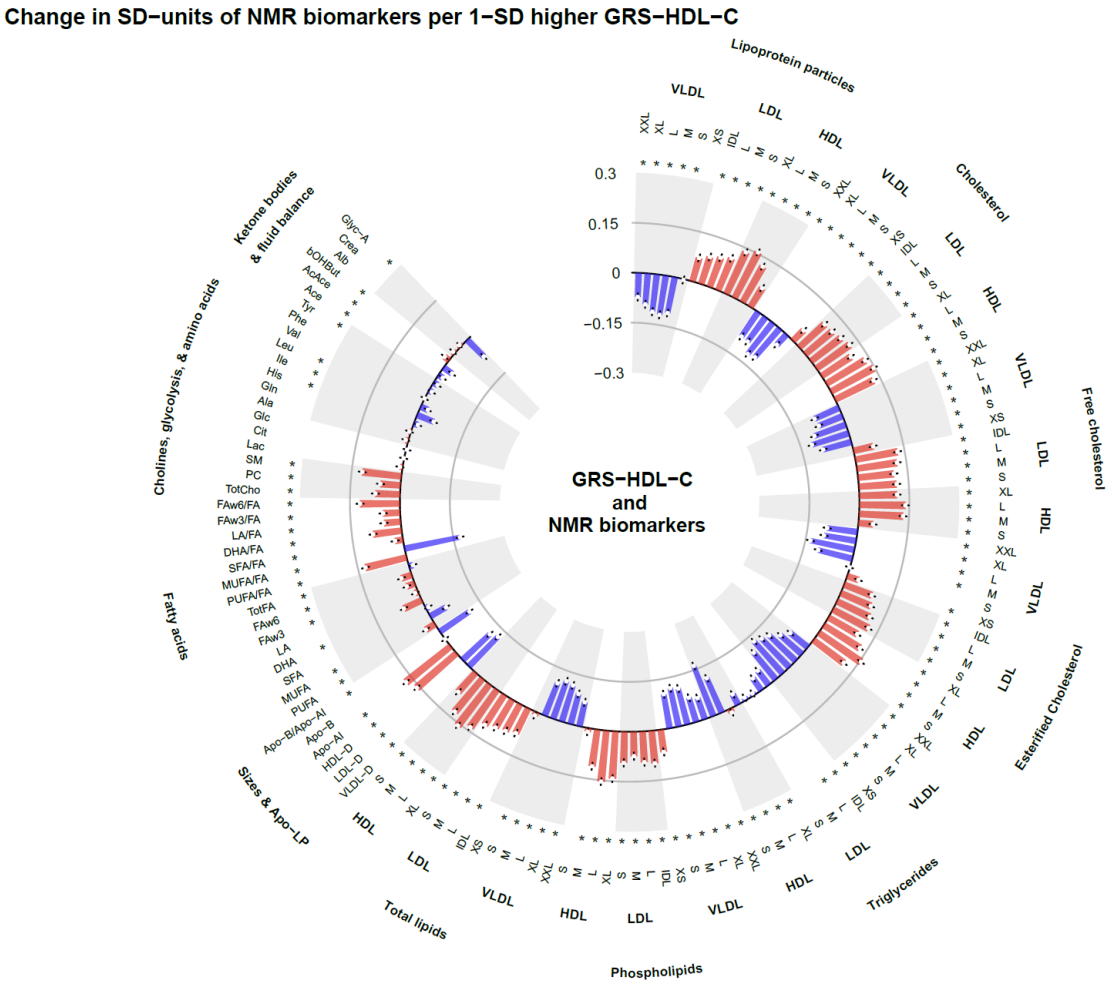
Figure 7.13 Cross−sectional associations of GRS−LDL−C and NMR biomarkers at recruitment and the effect of adjustment for confounders



No. of associations with FDR−adj. p−value <0.05: 128

All linear regression models are adjusted, where appropriate, for age, sex, and district of residence

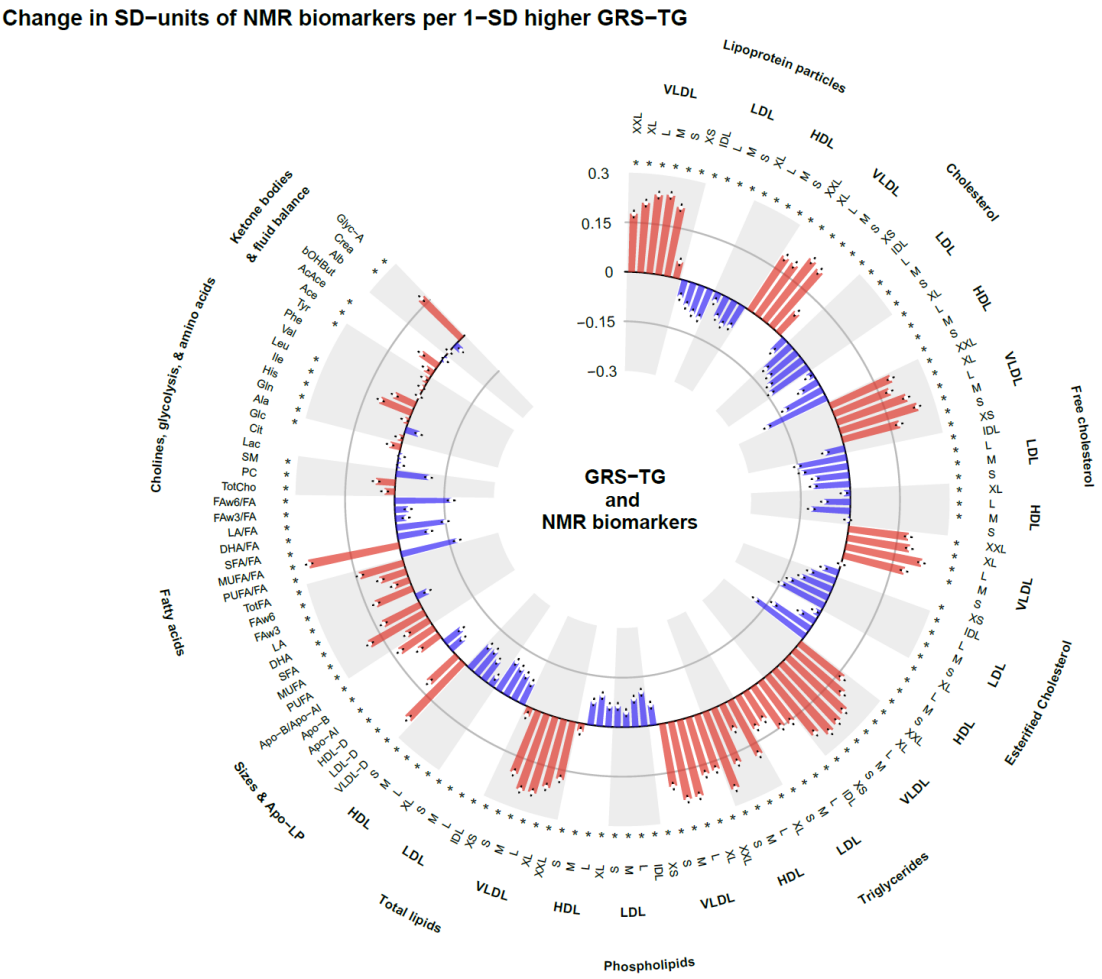
Figure 7.14 Cross−sectional associations of GRS−HDL−C and NMR biomarkers at recruitment and the effect of adjustment for confounders



No. of associations with FDR−adj. p−value <0.05: 121

All linear regression models are adjusted, where appropriate, for age, sex, and district of residence

Figure 7.15 Cross−sectional associations of GRS−TG and NMR biomarkers at recruitment and the effect of adjustment for confounders



No. of associations with FDR−adj. p−value <0.05: 131

All linear regression models are adjusted, where appropriate, for age, sex, and district of residence

Figure 7.16 Cross−sectional associations of GRS−ApoB and NMR biomarkers at recruitment and the effect of adjustment for confounders

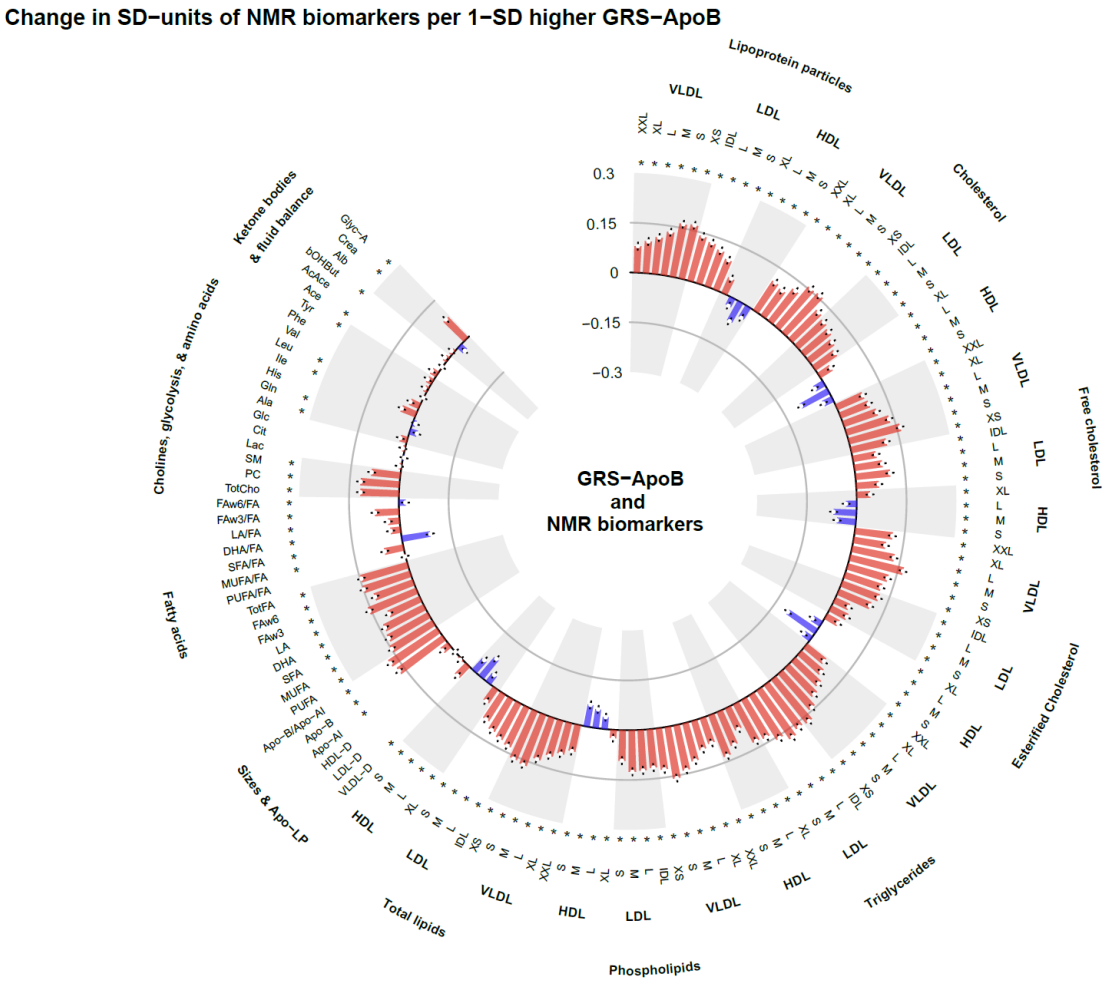
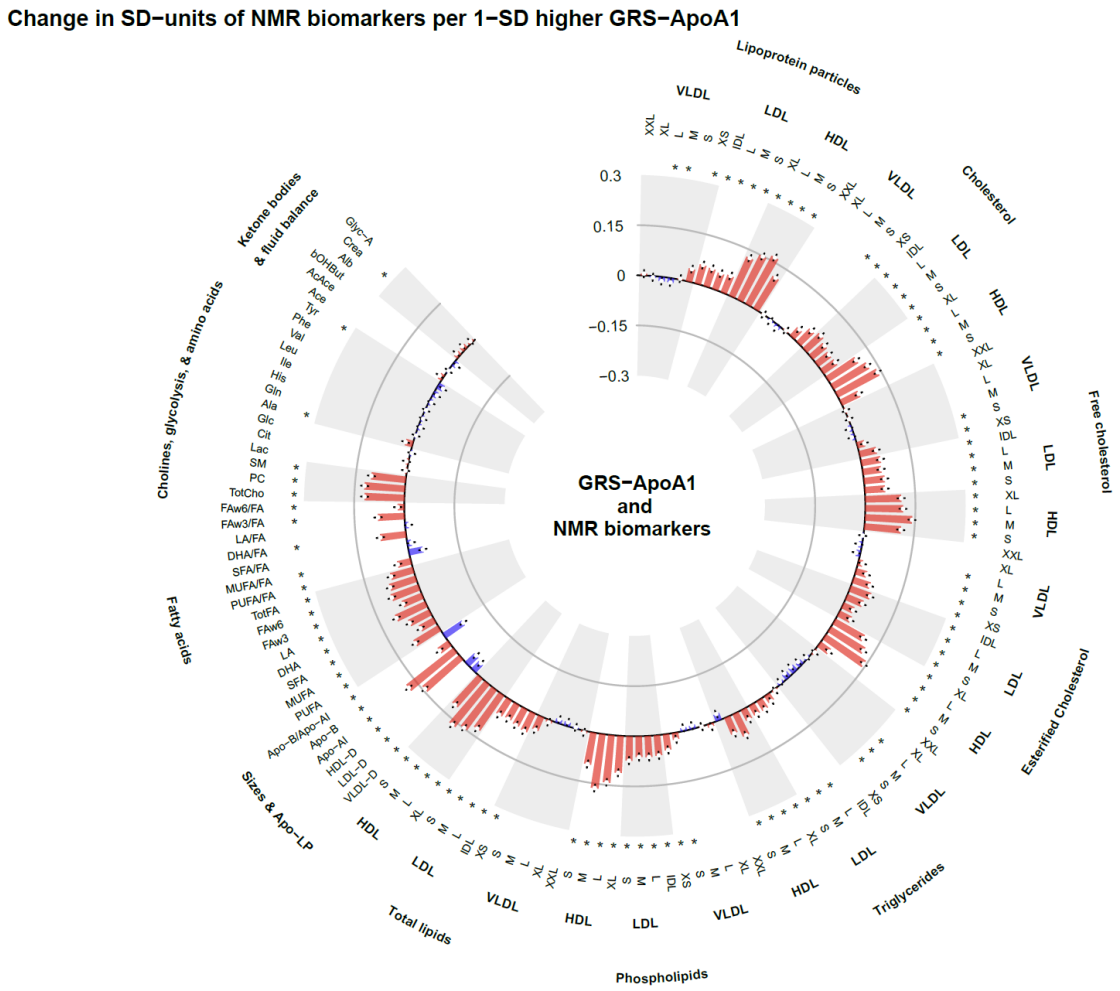


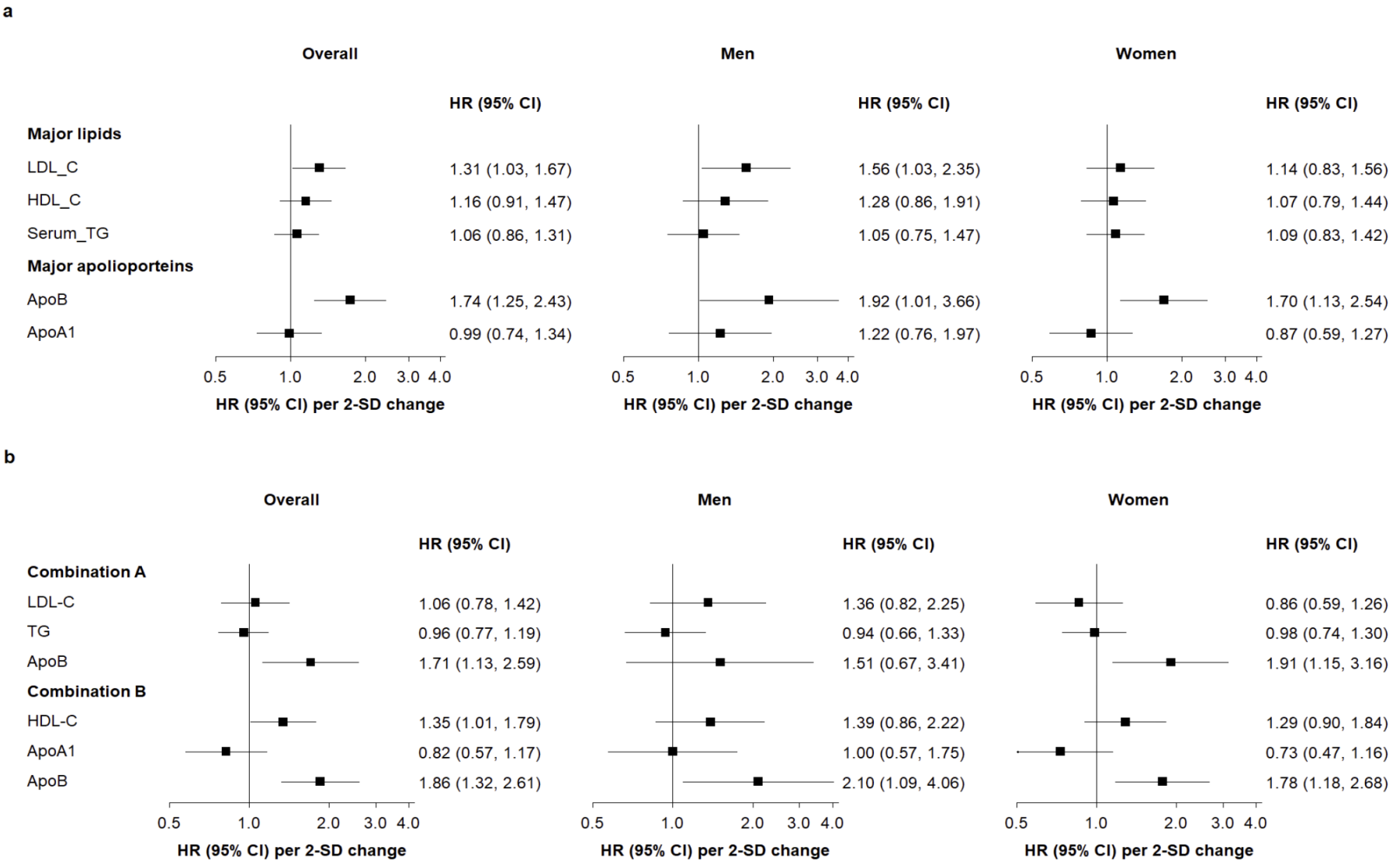
Figure 7.17 Cross−sectional associations of GRS−ApoA1 and NMR biomarkers at recruitment and the effect of adjustment for confounders



No. of associations with FDR−adj. p−value <0.05: 97

All linear regression models are adjusted, where appropriate, for age, sex, and district of residence

Figure 7.19 One-sample univariable and multivariable MR of major lipids and apolipoproteins and risk of ASCVD



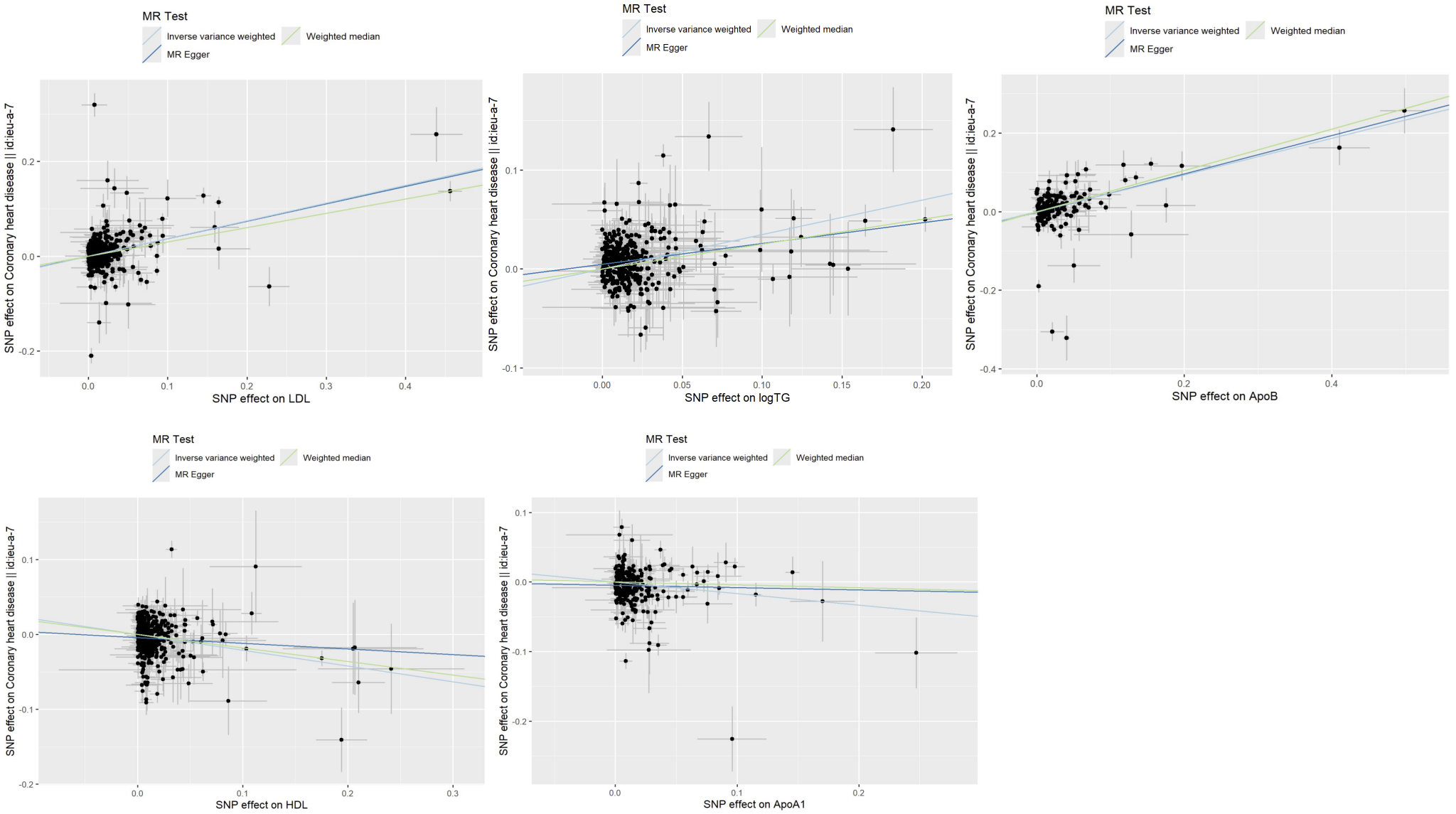
Univariable and multivariable MR estimates were derived using the 2SLS regression in MCPS using the validated GRS based on externally derived genetic variants for lipids and apolipoproteins. a, Univariable MR. b. Multivariable MR. The model for potentially atherogenic lipid traits included LDL-C, TG, and ApoB. The model for potentially protective lipid traits included HDL-C, ApoA1 (with adjustment for ApoB). The independent effect of ApoB on ASCVD was showed after accounting for ApoA1 and HDL-C in the `Protective lipid traits` group.

Figure 7.20 Sensitivity analyses of two-sample MR using pleiotropy-robust methods

Panel A: univariable MR analyses

Lipid-related traits	No. of SNPs	OR (95% CI)/ Egger intercept	P-value
LDL-C (per 2-SD increase [1 mmol/L])	367		
IVW		2.11 (1.77-2.52)	1.46E-16
Weighted median		1.83 (1.54-2.17)	4.74E-12
MR-Egger		2.08 (1.68-2.59)	1.16E-10
Intercept of MR-Egger regression		0.0003	0.85
TG (per 2-SD increase [0.8 mmol/L])	379		
IVW		2.01 (1.70-2.37)	1.40E-16
Weighted median		1.65 (1.33-2.05)	6.43E-06
MR-Egger		1.52 (1.23-1.88)	1.28E-04
Intercept of MR-Egger regression		0.005	9.31E-05
HDL-C (per 2-SD increase [0.4 mmol/L])	459		
IVW		0.66 (0.54-0.82)	1.92E-05
Weighted median		0.70 (0.56-0.87)	1.41E-03
MR-Egger		0.86 (0.68-1.09)	2.21E-01
Intercept of MR-Egger regression		-0.004	3.69E-04
ApoB (per 2-SD increase [0.4 g/L])	166		
IVW		2.54 (1.97-3.28)	5.63E-13
Weighted median		2.87 (2.31-3.56)	1.02E-07
MR-Egger		2.66 (1.88-3.75)	1.02E-21
Intercept of MR-Egger regression		-0.001	0.70
ApoA1 (per 2-SD increase [0.4 g/L])	256		
IVW		0.72 (0.56-0.91)	0.01
Weighted median		0.92 (0.73-1.16)	0.49
MR-Egger		0.94 (0.69-1.27)	0.09
Intercept of MR-Egger regression		-0.005	0.01

Panel B: Scatterplot of SNP potential effects on ASCVD vs lipid in univariable two-sample MR analyses



Panel C: multivariable MR analyses

Lipid-related traits	HR (95% CI)	P-value
Atherogenic lipid traits		
LDL-C (per 2-SD increase)	0.80 (0.53-1.19)	0.26
TG (per 2-SD increase)	1.21 (1.01-1.43)	0.03
ApoB (per 2-SD increase)	1.84 (1.30-2.60)	<0.001
Protective lipid traits with adjustment for ApoB		
HDL-C (per 2-SD increase)	1.24 (0.88-1.76)	0.22
ApoA1 (per 2-SD increase)	0.71 (0.50-1.00)	0.06
ApoB (per 2-SD increase)	1.84 (1.57-2.15)	2.23E-14

Panel A: univariable two-sample MR analyses. Panel B: Scatterplot of SNP potential effects on ASCVD vs lipid, with the slope of each line corresponding to estimated MR effect per method. Multivariable MR statistically adjusted for other lipid-related traits in the model. Panel C: multivariable two-sample MR analyses. ASCVD, atherosclerotic cardiovascular; CI, confidence interval; ApoB, apolipoprotein B; ApoA1, apolipoprotein AI; HDL, high-density lipoprotein; LDL, low-density lipoprotein; TG, triglycerides; SD, standard deviation; SNP, single nucleotide polymorphism.

References

1. 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. (In eng). DOI: 10.1002/sim.4780100703.
2. Sanderson E, Davey Smith G, Windmeijer F, Bowden J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. International Journal of Epidemiology 2018;48(3):713-727. DOI: 10.1093/ije/dyy262.
3. 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. (In eng). DOI: 10.1002/gepi.21965.
4. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. Int J Epidemiol 2017;46(6):1985-1998. (In eng). DOI: 10.1093/ije/dyx102.
5. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015;44(2):512-25. (In eng). DOI: 10.1093/ije/dyv080.
6. Sanderson E, Glymour MM, Holmes MV, et al. Mendelian randomization. Nature Reviews Methods Primers 2022;2(1):6. DOI: 10.1038/s43586-021-00092-5.
7. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol 2013;37(7):658-65. (In eng). DOI: 10.1002/gepi.21758.
8. Minelli C, Del Greco MF, van der Plaat DA, Bowden J, Sheehan NA, Thompson J. The use of two-sample methods for Mendelian randomization analyses on single large datasets. Int J Epidemiol 2021;50(5):1651-1659. (In eng). DOI: 10.1093/ije/dyab084.
9. Carnegie R, Zheng J, Sallis HM, et al. Mendelian randomisation for nutritional psychiatry. The Lancet Psychiatry 2020;7(2):208-216.
10. Brion M-JA, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. International Journal of Epidemiology 2012;42(5):1497-1501. DOI: 10.1093/ije/dyt179.
11. Graham SE, Clarke SL, Wu K-HH, et al. The power of genetic diversity in genome-wide association studies of lipids. Nature 2021;600(7890):675-679. DOI: 10.1038/s41586-021-04064-3.
12. 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 medicine 2020;17(3):e1003062. (In eng). DOI: 10.1371/journal.pmed.1003062.
13. 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. Journal of the American College of Cardiology 2012;60(25):2631-2639. DOI: https://doi.org/10.1016/j.jacc.2012.09.017.
14. Holmes MV, Asselbergs FW, Palmer TM, et al. Mendelian randomization of blood lipids for coronary heart disease. European heart journal 2014;36(9):539-550. DOI: 10.1093/eurheartj/eht571.
15. White J, Swerdlow DI, Preiss D, et al. Association of Lipid Fractions With Risks for Coronary Artery Disease and Diabetes. JAMA cardiology 2016;1(6):692-699. DOI: 10.1001/jamacardio.2016.1884.
16. Sniderman AD, Thanassoulis G, Glavinovic T, et al. Apolipoprotein B Particles and Cardiovascular Disease: A Narrative Review. JAMA cardiology 2019;4(12):1287-1295. (In eng). DOI: 10.1001/jamacardio.2019.3780.
17. Marston NA, Giugliano RP, Melloni GEM, 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 cardiology 2022;7(3):250-256. DOI: 10.1001/jamacardio.2021.5083.
18. Lincoff AM, Nicholls SJ, Riesmeyer JS, et al. Evacetrapib and Cardiovascular Outcomes in High-Risk Vascular Disease. The New England journal of medicine 2017;376(20):1933-1942. (In eng). DOI: 10.1056/NEJMoa1609581.
19. Ference BA, Kastelein JJP, Ginsberg HN, et al. Association of Genetic Variants Related to CETP Inhibitors and Statins With Lipoprotein Levels and Cardiovascular Risk. Jama 2017;318(10):947-956. (In eng). DOI: 10.1001/jama.2017.11467.
20. Frikke-Schmidt R, Nordestgaard BG, Stene MC, et al. Association of loss-of-function mutations in the ABCA1 gene with high-density lipoprotein cholesterol levels and risk of ischemic heart disease. Jama 2008;299(21):2524-32. (In eng). DOI: 10.1001/jama.299.21.2524.
21. Voight BF, Peloso GM, Orho-Melander M, et al. Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. Lancet (London, England) 2012;380(9841):572-80. (In eng). DOI: 10.1016/s0140-6736(12)60312-2.
22. Johannsen TH, Frikke-Schmidt R, Schou J, Nordestgaard BG, Tybjærg-

Hansen A. Genetic inhibition of CETP, ischemic vascular disease and mortality, and possible adverse effects. Journal of the American College of Cardiology 2012;60(20):2041-8. (In eng). DOI: 10.1016/j.jacc.2012.07.045.
23. Haase CL, Tybjærg-Hansen A, Qayyum AA, Schou J, Nordestgaard BG, Frikke-Schmidt R. LCAT, HDL cholesterol and ischemic cardiovascular disease: a Mendelian randomization study of HDL cholesterol in 54,500 individuals. The Journal of clinical endocrinology and metabolism 2012;97(2):E248-56. (In eng). DOI: 10.1210/jc.2011-1846.
24. Haase CL, Tybjærg-Hansen A, Qayyum AA, Schou J, Nordestgaard BG, Frikke-Schmidt R. LCAT, HDL cholesterol and ischemic cardiovascular disease: a Mendelian randomization study of HDL cholesterol in 54,500 individuals. The Journal of clinical endocrinology and metabolism 2012;97(2):E248-E256. (In eng). DOI: 10.1210/jc.2011-1846.
25. Voight BF, Peloso GM, Orho-Melander M, et al. Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. Lancet (London, England) 2012;380(9841):572-580. (In eng). DOI: 10.1016/S0140-6736(12)60312-2.
26. Landmesser U, Hazen S. HDL-cholesterol, genetics, and coronary artery disease: the myth of the 'good cholesterol'? European heart journal 2018;39(23):2179-2182. DOI: 10.1093/eurheartj/ehy299.
27. Zanoni P, Khetarpal SA, Larach DB, et al. Rare variant in scavenger receptor BI raises HDL cholesterol and increases risk of coronary heart disease. Science 2016;351(6278):1166-1171. DOI: doi:10.1126/science.aad3517.
28. Madsen CM, Varbo A, Nordestgaard BG. Extreme high high-density lipoprotein cholesterol is paradoxically associated with high mortality in men and women: two prospective cohort studies. European heart journal 2017;38(32):2478-2486. (In eng). DOI: 10.1093/eurheartj/ehx163.
29. Ko DT, Alter DA, Guo H, et al. High-Density Lipoprotein Cholesterol and Cause-Specific Mortality in Individuals Without Previous Cardiovascular Conditions: The CANHEART Study. Journal of the American College of Cardiology 2016;68(19):2073-2083. (In eng). DOI: 10.1016/j.jacc.2016.08.038.
30. Wilkins JT, Ning H, Stone NJ, et al. Coronary heart disease risks associated with high levels of HDL cholesterol. Journal of the American Heart Association 2014;3(2):e000519. (In eng). DOI: 10.1161/jaha.113.000519.
31. Mazidi M, Mikhailidis DP, Banach M. Associations between risk of overall mortality, cause-specific mortality and level of inflammatory factors with extremely low and high high-density lipoprotein cholesterol levels among American adults. International journal of cardiology 2019;276:242-247. (In

eng). DOI: 10.1016/j.ijcard.2018.11.095.
32. Keene D, Price C, Shun-Shin MJ, Francis DP. Effect on cardiovascular risk of high density lipoprotein targeted drug treatments niacin, fibrates, and CETP inhibitors: meta-analysis of randomised controlled trials including 117 411 patients. BMJ : British Medical Journal 2014;349:g4379. DOI: 10.1136/bmj.g4379.
33. 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). European heart journal 2019;41(1):111-188. DOI: 10.1093/eurheartj/ehz455.
34. 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 (London, England) 2010;376(9753):1670-81. (In eng). DOI: 10.1016/s0140-6736(10)61350-5.
35. Ference BA, Kastelein JJP, Ray KK, et al. Association of Triglyceride-Lowering LPL Variants and LDL-C-Lowering LDLR Variants With Risk of Coronary Heart Disease. Jama 2019;321(4):364-373. (In eng). DOI: 10.1001/jama.2018.20045.
36. Karjalainen MK, Karthikeyan S, Oliver-Williams C, et al. Genome-wide characterization of circulating metabolic biomarkers. Nature 2024;628(8006):130-138. DOI: 10.1038/s41586-024-07148-y.
37. Davies NM, Hill WD, Anderson EL, Sanderson E, Deary IJ, Davey Smith G. Multivariable two-sample Mendelian randomization estimates of the effects of intelligence and education on health. eLife 2019;8 (In eng). DOI: 10.7554/eLife.43990.

CHAPTER 8: Discussion

Chapter summary

This thesis has: 1) described the inter-correlations within the NMR lipid panel, and their correlations with traditional clinical chemistry lipids and baseline characteristics of participants (**Chapter 4**); 2) identified genetic variants associated with the levels of major lipid traits in a large, understudied population of Latin American ancestry (**Chapter 5**); 3) investigated the prospective association between NMR lipid traits and the risk of ASCVD death after accounting for confounding, reverse causality, and regression dilution bias (**Chapter 6**); and 4) used Mendelian randomization to assess the likely causal relationship of major lipid traits with ASCVD (**Chapter 7**). The main findings, methods and discussion under the wider context of literature, strengths and limitations, and conclusion are detailed in each chapter. The current discussion chapter summarizes the key findings from these chapters and how they interact and integrate with each other. The overall strengths and limitations, clinical and public health implications, and implications for future study are also described.

8.1 Summary of main findings

8.1.1 Complexity of plasma lipidome and its link with participant characteristics

By integrating the deep lipidome information measured by NMR metabolomics with the broad phenotypes collected in MCPS, the current study revealed the complexity and intercorrelated pattern within the NMR lipid profiles consisting of lipid, apolipoproteins, and lipoproteins with varying particle size, as well as their link with common sociodemographic and lifestyle factors related disease and health (**Chapter 4**). NMR lipid profiles largely clustered according to their biochemical structures and interaction in lipid metabolism. The complexity of the plasma lipidome makes the interpretation of the independent effects of individual lipid trait challenging but could elucidate the potential heterogeneous role of lipid subfractions in the development of ASCVD. The five major composite NMR-measured lipid traits (LDL-C, TG, HDL-C, ApoB, and ApoA1) that are the main exposures in this thesis were highly correlated with the traditional lipids measured by clinical chemistry assays, which supports the robustness of the NMR lipid quantification and their validity when comparing results with previous research mainly based on traditional lipid assays. The NMR lipid traits were related to a range of other baseline characteristics, such as age, sex, region, education, and smoking, which are key risk factors for ASCVD, reflecting that the lipid metabolites are broadly influenced by environmental exposome and these factors should be fully accounted for in the current and future research. Of note, age- and sex- differences were observed across lipid traits, possibly explained by the variation in sexual hormones and other sex-specific factors. Stratification on age and sex should be considered in the prediction of ASCVD risk.[1] In subsequent observational analyses, certain sex differences in the relationships of the lipids to risk were observed, particularly at younger ages at risk.

8.1.2 Genetic discovery of major lipid traits

GWAS for five major NMR-measured lipid traits identified multiple previously known

genetic loci involved in lipid metabolism and numerous potentially novel loci. The current GWAS identified 135 conditionally-independent signals (36 potentially novel) at 92 loci for LDL-C (e.g., *APOE, CELSR2, PCSK9, APOB, HMGCR, LPA, LPL, LDLR*), 202 independent signals (70 novel) at 135 loci for HDL-C (e.g., *CETP, SIK3, LIPC, LPL, APOA5*), 118 independent signals (36 novel) at 75 loci for TG (e.g., *ZPR1, GCKR, APOC1P1, APOC3, LPA, DOCK6*), 129 independent signals (51 novel) at 92 loci for ApoA1 (e.g., *SIK3, CETP, ANGPTL8, LIPC, ABCA1*), and 160 independent signals (67 novel) at 113 loci for ApoB (e.g., *CELSR2, APOE, APOB, LDLR, PCSK9, CETP, LPA*). The numbers of independent, significant signals identified for each lipid trait in the current study are generally higher than findings in previous studies of a similar size,[2-4] suggesting that including underrepresented populations with greater genetic diversity (rather than studying additional individuals of European ancestry) may improve the genetic discovery and enhance the fine-mapping of functional variants. Considering the lack of representation of diverse non-European ancestry population (e.g., Latin-American) in previous GWAS of lipid traits, future investigations should further evaluate the improvement of genetic discovery by including multi-ancestry populations to increase genetic diversity.

8.1.3 Association of LDL-C with ASCVD

Evidence from genetic studies, prospective cohort studies, Mendelian randomization studies, and randomized clinical trials of lipid-lowering therapies have shown that LDL and other ApoB-containing lipoproteins (VLDL remnants, IDL, and Lp(a)) have causal effect on ASCVD.[5] In the current Latin American population with a high prevalence of obesity and diabetes, when assessed individually in observational analyses, LDL-C was *not* associated with the risk of ASCVD death. After accounting for the impact of other lipids, the main findings were that, given ApoB, usual LDL-C was *negatively* associated with the risk and given LDL-C, usual ApoB was strongly, *positively* associated with risk. These findings implicate smaller, denser LDL particles as being particularly atherogenic,

which is supported by examination of LDL diameter directly. When looking at direct measures of LDL subtypes and their cholesterol contents, the cholesterol in small but not large LDL was related to increased risk. These findings are in line with the variability of ApoB-containing particles (cholesterol content of LDL particles varied) and the *discordance* changes between the levels of LDL-C and ApoB/LDL particles under certain conditions where TG levels exceeded 150 mg/dL (e.g., the metabolic syndrome, Type 2 diabetes, and hypertriglyceridemia), as detailed in **Chapter 1**. LDL-C was positively related to the ASCVD risk in univariable MR analyses, but the effect decreased to null after accounting for the effects of ApoB and TG in multivariable MR analyses.

In most cases (including the current study), the LDL-C concentration is correlated with the number of LDL particles and, hence, the ApoB level, and thus provides similar information on the risk of ASCVD.[5] However, the current study suggests that relying solely on LDL-C measurements in the context of insulin resistance—such as in cases of obesity, diabetes, or elevated triglycerides, which may increase the proportion of smaller, denser LDL-C—could lead to discordance between LDL-C and ApoB levels, potentially resulting in suboptimal ASCVD risk identification. Direct measures of lipoprotein particle subfractions could be more informative than conventional clinical lipid measures.[5-7] ApoB, as the most accurate and direct measure of ApoB-containing particle concentration, is more closely associated with the cardiovascular risk. The superiority of ApoB over LDL-C and other lipid measures is discussed in detail later in the discussion.

8.1.4 Association of TG with ASCVD

Previous evidence from observational studies and randomized trials does not provide consistent evidence that lowering TG reduces the risk of CVD events.[8-12] In the current observational analyses, when assessed individually, TG levels was related to increased risk of ASCVD, but this association was attenuated to the null after accounting for ApoB and LDL-C. This could be explained by the interrelated atherogenic changes in

lipoproteins (e.g., low HDL-C and high level of, small, dense cholesterol-depleted LDL particles) observed at a high TG levels. In both the univariable and multivariable MR analyses, TG was not significantly related to the ASCVD risk. In multivariable MR including ApoB, TG, and LDL-C, only ApoB was positively and causally related to risk. The current results are consistent with other previous observational and genetic studies that have found that the association of TG within TG rich lipoproteins with risk of ASCVD often becomes non-significant after adjustment for HDL-C and non HDL-C/ApoB. The clinical benefit of TG-lowering therapy (e.g., fibrates, niacin, or OM3FAs) was only shown in subgroups of patients with high baseline TG levels (the corresponding reduction in apoB-containing particles would only reach clinical significance in those with hypertriglyceridemia).[13-19] For example, in the ACCORD trials of fenofibrate vs placebo in patients with type 2 diabetes who were at high risk of ASCVD, there was a possible interaction according to lipid subgroup and a potential benefit for patients with high TG levels and low HDL-C.[14] Recent MR analyses have further demonstrated that TG-lowering *LPL* variants and LDL-C-lowering *LDLR* variants were associated with similar lower risk of ASCVD per unit change in apoB-containing lipoproteins and their effects were independent and additive, suggesting the potential clinical benefit of lowering TG levels is similar to the benefit of lowering LDL-C per unit change in apoB.[12] Overall, the current study, in line with previous evidence, suggests that the effect of TG may not be causal and the risk reduction of ASCVD related to TG within the ApoB-containing particles should be proportional to the relevant reduction in ApoB per se.

8.1.5 Association of HDL-C and ApoA1 with ASCVD

Current evidence from epidemiological and genetic studies on the HDL-related traits (HDL particles, HDL-C, and ApoA1) has changed dramatically over recent years and remains inconsistent, with inverse, U-shaped, or positive (at high level of HDL-C) associations observed.[20-29] In the current observational analyses of MCPS, although higher HDL-P and HDL-C were associated with a lower risk of ASCVD overall, larger

HDL particles (which make up most of the HDL cholesterol in plasma) were associated with *increased* risk. In multivariable MR analyses including ApoB, ApoA1 and HDL-C, a positive association with HDL-C was observed. These findings are partially consistent with the previous understanding of HDL function in reverse cholesterol transport,[30,31] suggesting that HDL particles with different particle sizes may have a heterogeneous role in ASCVD (due to different potential roles in pathophysiological functions and atherosclerosis). The association of HDL-P and HDL-C with ASCVD risk warrants further investigation, especially using the direct of measures of more specific HDL subfractions, to elucidate their heterogeneity in biological functions and causality in the development of ASCVD.

8.1.6 Causality of ApoB on ASCVD

In line with previous evidence,[32-35] the current observational and MR findings collectively suggest the superiority of ApoB over cholesterol as a marker for ASCVD based on the variability in ApoB-containing particles, the discordant changes in LDL-C and ApoB, and the observed differences in their association with risk of ASCVD. In observational analyses, after accounting for LDL-C, ApoB was strongly and log-linearly positively associated with risk of ASCVD in both men and women. S-LDL-C (small, dense LDL-C) was more strongly related to ASCVD mortality than L-LDL-C, which was further supported by the observation that per unit change of S-LDL-P and L-LDL-P had similar effects to each other. These findings, in conjunction with previous substantial evidence and atherosclerosis model of ApoB particles,[32,33] suggests that each ApoB-containing particle is related to similar risk and are equally atherogenic. Further MR analyses confirmed that ApoB, as a direct and accurate measures of the number of apoB particles, is causally related to the risk of ASCVD, independent of lipid contents (such as TG and cholesterol) and of HDL or ApoA1. Collectively, these findings indicate that the risk of ASCVD is more closely and causally related to the number of circulating apoB particles than with the variable mass of cholesterol or TG within them. Reducing the number of

ApoB-containing particles should remain the focus of efforts to reduce the risk and increasing burden of ASCVD.

8.1.7 Overall strengths and limitations

The main strength of the current study is that it is derived from a large prospective cohort with detailed genetic and non-genetic data and a prolonged duration of follow-up in a previously understudied Latin-American population. The analyses shown comprehensively and robustly investigate the observational and causal associations between NMR-measured lipid traits and ASCVD death, after accounting for potential confounding, reverse causality, and regression dilution bias. The use of NMR-measured plasma lipids provides additional information on the heterogeneity of lipids and lipoproteins with varying size and their relationships with ASCVD risk. The main exposures included a range of NMR lipid-related traits from a clinical perspective (using composite measures similar to chemical assays) and at a more refined resolution (using lipid subfractions of varying sizes and potential biological functions). Use of lipid-lowering drugs at recruitment was low and so exclusion of such participants from the analyses would not have introduced any material theoretical selection biases. The MR analyses utilized the most recent data from a large multi-ancestry lipid GWAS to construct genetic instruments for composite lipid-related traits. Although the variants significantly associated with lipid traits were identified in external populations predominantly of European ancestries, they were validated in MCPS to be robust instruments (e.g., they were strongly related to the measured baseline lipid levels and were not associated with other baseline characteristics). Multivariable MR was conducted to account for horizontal pleiotropy and assess the independent causal effects of each blood lipid measure. The large sample size, rigorous statistical methods, and integration of observational and Mendelian randomization findings enhance the causal inference and strengthen the robustness of the current findings.

However, the findings from the current study should be interpreted with caution in the context of its key limitations. With respect to the potentially novel loci identified in the GWAS analyses, no attempt was made to replicate these findings or conduct further downstream analyses to investigate gene function and enrichment (as this was deemed beyond the scope of the current thesis). Further research is therefore needed to discover which of these loci replicate with subsequent downstream analyses exploring the function of any that do. The ASCVD mortality outcome in the current analyses mainly comprised CHD, which limited the ability to conduct explore variation by vascular domain (e.g., CHD, ischemic stroke, intracerebral hemorrhage, and PAD). In addition, no information was available on non-fatal ASCVD events and so the conclusions apply directly only to fatal ASCVD. Regarding the MR analyses, due to the lack of valid genetic instruments for lipid subfractions (current GWAS on NMR lipid subfractions are of limited sample sizes compared to those from large lipid consortium on composite lipids), composite major lipid traits but not lipid subfractions were used as the main exposures. It is likely that specific HDL subtypes could play an important role as observed for small HDL particles in observational analyses, which warrant further bespoke MR studies on lipid subfractions when more GWAS of ASCVD in Latin American populations are available. Finally, at the time of analysis, the total study sample included only approximately 80% of the full cohort (~120,000 of ~150,000 participants). Future research utilizing the full cohort with NMR measurements should further enhance discovery potential, although the current sample size is still large by most standards.

8.2 Clinical and public health implications

8.2.1 Lipid management and ASCVD control in Mexico

Latin-American populations have been found to have higher ASCVD risk compared with North American and European populations.[36] The findings of the current study provide direct evidence for lipid management, ASCVD risk control, and health resource allocation

in Mexico, where the prevalence of overweight and obesity, diabetes, and chronic kidney disease are particularly high. In Mexico, ASCVD remains the leading cause of death and the second highest cause of outpatient visits.[36] Despite dyslipidemia and other related risk factors such as obesity and diabetes being identified as significant risk factors for ASCVD in Mexico a large proportion of individuals are still not treated with lipid-lowering therapy while, for those who are treated, many do not reach LDL-C thresholds recommended by guideline.[36,37] In MCPS, among participants with a previous diagnosis of diabetes, only 1% were using lipid-lowering drugs at the time of recruitment (1998-2004) while this was still only 14% at resurvey (2015-2019). Indeed, even among those with both diabetes and cardiovascular disease at resurvey, only about one third were taking a lipid-lowering drug.[20] A Mexican registry cohort[36] reported that 96% of the population were classified as being at high or very high risk for ASCVD, yet 63% were not receiving any lipid-lowering treatment. Among patients with ASCVD, only about a quarter had LDL-C levels below 55 mg/dL. Given the rising trend of ASCVD in Mexico[38] and the poor lipid management, targeted intervention strategies—such as promoting adherence to modifiable healthy lifestyle factors like physical activity and smoking cessation, along with the use of lipid-lowering drugs like statins—are urgently needed to reduce the burden of ASCVD and related diseases (e.g., diabetes and chronic kidney diseases).

8.2.2 ApoB vs LDL-C as the primary measures of risk related to apoB-containing particles

In most cases, LDL-C is physiologically related and highly correlated with ApoB (concordant change of two markers) and provide similar information on ASCVD risk.[5-7] Therefore, LDL-C is widely used as an acceptable surrogate for LDL and apoB and has been the primary markers of clinical effectiveness in major clinical trials of lipid-lowering drugs.[39-41] However, due to the variability of apoB-containing particles (cholesterol content of LDL particles varied but each has one exact ApoB), in certain conditions, discordant change of ApoB and LDL-C could bias the risk related to apoB-containing

particles. As observed in the current and previous studies,[5,7,35] in population with extremely high levels of TG, obesity, and diabetes, the proportion of smaller and denser LDL-C is increased and risk prediction based on LDL-C only could lead to biased risk estimates. In line with the steadily mounted evidence from clinical and genetic studies, [5,7,34,35] the current findings from MCPS highlight ApoB quantifying the total concentration of atherosclerotic apoB-containing particles is a more accurate and causal maker for risk prediction of ASCVD beyond their lipid contents (e.g., LDL-C, TG, and non-HDL-C). Also, evidence from clinical trials and drug simulation MR study supported the superiority of apoB in both primary and secondary prevention.[12,33-35,42] For example, FOURIER (Further Cardiovascular Outcomes Research With PCSK9 Inhibition in Subjects With Elevated Risk) and IMPROVE-IT (Improved Reduction of Outcomes: Vytorin Efficacy International Trial) demonstrated that ApoB more accurately predict ASCVD risk than LDL-C and non-HDL-C.[34,35] Drug simulation MR studies on combination of CETP inhibitors and statins, or combination of cholesterol- or TG-lowering treatments confirm that any clinical benefit of lowering TG or cholesterol should be proportional to the absolute change in ApoB.[12,33] In addition, LDL-C is an incomplete—and not the most accurate—marker for assessing the clinical benefits of statins, ezetimibe, or PCSK9 inhibitors, as these drugs reduce both LDL ApoB and VLDL ApoB levels. Although LDL ApoB accounts for the majority of total ApoB, the absolute risk attributable to triglyceride-rich lipoproteins may be particularly high in populations with extremely elevated TG levels and high prevalence of obesity and diabetes (e.g., MCPS). Another proposed argument for the superiority of ApoB is its significantly lower susceptibility to measurement error. ApoB can be measured more accurately, affordably, and rapidly using standardized, automated methods that are widely available, whereas accurate LDL-C measurement remains challenging, with still ongoing efforts to improve its estimation through new equations.[35,43] Overall, ApoB can be accurately quantified as a direct measure of total concentration of apoB-containing particles that has causal effect

on ASCVD, and better predict risk beyond LDL-C. ApoB should be the primary marker of atherogenic risk related to apoB particles and adequacy of lipid-modifying therapy.

However, prioritizing ApoB does not mean abandoning the conventional lipid panel such as LDL-C and non HDL-C. Despite not ideal, LDL-C and non HDL-C still can be considered as less accurate surrogate marker of apoB particle number but still useful and practical measurement for risk prediction. The additional cost of measuring ApoB beyond the current clinical lipid panel should also be considered, which may impact the feasibility and cost-effectiveness of the routine implementation of ApoB in clinical practice.[44] Over the past few decades, a substantial body of evidence on lipids and ASCVD has been built, with much of the focus centred on LDL-C. For example, a meta-analysis of individual participants from 26 randomised trials of statin therapy including almost 170,000 individuals demonstrated that lowering LDL-C reduces the risk of CVD events proportionate to the absolute reduction in LDL-C, and independently of other risk factors (22% reduction in CVD risk per mmol/L reduction in LDL-C over a median treatment period of 5 years).[39,40] Building on this extensive body of trials focused on LDL reduction and ASCVD, the safety and efficacy of statin treatment have been reliably evaluated across various patient populations.[40] The superiority of ApoB support the paradigm to shift in the clinical practice (e.g., transiting from the “bad cholesterol content” to “number of bad cholesterol particles”).[35] However, this does not undermine the substantial body of evidence derived from reliable LDL-C studies.[39,40] The 2019 guidelines from the European Society of Cardiology and the European Atherosclerosis Society clearly assert that apoB is the most effective marker for identifying atherogenic lipoproteins.[45] Further reference ranges for apoB to determine when to initiate or enhance lipid-lowering therapy are needed, similar to the current practice with LDL-C. Specifically, in resource-limited environments in Mexico, where healthcare professionals may be more accustomed to using LDL-C, it is crucial to apply LDL-C lowering treatment cautiously. This approach should consider the potential biases introduced by the highly

prevalent insulin-resistant conditions such as obesity and diabetes.

It is worth noting that non-HDL-C may be more informative than LDL-C, as it is more closely associated with the total number of apoB-containing particles. In the current study, the patterns of association with non-HDL-C were largely similar to those observed with ApoB, though the associations were generally weaker. One recent study consistently reported that residual ApoB remained significantly associated with the risk of ASCVD after accounting for LDL-C and non-HDL-C.[46] Overall, the use of ApoB or non-HDL-C (or LDL-C) should take into count their relative cost-effectiveness and feasibility for implementation in clinical practice, especially in resource-limiting setting.[44]

8.2.3 Lowering the burden of cardiovascular diseases by targeting apoB-containing particles

The current study in combination with previous research suggests that ApoB is the strongest causal lipid-related factor and should be the primary measure of the adequacy of therapies that target apoB-containing lipoproteins. Statins are highly effective in managing cardiovascular disease by reducing the production of VLDL in the liver and enhancing the clearance of VLDL remnants and LDL through the upregulation of LDL receptors, which lowers the circulating levels of atherogenic apoB-containing particles.[47] Intermediate dose statin treatment has an acceptable cost-effectiveness profile and has been prioritized as a key intervention in the primary and secondary prevention of cardiovascular events.[39,40,48-50] Given the accumulation of ASCVD burden and other related chronic diseases, even the modest reductions in the ApoB and LDL-C levels among the large populations (e.g., by scaling up statin use) are likely to significantly prevent ASCVD events and save health care costs.

8.3 Implications for future research

Future research is warranted to fill the knowledge gap and limitations of the current study. First, further GWAS on blood lipids should include diverse ancestry groups and increase

the number and representativeness of non-European populations to improve the genetic discovery and enhance the fine-mapping of functional variants. Second, future studies including incident disease outcomes beyond mortality only and have longer duration of follow-up to enable more detailed investigation of the association between plasma lipid traits and domain-specific ASCVD. Third, as the current study indicated that LDL and HDL particles of varying sizes may have different functions and role in the development of ASCVD, future MR studies with bespoke study design and statistical methods such as MR-BMA are needed to elucidate the causal effects of blood lipids at subfraction resolution given the valid genetic instruments for lipid subfractions are increasing available. Finally, the current NMR panel included only about 230 metabolites including mainly lipids and other small molecules. The current study focused solely on lipid related traits and future studies should assess the association between additional metabolites and ASCVD, possibly integrating other omics layers such as proteomics and epigenetics, to uncover the molecular mechanism underlying atherosclerosis and ASCVD for better prevention and treatment.

Another area worthy of future research is the potential heterogeneity of associations of lipids with ASCVD depending on sex. Although there is no consensus that atherosclerotic lipids such as LDL-C is better (or worse) predictors of ASCVD in women than in men,[1] the current findings from MCPS are largely in line with insights from other large prospective studies.[50] These findings of sex heterogeneity may be partly explained by a differential impact of traditional CVD risk factors in men compared with women (e.g., diabetes and obesity), and the impact of sex-specific risk factors (e.g., polycystic ovary syndrome, gestational diabetes, and early menopause). Women are also disproportionately affected by chronic kidney disease and auto-immune inflammatory diseases (both risk factors for ASCVD).[1] Gender-related sociocultural contributors may also contribute to the observed heterogeneity. Emerging experimental evidence from basic research suggests that sex (e.g., sex hormones and X-chromosome) play a

significant role in the development of atherosclerosis. Further research is required to explore the detailed mechanisms behind these sex differences in ASCVD risk related to blood lipids.

8.4 Conclusion

In this large blood-based prospective study of Mexican adults complementary findings from traditional and genetic epidemiological analyses implicate that ApoB as the strongest causal lipid factor for ASCVD. Direct measures of the number of different lipoproteins (e.g., LDL) and their subtypes may be more relevant to risk than the total amount of cholesterol within these particles. Future studies are needed to determine the causality of these lipid subfractions. Enhancing lipid management and controlling ASCVD through the wider use of lipid-lowering drugs, such as statins targeting ApoB particles, is urgently needed in this understudied and undertreated Latin American population. This approach is crucial to alleviating the accumulating burden of ASCVD and related chronic diseases on a large scale and improving overall population well-being.

Reference

1 Roeters van Lennep, J. E. *et al.* Women, lipids, and atherosclerotic cardiovascular disease: a call to action from the European Atherosclerosis Society. *European heart journal* **44**, 4157-4173, doi:10.1093/eurheartj/ehad472 (2023).

2 Graham, S. E. *et al.* The power of genetic diversity in genome-wide association studies of lipids. *Nature* **600**, 675-679, doi:10.1038/s41586-021-04064-3 (2021).

3 Ziyatdinov, A. *et al.* Genotyping, sequencing and analysis of 140,000 adults from Mexico City. *Nature* **622**, 784-793, doi:10.1038/s41586-023-06595-3 (2023).

4 Kamiza, A. B. *et al.* Multi-trait discovery and fine-mapping of lipid loci in 125,000 individuals of African ancestry. *Nature communications* **14**, 5403, doi:10.1038/s41467-023-41271-0 (2023).

5 Ference, B. A. *et al.* Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. *European heart journal* **38**, 2459-2472, doi:10.1093/eurheartj/ehx144 (2017).

6 Welsh, C. *et al.* Comparison of Conventional Lipoprotein Tests and Apolipoproteins in the Prediction of Cardiovascular Disease. *Circulation* **140**, 542-552, doi:10.1161/circulationaha.119.041149 (2019).

7 Sniderman, A. D. *et al.* Apolipoprotein B Particles and Cardiovascular Disease: A Narrative Review. *JAMA cardiology* **4**, 1287-1295, doi:10.1001/jamacardio.2019.3780 (2019).

8 Miller, M. *et al.* Triglycerides and cardiovascular disease: a scientific statement from the American Heart Association. *Circulation* **123**, 2292-2333, doi:10.1161/CIR.0b013e3182160726 (2011).

9 Sandesara, P. B., Virani, S. S., Fazio, S. & Shapiro, M. D. The Forgotten Lipids: Triglycerides, Remnant Cholesterol, and Atherosclerotic Cardiovascular Disease Risk. *Endocrine reviews* **40**, 537-557, doi:10.1210/er.2018-00184 (2019).

10 Sarwar, N. *et al.* Triglycerides and the risk of coronary heart disease: 10,158 incident cases among 262,525 participants in 29 Western prospective studies. *Circulation* **115**, 450-458, doi:10.1161/circulationaha.106.637793 (2007).

11 Hokanson, J. E. & Austin, M. A. Plasma triglyceride level is a risk factor for cardiovascular disease independent of high-density lipoprotein cholesterol level: a meta-analysis of population-based prospective studies. *Journal of cardiovascular risk* **3**, 213-219 (1996).

12 Ference, B. A. *et al.* Association of Triglyceride-Lowering LPL Variants and LDL-C-Lowering LDLR Variants With Risk of Coronary Heart Disease. *Jama* **321**, 364-373, doi:10.1001/jama.2018.20045 (2019).
13 Manninen, V. *et al.* Joint effects of serum triglyceride and LDL cholesterol and HDL cholesterol concentrations on coronary heart disease risk in the Helsinki Heart Study. Implications for treatment. *Circulation* **85**, 37-45, doi:10.1161/01.cir.85.1.37 (1992).
14 Ginsberg, H. N. *et al.* Effects of combination lipid therapy in type 2 diabetes mellitus. *The New England journal of medicine* **362**, 1563-1574, doi:10.1056/NEJMoa1001282 (2010).
15 Robins, S. J. *et al.* Relation of gemfibrozil treatment and lipid levels with major coronary events: VA-HIT: a randomized controlled trial. *Jama* **285**, 1585-1591, doi:10.1001/jama.285.12.1585 (2001).
16 Guyton, J. R. *et al.* Relationship of lipoproteins to cardiovascular events: the AIM-HIGH Trial (Atherothrombosis Intervention in Metabolic Syndrome With Low HDL/High Triglycerides and Impact on Global Health Outcomes). *Journal of the American College of Cardiology* **62**, 1580-1584, doi:10.1016/j.jacc.2013.07.023 (2013).
17 Bucher, H. C., Hengstler, P., Schindler, C. & Meier, G. N-3 polyunsaturated fatty acids in coronary heart disease: a meta-analysis of randomized controlled trials. *The American journal of medicine* **112**, 298-304, doi:10.1016/s0002-9343(01)01114-7 (2002).
18 Rizos, E. C., Ntzani, E. E., Bika, E., Kostapanos, M. S. & Elisaf, M. S. Association between omega-3 fatty acid supplementation and risk of major cardiovascular disease events: a systematic review and meta-analysis. *Jama* **308**, 1024-1033, doi:10.1001/2012.jama.11374 (2012).
19 Yokoyama, M. *et al.* Effects of eicosapentaenoic acid on major coronary events in hypercholesterolaemic patients (JELIS): a randomised open-label, blinded endpoint analysis. *Lancet (London, England)* **369**, 1090-1098, doi:10.1016/s0140-6736(07)60527-3 (2007).
20 Gofman, J. W., Young, W. & Tandy, R. Ischemic heart disease, atherosclerosis, and longevity. *Circulation* **34**, 679-697, doi:10.1161/01.cir.34.4.679 (1966).
21 Gordon, T., Castelli, W. P., Hjortland, M. C., Kannel, W. B. & Dawber, T. R. High density lipoprotein as a protective factor against coronary heart disease. The Framingham Study. *The American journal of medicine* **62**, 707-714, doi:10.1016/0002-9343(77)90874-9 (1977).
22 Castelli, W. P. *et al.* Incidence of coronary heart disease and lipoprotein cholesterol levels. The Framingham Study. *Jama* **256**, 2835-2838 (1986).
23 Di Angelantonio, E. *et al.* Major lipids, apolipoproteins, and risk of vascular disease. *Jama* **302**, 1993-2000, doi:10.1001/jama.2009.1619 (2009).

24 Barter, P. J. *et al.* Effects of torcetrapib in patients at high risk for coronary events. *The New England journal of medicine* **357**, 2109-2122, doi:10.1056/NEJMoa0706628 (2007).
25 Williams, S. A. *et al.* Improving Assessment of Drug Safety Through Proteomics: Early Detection and Mechanistic Characterization of the Unforeseen Harmful Effects of Torcetrapib. *Circulation* **137**, 999-1010, doi:10.1161/circulationaha.117.028213 (2018).
26 Schwartz, G. G. *et al.* Effects of dalcetrapib in patients with a recent acute coronary syndrome. *The New England journal of medicine* **367**, 2089-2099, doi:10.1056/NEJMoa1206797 (2012).
27 Lincoff, A. M. *et al.* Evacetrapib and Cardiovascular Outcomes in High-Risk Vascular Disease. *The New England journal of medicine* **376**, 1933-1942, doi:10.1056/NEJMoa1609581 (2017).
28 Bowman, L. *et al.* Effects of Anacetrapib in Patients with Atherosclerotic Vascular Disease. *The New England journal of medicine* **377**, 1217-1227, doi:10.1056/NEJMoa1706444 (2017).
29 Tall, A. R. & Rader, D. J. Trials and Tribulations of CETP Inhibitors. *Circulation research* **122**, 106-112, doi:10.1161/circresaha.117.311978 (2018).
30 Pownall, H. J., Rosales, C., Gillard, B. K. & Gotto, A. M., Jr. High-density lipoproteins, reverse cholesterol transport and atherogenesis. *Nature reviews. Cardiology* **18**, 712-723, doi:10.1038/s41569-021-00538-z (2021).
31 Casula, M., Colpani, O., Xie, S., Catapano, A. L. & Baragetti, A. HDL in Atherosclerotic Cardiovascular Disease: In Search of a Role. *Cells* **10**, doi:10.3390/cells10081869 (2021).
32 Akbaraly, T. *et al.* Association of circulating metabolites with healthy diet and risk of cardiovascular disease: analysis of two cohort studies. *Scientific reports* **8(1)**, 8620 (2018).
33 Ference, B. A. *et al.* Association of Genetic Variants Related to CETP Inhibitors and Statins With Lipoprotein Levels and Cardiovascular Risk. *Jama* **318**, 947-956, doi:10.1001/jama.2017.11467 (2017).
34 Marston, N. A. *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 cardiology* **7**, 250-256, doi:10.1001/jamacardio.2021.5083 (2022).
35 Sniderman, A. D., Navar, A. M. & Thanassoulis, G. Apolipoprotein B vs Low-Density Lipoprotein Cholesterol and Non–High-Density Lipoprotein Cholesterol as the Primary Measure of Apolipoprotein B Lipoprotein-Related Risk: The Debate Is Over. *JAMA cardiology* **7**, 257-258,

doi:10.1001/jamacardio.2021.5080 (2022).

36 De Los Ríos-Ibarra, M. O. *et al.* Risk stratification and lipid evaluation in mexican patients, evidence of lipid and cardiovascular analysis in REMECAR. The mexican registry of cardiovascular diseases (REMECAR group). *Atheroscler Plus* **50**, 32-39, doi:10.1016/j.athplu.2022.08.002 (2022).

37 Aguilar-Ramirez, D. *et al.* Changes in the Diagnosis and Management of Diabetes in Mexico City Between 1998–2004 and 2015–2019. *Diabetes care* **44**, 944-951, doi:10.2337/dc20-2276 (2021).

38 Fajardo Dolci, G. E., Anda, F. J. V., Ortiz-Vázquez, P. & Olaiz-Fernández, G. The burden of cardiovascular disease in Mexico 1990-2021. Summary of the Global Burden of Disease 2021 study. *Gac Med Mex* **159**, 557-564, doi:10.24875/gmm.M24000837 (2023).

39 Baigent, C. *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 (London, England)* **376**, 1670-1681, doi:10.1016/s0140-6736(10)61350-5 (2010).

40 Collins, R. *et al.* Interpretation of the evidence for the efficacy and safety of statin therapy. *Lancet (London, England)* **388**, 2532-2561, doi:10.1016/s0140-6736(16)31357-5 (2016).

41 Mihaylova, B. *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 (London, England)* **380**, 581-590, doi:10.1016/s0140-6736(12)60367-5 (2012).

42 Thanassoulis, G. *et al.* Relations of change in plasma levels of LDL-C, non-HDL-C and apoB with risk reduction from statin therapy: a meta-analysis of randomized trials. *Journal of the American Heart Association* **3**, e000759, doi:10.1161/jaha.113.000759 (2014).

43 Langlois, M. R. *et al.* Quantifying atherogenic lipoproteins for lipid-lowering strategies: consensus-based recommendations from EAS and EFLM. *Clin Chem Lab Med* **58**, 496-517, doi:10.1515/cclm-2019-1253 (2020).

44 Welsh, P. & Sattar, N. To ApoB or Not to ApoB: New Arguments, but Basis for Widespread Implementation Remains Elusive. *Clinical chemistry* **69**, 3-5, doi:10.1093/clinchem/hvac183 (2022).

45 Mach, F. *et al.* 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk. *European heart journal* **41**, 111-188, doi:10.1093/eurheartj/ehz455 (2020).

46 Sniderman, A. D. *et al.* Discordance among apoB, non–high-density lipoprotein cholesterol, and triglycerides: implications for cardiovascular prevention. *European heart journal* **45**, 2410-2418, doi:10.1093/eurheartj/ehae258 (2024).

47 Goldstein, J. L. & Brown, M. S. A century of cholesterol and coronaries: from plaques to genes to statins. *Cell* **161**, 161-172, doi:10.1016/j.cell.2015.01.036 (2015).

48 Pandya, A., Sy, S., Cho, S., Weinstein, M. C. & Gaziano, T. A. Cost-effectiveness of 10-Year Risk Thresholds for Initiation of Statin Therapy for Primary Prevention of Cardiovascular Disease. *Jama* **314**, 142-150, doi:10.1001/jama.2015.6822 (2015).

49 Kohli-Lynch, C. N. *et al.* Cost-effectiveness of Low-density Lipoprotein Cholesterol Level–Guided Statin Treatment in Patients With Borderline Cardiovascular Risk. *JAMA cardiology* **4**, 969-977, doi:10.1001/jamacardio.2019.2851 (2019).

50 Mihaylova, B. *et al.* Lifetime effects and cost-effectiveness of statin therapy for older people in the United Kingdom: a modelling study. *Heart*, doi:10.1136/heartjnl-2024-324052 (2024).

Supplementary appendix

List of Supplementary Tables and Figures

	Description
	Table 6.S1. ICD-10 code for ASCVD and number of death
	Table 7.S1 Baseline characteristics of participants without prior chronic disease by GRS groups of LDL-C
	Table 7.S2 Baseline characteristics of participants without prior chronic disease by GRS groups of HDL-C
	Table 7.S3 Baseline characteristics of participants without prior chronic disease by GRS groups of TG
	Table 7.S4 Baseline characteristics of participants without prior chronic disease by GRS groups of ApoB
	Table 7.S5 Baseline characteristics of participants without prior chronic disease by GRS groups of ApoA1
	Figure 6.S1. The association between clinical LDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period
	Figure 6.S2. The association between remnant-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period
	Figure 6.S3. The association between non-HDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period
	Figure 6.S4. The association between non-HDL-C/HDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period
	Figure 6.S5. The association between ApoB/ApoA1 and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period
	Figure 6.S6. Subgroup analyses for the association between ApoB given LDL-C in participants without prior chronic disease at baseline during full follow-up period
	Figure 6.S7. The association between ApoB and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period in participants without diabetes at baseline
	Figure 6.S8. The association between LDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period in participants without diabetes at baseline
	Figure 6.S9. The association between TG and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period in participants without diabetes at baseline
	Figure 6.S10. The association between ApoA1 and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period in participants without diabetes at baseline
	Figure 6.S11. The association between HDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period in participants without diabetes at baseline
	Figure 6.S12. The association between clinical LDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period in participants without diabetes at baseline

Figure 6.S13. The association between remnant-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period in participants without diabetes at baseline

Figure 6.S14. The association between non-HDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period in participants without diabetes at baseline

Figure 6.S15. The association between non-HDL-C/HDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period in participants without diabetes at baseline

Figure 6.S16. The association between ApoB/ApoA1 and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period in participants without diabetes at baseline

Figure 6.S17. Extent of regression to the mean in 10 major composite lipid-related traits between baseline and resurvey, for each fifth of the baseline distribution

Figure 7.S1 Mean baseline NMR-measured LDL-C by GRS for LDL-C using SNPs from GLGC 2021 in male and female

Figure 7.S2 Mean baseline NMR-measured LDL-C by GRS for LDL-C using SNPs from GLGC 2021 by age groups

Figure 7.S3 Mean baseline NMR-measured LDL-C by GRS for LDL-C using SNPs from GLGC 2021 by ancestry groups

Figure 7.S4 Mean baseline NMR-measured LDL-C by GRS for LDL-C using SNPs from GLGC 2021 by study region

Figure 7.S5 Mean baseline NMR-measured HDL-C by GRS for HDL-C using SNPs from GLGC 2021 in male and female

Figure 7.S6 Mean baseline NMR-measured HDL-C by GRS for HDL-C using SNPs from GLGC 2021 by age groups

Figure 7.S7 Mean baseline NMR-measured HDL-C by GRS for HDL-C using SNPs from GLGC 2021 by ancestry groups

Figure 7.S8 Mean baseline NMR-measured HDL-C by GRS for HDL-C using SNPs from GLGC 2021 by study region

Figure 7.S9 Mean baseline NMR-measured TG by GRS for TG using SNPs from GLGC 2021 in male and female

Figure 7.S10 Mean baseline NMR-measured TG by GRS for TG using SNPs from GLGC 2021 by age groups

Figure 7.S11 Mean baseline NMR-measured TG by GRS for TG using SNPs from GLGC 2021 by ancestry groups

Figure 7.S12 Mean baseline NMR-measured TG by GRS for TG using SNPs from GLGC 2021 by study region

Figure 7.S13 Mean baseline NMR-measured ApoB by GRS for ApoB using SNPs from UKB 2020 in male and female

Figure 7.S14 Mean baseline NMR-measured ApoB by GRS for ApoB using SNPs from UKB 2020 by age groups

Figure 7.S15 Mean baseline NMR-measured ApoB by GRS for ApoB using SNPs from UKB 2020 by ancestry groups

Figure 7.S16 Mean baseline NMR-measured ApoB by GRS for ApoB using SNPs from UKB 2020 by study region

Figure 7.S17 Mean baseline NMR-measured ApoA1 by GRS for ApoA1 using SNPs from UKB 2020 in male and female

Figure 7.S18 Mean baseline NMR-measured ApoA1 by GRS for ApoA1 using SNPs from UKB 2020 by age groups

Figure 7.S19 Mean baseline NMR-measured ApoA1 by GRS for ApoA1 using SNPs from UKB 2020 by ancestry groups

Figure 7.S20 Mean baseline NMR-measured ApoA1 by GRS for ApoA1 using SNPs from UKB 2020 by study region

Figure 7.S21 One-sample univariable and multivariable MR of major lipids and apolipoproteins and risk of ASCVD by age groups

Figure 7.S22 One-sample univariable and multivariable MR of major lipids and apolipoproteins and risk of ASCVD by regions

Figure 7.S23 One-sample univariable and multivariable MR of major lipids and apolipoproteins and risk of ASCVD by ancestry groups

Figure 7.S24 One-sample univariable and multivariable MR of major lipids and apolipoproteins and risk of ASCVD by diabetes

Table 6.S1. ICD-10 code for ASCVD and number of death

ASCVD	ICD-10 codes
Ischaemic heart disease	I20-I25
Ischaemic stroke	I63, I693
Peripheral arterial death	I70-I74, I77, I79, E105, E115, E125, E135, E145

Outcome*	Men	Women	Overall
ASCVD mortality	1303	1410	2713
Ischaemic heart disease	1192	1287	2479
Ischaemic stroke	32	81	113
Peripheral arterial death	79	42	121
All-cause mortality	6459	8264	14723

*Death at age of 35-84.

Table 7.S1 Baseline characteristics of participants without prior chronic disease by GRS groups of LDL-C

	Genetic risk scores for LDL-C (participants)										Overall (n=107 504)
	Group 1 (n=9 735)	Group 2 (n=9 736)	Group 3 (n=9 738)	Group 4 (n=9 741)	Group 5 (n=9 742)	Group 6 (n=9 738)	Group 7 (n=9 736)	Group 8 (n=9 732)	Group 9 (n=9 740)	Group 10 (n=9 741)	
Age, years	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)
Male	3198 (33%)	3231 (33%)	3234 (33%)	3180 (33%)	3185 (33%)	3140 (32%)	3243 (33%)	3216 (33%)	3163 (32%)	3187 (33%)	35 518 (33%)
Blood pressure											
Systolic blood pressure, mmHg	128 (17)	127 (16)	127 (17)	128 (17)	127 (17)	127 (16)	127 (16)	127 (16)	127 (16)	127 (16)	127 (17)
Diastolic blood pressure, mmHg	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)
Taking blood pressure lowering medication	1467 (15%)	1426 (15%)	1433 (15%)	1417 (15%)	1439 (15%)	1393 (14%)	1371 (14%)	1399 (14%)	1364 (14%)	1353 (14%)	15 441 (14%)
Socio-economic status and lifestyle behaviours											
Resident of Coyoacan	4464 (46%)	4220 (43%)	4231 (43%)	4062 (42%)	4073 (42%)	4018 (41%)	4033 (41%)	4029 (41%)	4110 (42%)	4317 (44%)	47 683 (44%)
University/college educated	1821 (19%)	1604 (16%)	1584 (16%)	1497 (15%)	1466 (15%)	1441 (15%)	1527 (16%)	1472 (15%)	1466 (15%)	1485 (15%)	16 897 (16%)
Current smoker	3303 (34%)	3220 (33%)	3141 (32%)	3101 (32%)	3096 (32%)	3004 (31%)	3108 (32%)	3011 (31%)	3117 (32%)	3110 (32%)	34 552 (32%)
Current drinker	15 221 (156%)	15 295 (157%)	15 287 (157%)	15 091 (155%)	15 057 (155%)	15 145 (156%)	15 075 (155%)	15 072 (155%)	14 999 (154%)	15 183 (156%)	167417 (156%)
Any regular leisure-time physical activity	3893 (40%)	3754 (39%)	3691 (38%)	3665 (38%)	3570 (37%)	3384 (35%)	3608 (37%)	3478 (36%)	3558 (37%)	3494 (36%)	39 703 (37%)
Physical measurements											
BMI, kg/m^2	29.0 (5.0)	29.0 (4.9)	29.0 (4.9)	29.2 (5.0)	29.0 (4.9)	29.0 (4.9)	28.9 (4.8)	28.9 (4.9)	29.0 (4.8)	29.0 (4.9)	29.0 (4.9)
Waist circumference, cm	94 (12)	94 (11)	94 (12)	94 (12)	94 (11)	94 (11)	94 (11)	94 (11)	94 (11)	94 (11)	94 (11)
Hip circumference, cm	105 (11)	105 (11)	105 (11)	105 (11)	105 (10)	104 (10)	104 (10)	104 (10)	104 (10)	104 (11)	105 (10)
Diabetes and glycosylated hemoglobin											
Diabetes status											
No diabetes	8060 (83%)	8035 (83%)	8025 (82%)	7973 (82%)	7909 (81%)	7998 (82%)	7983 (82%)	7964 (82%)	7967 (82%)	8009 (82%)	88 314 (82%)
Undiagnosed diabetes	429 (4%)	457 (5%)	446 (5%)	493 (5%)	511 (5%)	499 (5%)	467 (5%)	472 (5%)	499 (5%)	478 (5%)	5227 (5%)
Diagnosed diabetes with HbA1c <9%	447 (5%)	492 (5%)	455 (5%)	513 (5%)	518 (5%)	452 (5%)	485 (5%)	525 (5%)	521 (5%)	478 (5%)	5375 (5%)
Diagnosed diabetes with HbA1c ≥11%	799 (8%)	752 (8%)	812 (8%)	762 (8%)	804 (8%)	789 (8%)	801 (8%)	771 (8%)	753 (8%)	776 (8%)	8588 (8%)

Mean (SD) or n (column %) shown. BMI=Body mass index. Table excludes those with prior chronic disease at recruitment (ischemic heart disease, stroke, chronic kidney disease, cirrhosis, cancer, emphysema), missing or out-of-range data on any analysis covariate (sex, district of residence, educational level attained, anthropometry, smoking status, alcohol intake, leisure time physical activity, diabetes status), uncertain follow-up, or missing or out-of-range blood pressure. The difference (mean or percentage) between group 1 and 10 was low for all characteristics.

Table 7.S2 Baseline characteristics of participants without prior chronic disease by GRS groups of HDL-C

	Genetic risk scores for HDL-C (participants)										Overall (n=107 504)
	Group 1 (n=9 735)	Group 2 (n=9 736)	Group 3 (n=9 738)	Group 4 (n=9 741)	Group 5 (n=9 742)	Group 6 (n=9 738)	Group 7 (n=9 736)	Group 8 (n=9 732)	Group 9 (n=9 740)	Group 10 (n=9 741)	
Age, years	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)
Male	3200 (33%)	3154 (32%)	3249 (33%)	3189 (33%)	3214 (33%)	3241 (33%)	3182 (33%)	3180 (33%)	3206 (33%)	3162 (32%)	35 518 (33%)
Blood pressure											
Systolic blood pressure, mmHg	127 (16)	128 (17)	127 (16)	127 (17)	127 (17)	127 (16)	127 (17)	127 (17)	127 (16)	127 (16)	127 (17)
Diastolic blood pressure, mmHg	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)
Taking blood pressure lowering medication	1451 (15%)	1435 (15%)	1402 (14%)	1362 (14%)	1405 (14%)	1417 (15%)	1428 (15%)	1416 (15%)	1398 (14%)	1348 (14%)	15 441 (14%)
Socio-economic status and lifestyle behaviours											
Resident of Coyoacan	4136 (42%)	4222 (43%)	4176 (43%)	4148 (43%)	4044 (42%)	4172 (43%)	4111 (42%)	4130 (42%)	4220 (43%)	4198 (43%)	47 683 (44%)
University/college educated	1447 (15%)	1516 (16%)	1576 (16%)	1487 (15%)	1496 (15%)	1551 (16%)	1563 (16%)	1577 (16%)	1565 (16%)	1585 (16%)	16 897 (16%)
Current smoker	3124 (32%)	3096 (32%)	3048 (31%)	3062 (31%)	3119 (32%)	3141 (32%)	3217 (33%)	3079 (32%)	3124 (32%)	3201 (33%)	34 552 (32%)
Current drinker	15 102 (155%)	15 101 (155%)	15 018 (154%)	15 093 (155%)	15 216 (156%)	15 169 (156%)	15 173 (156%)	15 102 (155%)	15 262 (157%)	15 189 (156%)	167 417 (156%)
Any regular leisure-time physical activity	3560 (37%)	3622 (37%)	3644 (37%)	3508 (36%)	3548 (36%)	3718 (38%)	3528 (36%)	3562 (37%)	3716 (38%)	3689 (38%)	39 703 (37%)
Physical measurements											
BMI, kg/m^2	28.9 (4.9)	28.9 (4.9)	29.0 (4.9)	29.0 (4.9)	28.9 (4.9)	28.9 (4.9)	29.0 (4.9)	29.0 (4.9)	29.1 (4.9)	29.0 (5.0)	29.0 (4.9)
Waist circumference, cm	94 (11)	94 (11)	94 (11)	94 (12)	94 (11)	94 (11)	94 (12)	94 (12)	95 (11)	94 (12)	94 (11)
Hip circumference, cm	104 (10)	104 (10)	105 (11)	105 (11)	105 (10)	105 (10)	105 (10)	105 (11)	105 (10)	105 (11)	105 (10)
Diabetes and glycosylated hemoglobin											
Diabetes status											
No diabetes	7744 (80%)	7902 (81%)	7895 (81%)	7930 (81%)	7977 (82%)	8035 (83%)	7990 (82%)	8085 (83%)	8134 (84%)	8231 (85%)	88 314 (82%)
Undiagnosed diabetes	555 (6%)	496 (5%)	493 (5%)	471 (5%)	434 (4%)	471 (5%)	499 (5%)	467 (5%)	435 (4%)	430 (4%)	5227 (5%)
Diagnosed diabetes with HbA1c <9%	553 (6%)	509 (5%)	502 (5%)	502 (5%)	510 (5%)	455 (5%)	478 (5%)	472 (5%)	476 (5%)	429 (4%)	5375 (5%)
Diagnosed diabetes with HbA1c ≥11%	884 (9%)	835 (9%)	846 (9%)	834 (9%)	823 (8%)	773 (8%)	766 (8%)	720 (7%)	691 (7%)	647 (7%)	8588 (8%)

Mean (SD) or n (column %) shown. BMI=Body mass index. Table excludes those with prior chronic disease at recruitment (ischemic heart disease, stroke, chronic kidney disease, cirrhosis, cancer, emphysema), missing or out-of-range data on any analysis covariate (sex, district of residence, educational level attained, anthropometry, smoking status, alcohol intake, leisure time physical activity, diabetes status), uncertain follow-up, or missing or out-of-range blood pressure. The difference (mean or percentage) between group 1 and 10 was low for all characteristics.

Table 7.S3 Baseline characteristics of participants without prior chronic disease by GRS groups of TG

	Genetic risk scores for TG (participants)										Overall (n=107 504)
	Group 1 (n=9 735)	Group 2 (n=9 736)	Group 3 (n=9 738)	Group 4 (n=9 741)	Group 5 (n=9 742)	Group 6 (n=9 738)	Group 7 (n=9 736)	Group 8 (n=9 732)	Group 9 (n=9 740)	Group 10 (n=9 741)	
Age, years	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)
Male	3146 (32%)	3331 (34%)	3082 (32%)	3153 (32%)	3218 (33%)	3247 (33%)	3155 (32%)	3162 (32%)	3288 (34%)	3206 (33%)	35 518 (33%)
Blood pressure											
Systolic blood pressure, mmHg	127 (16)	127 (17)	127 (16)	127 (17)	127 (17)	127 (16)	127 (17)	127 (16)	127 (17)	127 (17)	127 (17)
Diastolic blood pressure, mmHg	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)
Taking blood pressure lowering medication	1463 (15%)	1423 (15%)	1418 (15%)	1432 (15%)	1462 (15%)	1448 (15%)	1402 (14%)	1354 (14%)	1341 (14%)	1325 (14%)	15 441 (14%)
Socio-economic status and lifestyle behaviours											
Resident of Coyoacan	4433 (46%)	4377 (45%)	4284 (44%)	4210 (43%)	4101 (42%)	4049 (42%)	4124 (42%)	4114 (42%)	4033 (41%)	3863 (40%)	47 683 (44%)
University/college educated	1830 (19%)	1706 (17%)	1589 (16%)	1624 (17%)	1517 (16%)	1477 (15%)	1459 (15%)	1441 (15%)	1395 (14%)	1337 (14%)	16 897 (16%)
Current smoker	3395 (35%)	3390 (35%)	3200 (33%)	3148 (32%)	3088 (32%)	3125 (32%)	3036 (31%)	2998 (31%)	2971 (31%)	2877 (30%)	34 552 (32%)
Current drinker	15 228 (156%)	15 371 (158%)	15 231 (157%)	15 149 (156%)	15 263 (156%)	14 967 (154%)	15 176 (156%)	15 106 (155%)	14 985 (154%)	15 009 (154%)	167 417 (156%)
Any regular leisure-time physical activity	3907 (40%)	3792 (39%)	3749 (39%)	3650 (38%)	3542 (36%)	3589 (37%)	3567 (37%)	3518 (36%)	3508 (36%)	3287 (34%)	39 703 (37%)
Physical measurements											
BMI, kg/m^2	29.0 (4.9)	29.0 (4.9)	29.0 (4.9)	29.0 (5.0)	28.9 (4.9)	29.1 (5.0)	29.0 (4.9)	29.0 (4.9)	28.9 (4.8)	29.0 (4.9)	29.0 (4.9)
Waist circumference, cm	94 (12)	94 (12)	94 (12)	94 (12)	94 (12)	95 (12)	94 (11)	94 (11)	94 (11)	94 (11)	94 (11)
Hip circumference, cm	105 (11)	105 (10)	105 (11)	105 (11)	104 (11)	105 (11)	104 (10)	104 (11)	104 (10)	104 (10)	105 (10)
Diabetes and glycosylated hemoglobin											
Diabetes status											
No diabetes	8246 (85%)	8092 (83%)	8057 (83%)	7960 (82%)	8037 (82%)	7860 (81%)	8010 (82%)	7932 (81%)	7863 (81%)	7896 (81%)	88 314 (82%)
Undiagnosed diabetes	402 (4%)	463 (5%)	454 (5%)	484 (5%)	451 (5%)	464 (5%)	495 (5%)	515 (5%)	504 (5%)	522 (5%)	5227 (5%)
Diagnosed diabetes with HbA1c <9%	415 (4%)	440 (5%)	497 (5%)	507 (5%)	529 (5%)	540 (6%)	459 (5%)	489 (5%)	516 (5%)	496 (5%)	5375 (5%)
Diagnosed diabetes with HbA1c ≥11%	672 (7%)	762 (8%)	721 (7%)	778 (8%)	761 (8%)	834 (9%)	792 (8%)	818 (8%)	851 (9%)	832 (9%)	8588 (8%)

Mean (SD) or n (column %) shown. BMI=Body mass index. Table excludes those with prior chronic disease at recruitment (ischemic heart disease, stroke, chronic kidney disease, cirrhosis, cancer, emphysema), missing or out-of-range data on any analysis covariate (sex, district of residence, educational level attained, anthropometry, smoking status, alcohol intake, leisure time physical activity, diabetes status), uncertain follow-up, or missing or out-of-range blood pressure. The difference (mean or percentage) between group 1 and 10 was low for all characteristics.

Table 7.S4 Baseline characteristics of participants without prior chronic disease by GRS groups of ApoB

	Genetic risk scores for ApoB (participants)										Overall (n=107 504)
	Group 1 (n=9 735)	Group 2 (n=9 736)	Group 3 (n=9 738)	Group 4 (n=9 741)	Group 5 (n=9 742)	Group 6 (n=9 738)	Group 7 (n=9 736)	Group 8 (n=9 732)	Group 9 (n=9 740)	Group 10 (n=9 741)	
Age, years	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)
Male	3162 (32%)	3246 (33%)	3177 (33%)	3185 (33%)	3182 (33%)	3242 (33%)	3158 (32%)	3209 (33%)	3220 (33%)	3207 (33%)	35 518 (33%)
Blood pressure											
Systolic blood pressure, mmHg	127 (16)	127 (17)	127 (17)	127 (16)	127 (17)	127 (17)	127 (16)	127 (16)	128 (17)	127 (16)	127 (17)
Diastolic blood pressure, mmHg	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)
Taking blood pressure lowering medication	1464 (15%)	1516 (16%)	1459 (15%)	1376 (14%)	1390 (14%)	1402 (14%)	1386 (14%)	1334 (14%)	1404 (14%)	1337 (14%)	15 441 (14%)
Socio-economic status and lifestyle behaviours											
Resident of Coyoacan	4446 (46%)	4326 (44%)	4198 (43%)	4163 (43%)	4188 (43%)	4084 (42%)	4083 (42%)	3900 (40%)	4045 (42%)	4155 (43%)	47 683 (44%)
University/college educated	1802 (18%)	1614 (17%)	1632 (17%)	1536 (16%)	1488 (15%)	1544 (16%)	1416 (15%)	1463 (15%)	1457 (15%)	1423 (15%)	16 897 (16%)
Current smoker	3307 (34%)	3219 (33%)	3123 (32%)	3161 (32%)	3131 (32%)	3154 (32%)	2971 (30%)	3077 (32%)	3031 (31%)	3054 (31%)	34 552 (32%)
Current drinker	15 233 (156%)	15 358 (158%)	15 203 (156%)	15 034 (154%)	15 071 (155%)	15 141 (155%)	15 206 (156%)	15 030 (154%)	15 114 (155%)	15 095 (155%)	167 417 (156%)
Any regular leisure-time physical activity	3854 (40%)	3771 (39%)	3681 (38%)	3655 (38%)	3527 (36%)	3581 (37%)	3653 (38%)	3492 (36%)	3469 (36%)	3426 (35%)	39 703 (37%)
Physical measurements											
BMI, kg/m^2	28.9 (5.0)	29.0 (4.9)	29.0 (4.9)	29.0 (4.9)	28.9 (5.0)	29.0 (4.9)	29.1 (4.9)	28.9 (4.8)	29.0 (4.8)	29.0 (4.9)	29.0 (4.9)
Waist circumference, cm	94 (12)	94 (12)	94 (11)	94 (11)	94 (12)	94 (12)	94 (11)	94 (11)	94 (11)	94 (11)	94 (11)
Hip circumference, cm	105 (11)	105 (11)	105 (10)	105 (10)	105 (11)	105 (10)	105 (11)	104 (10)	104 (10)	104 (11)	105 (10)
Diabetes and glycosylated hemoglobin											
Diabetes status											
No diabetes	8023 (82%)	8059 (83%)	7972 (82%)	7936 (81%)	8001 (82%)	7985 (82%)	7962 (82%)	7993 (82%)	8004 (82%)	8018 (82%)	88 314 (82%)
Undiagnosed diabetes	453 (5%)	437 (4%)	483 (5%)	492 (5%)	462 (5%)	500 (5%)	502 (5%)	470 (5%)	469 (5%)	486 (5%)	5227 (5%)
Diagnosed diabetes with HbA1c <9%	480 (5%)	500 (5%)	543 (6%)	489 (5%)	502 (5%)	480 (5%)	477 (5%)	467 (5%)	499 (5%)	451 (5%)	5375 (5%)
Diagnosed diabetes with HbA1c ≥11%	785 (8%)	746 (8%)	744 (8%)	824 (8%)	777 (8%)	777 (8%)	800 (8%)	812 (8%)	770 (8%)	786 (8%)	8588 (8%)

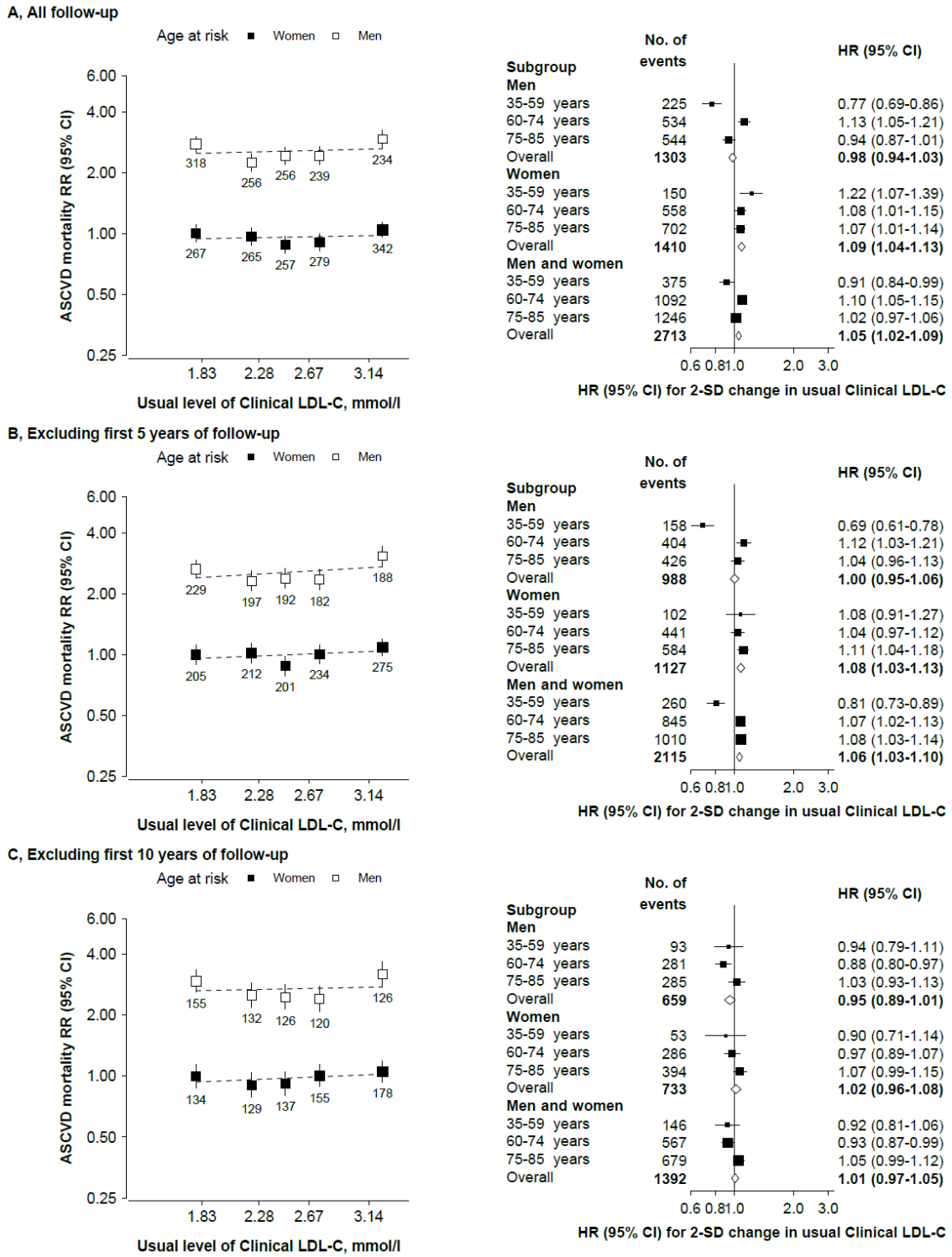
Mean (SD) or n (column %) shown. BMI=Body mass index. Table excludes those with prior chronic disease at recruitment (ischemic heart disease, stroke, chronic kidney disease, cirrhosis, cancer, emphysema), missing or out-of-range data on any analysis covariate (sex, district of residence, educational level attained, anthropometry, smoking status, alcohol intake, leisure time physical activity, diabetes status), uncertain follow-up, or missing or out-of-range blood pressure. The difference (mean or percentage) between group 1 and 10 was low for all characteristics.

Table 7.S5 Baseline characteristics of participants without prior chronic disease by GRS groups of ApoA1

	Genetic risk scores for ApoA1 (participants)										Overall (n=107 504)
	Group 1 (n=9 735)	Group 2 (n=9 736)	Group 3 (n=9 738)	Group 4 (n=9 741)	Group 5 (n=9 742)	Group 6 (n=9 738)	Group 7 (n=9 736)	Group 8 (n=9 732)	Group 9 (n=9 740)	Group 10 (n=9 741)	
Age, years	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)	52 (12)
Male	3247 (33%)	3204 (33%)	3213 (33%)	3172 (33%)	3212 (33%)	3160 (32%)	3189 (33%)	3217 (33%)	3204 (33%)	3170 (33%)	35 518 (33%)
Blood pressure											
Systolic blood pressure, mmHg	127 (16)	127 (17)	127 (17)	128 (17)	128 (17)	127 (16)	127 (17)	127 (17)	127 (16)	127 (16)	127 (17)
Diastolic blood pressure, mmHg	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)	83 (10)
Taking blood pressure lowering medication	1403 (14%)	1421 (15%)	1337 (14%)	1437 (15%)	1402 (14%)	1430 (15%)	1417 (15%)	1335 (14%)	1404 (14%)	1482 (15%)	15 441 (14%)
Socio-economic status and lifestyle behaviours											
Resident of Coyoacan	4122 (42%)	4034 (41%)	4042 (41%)	4119 (42%)	4222 (43%)	4178 (43%)	4195 (43%)	4186 (43%)	4190 (43%)	4300 (44%)	47 683 (44%)
University/college educated	1475 (15%)	1435 (15%)	1487 (15%)	1508 (15%)	1533 (16%)	1465 (15%)	1546 (16%)	1543 (16%)	1636 (17%)	1747 (18%)	16 897 (16%)
Current smoker	3090 (32%)	3055 (31%)	3133 (32%)	3083 (32%)	3149 (32%)	3026 (31%)	3142 (32%)	3141 (32%)	3178 (33%)	3231 (33%)	34 552 (32%)
Current drinker	15 155 (156%)	14 995 (154%)	15 124 (155%)	15 186 (156%)	15 073 (155%)	15 223 (156%)	15 195 (156%)	15 220 (156%)	15 210 (156%)	15 104 (155%)	167 417 (156%)
Any regular leisure-time physical activity	3629 (37%)	3592 (37%)	3399 (35%)	3429 (35%)	3666 (38%)	3619 (37%)	3566 (37%)	3624 (37%)	3825 (39%)	3760 (39%)	39 703 (37%)
Physical measurements											
BMI, kg/m^2	28.9 (4.9)	29.0 (4.9)	28.9 (4.9)	29.1 (4.9)	29.0 (4.9)	28.9 (4.9)	29.0 (4.8)	29.0 (4.9)	29.0 (4.9)	29.0 (5.0)	29.0 (4.9)
Waist circumference, cm	94 (12)	94 (11)	94 (11)	94 (11)	94 (12)	94 (11)	94 (11)	94 (12)	94 (11)	94 (12)	94 (11)
Hip circumference, cm	104 (10)	105 (11)	104 (10)	105 (11)	105 (11)	105 (10)	105 (10)	105 (11)	105 (11)	105 (11)	105 (10)
Diabetes and glycosylated hemoglobin											
Diabetes status											
No diabetes	7887 (81%)	7947 (82%)	7915 (81%)	7938 (81%)	7965 (82%)	7977 (82%)	7983 (82%)	8017 (82%)	8130 (83%)	8194 (84%)	88 314 (82%)
Undiagnosed diabetes	503 (5%)	469 (5%)	504 (5%)	511 (5%)	513 (5%)	422 (4%)	473 (5%)	470 (5%)	466 (5%)	423 (4%)	5227 (5%)
Diagnosed diabetes with HbA1c <9%	518 (5%)	480 (5%)	508 (5%)	537 (6%)	491 (5%)	524 (5%)	462 (5%)	496 (5%)	438 (4%)	434 (4%)	5375 (5%)
Diagnosed diabetes with HbA1c ≥11%	833 (9%)	846 (9%)	814 (8%)	756 (8%)	773 (8%)	819 (8%)	823 (8%)	759 (8%)	708 (7%)	690 (7%)	8588 (8%)

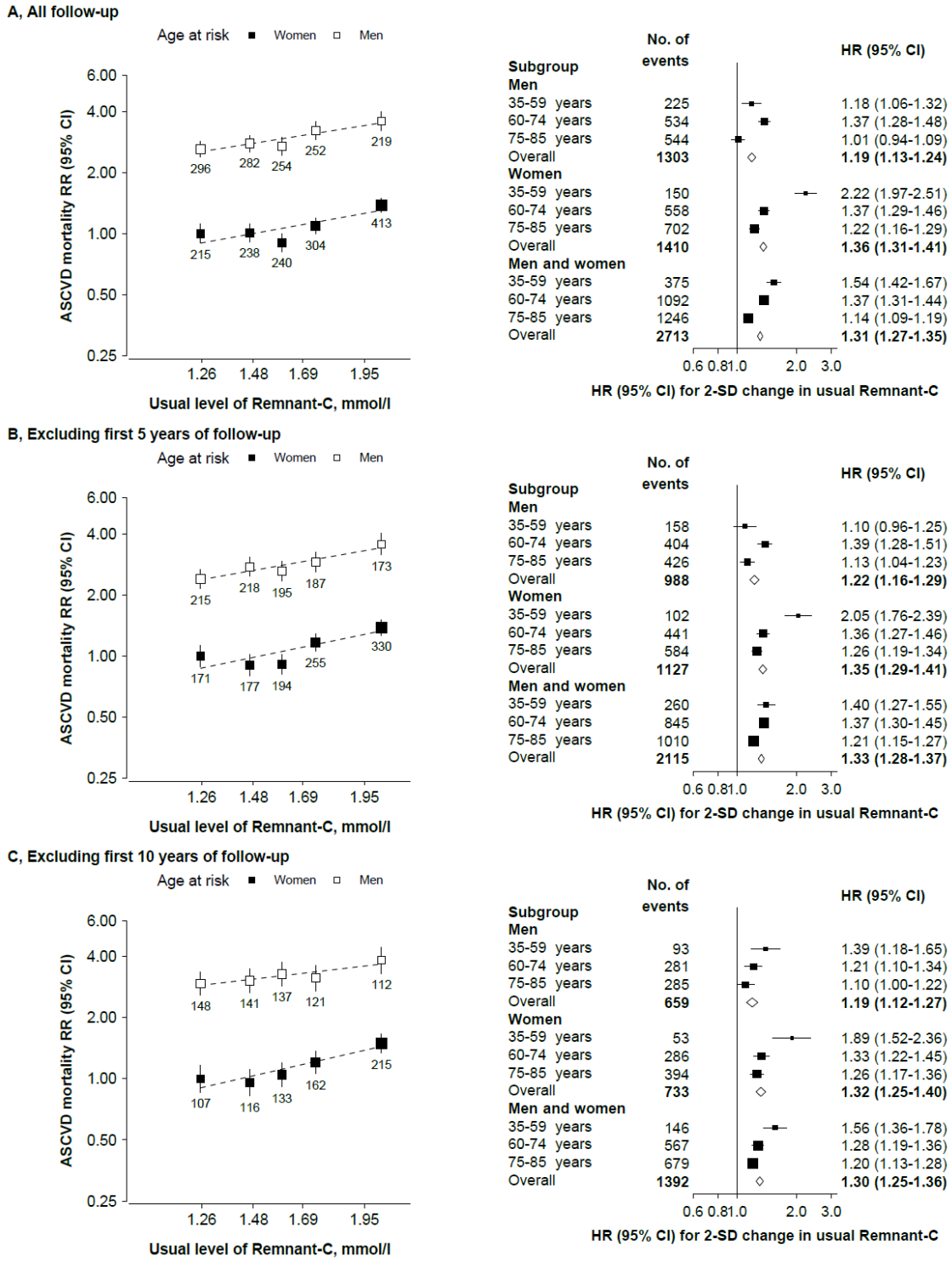
Mean (SD) or n (column %) shown. BMI=Body mass index. Table excludes those with prior chronic disease at recruitment (ischemic heart disease, stroke, chronic kidney disease, cirrhosis, cancer, emphysema), missing or out-of-range data on any analysis covariate (sex, district of residence, educational level attained, anthropometry, smoking status, alcohol intake, leisure time physical activity, diabetes status), uncertain follow-up, or missing or out-of-range blood pressure. The difference (mean or percentage) between group 1 and 10 was low for all characteristics.

Figure 6.S1. The association between clinical LDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period



Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

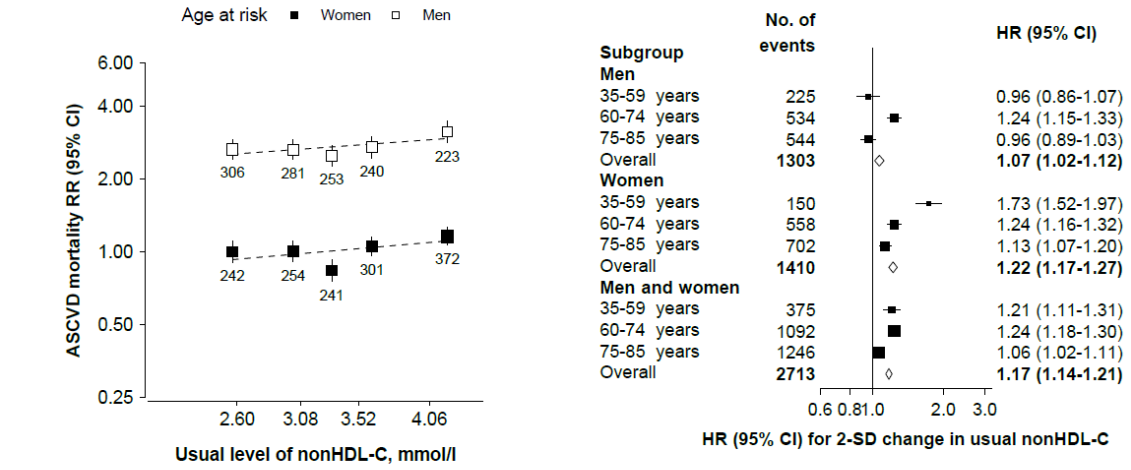
Figure 6.S2. The association between remnant-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period



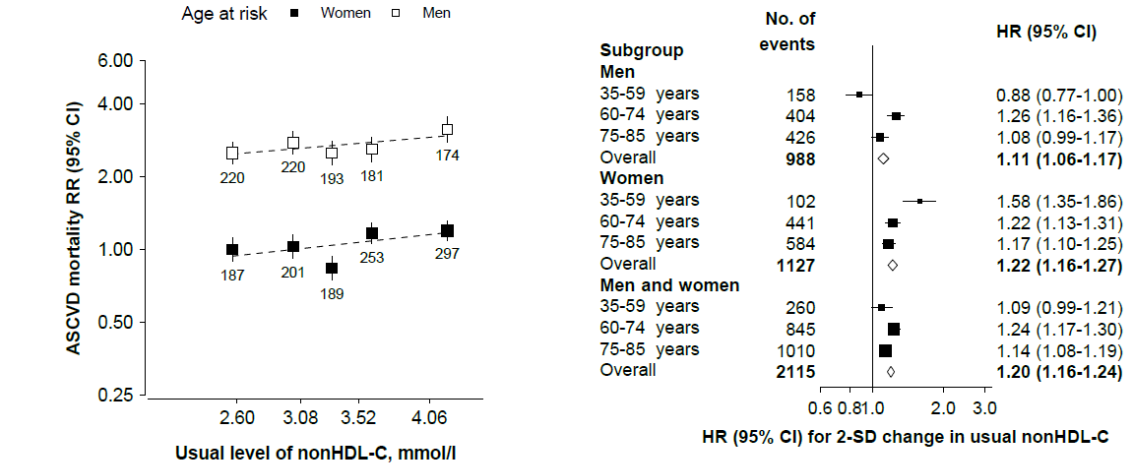
Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.S3. The association between non-HDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period

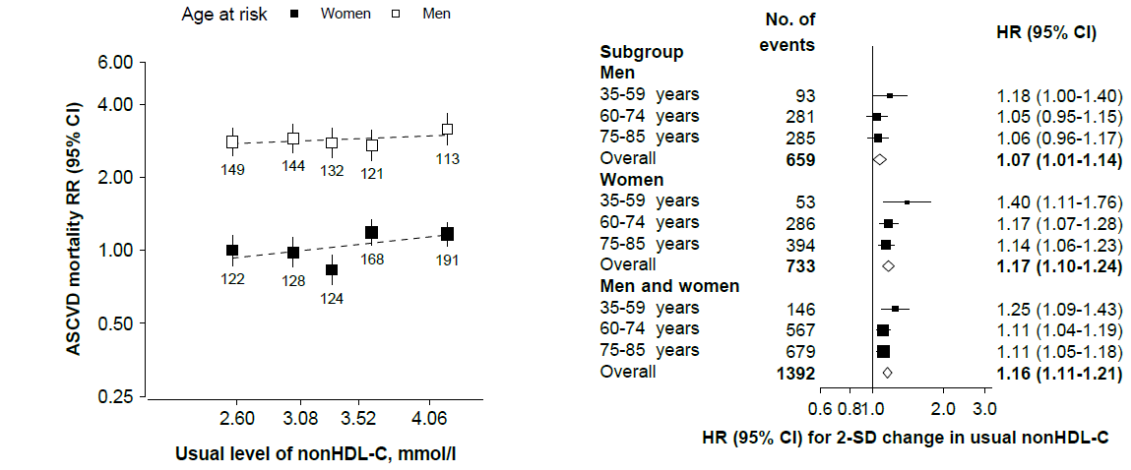
A, All follow-up



B, Excluding first 5 years of follow-up

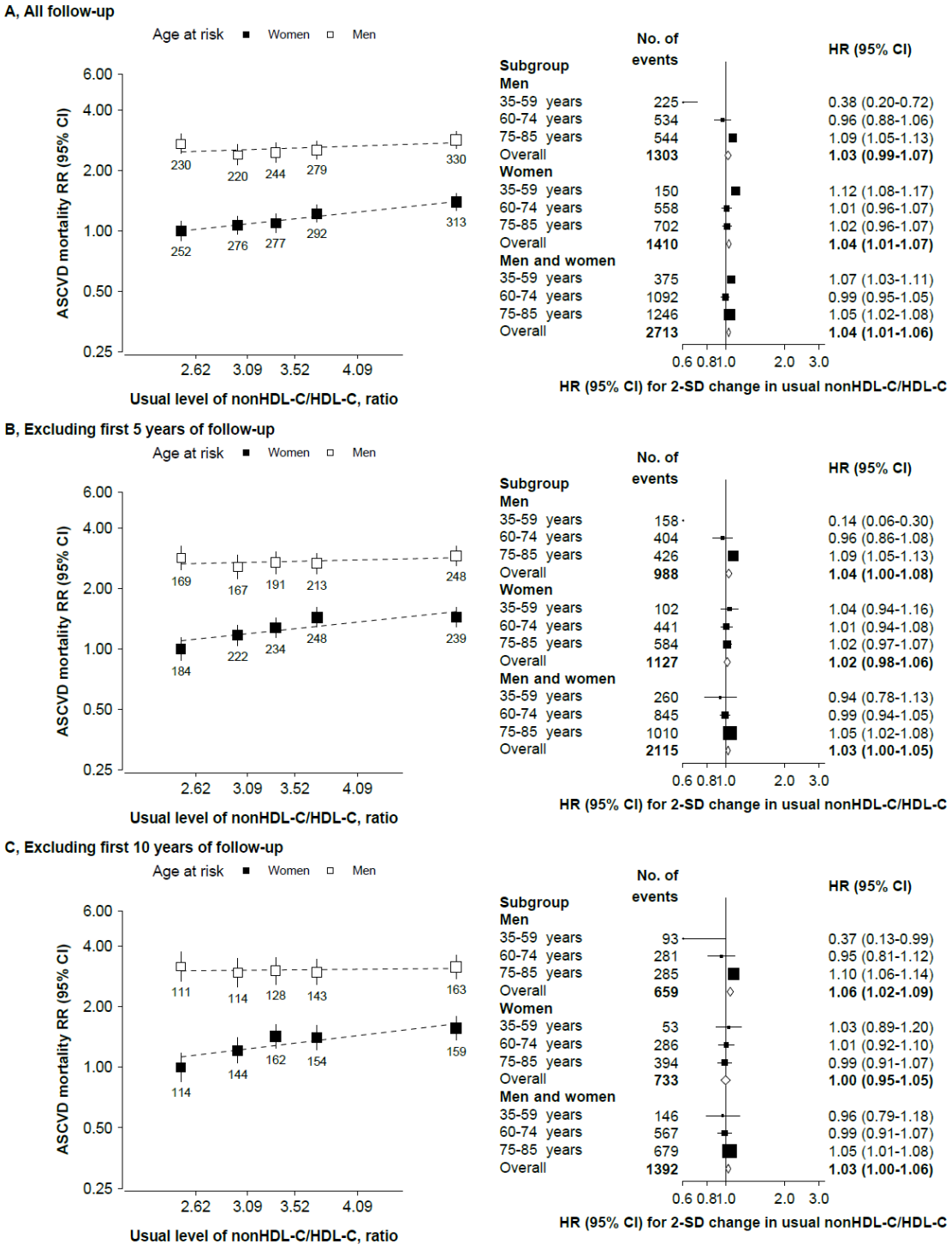


C, Excluding first 10 years of follow-up



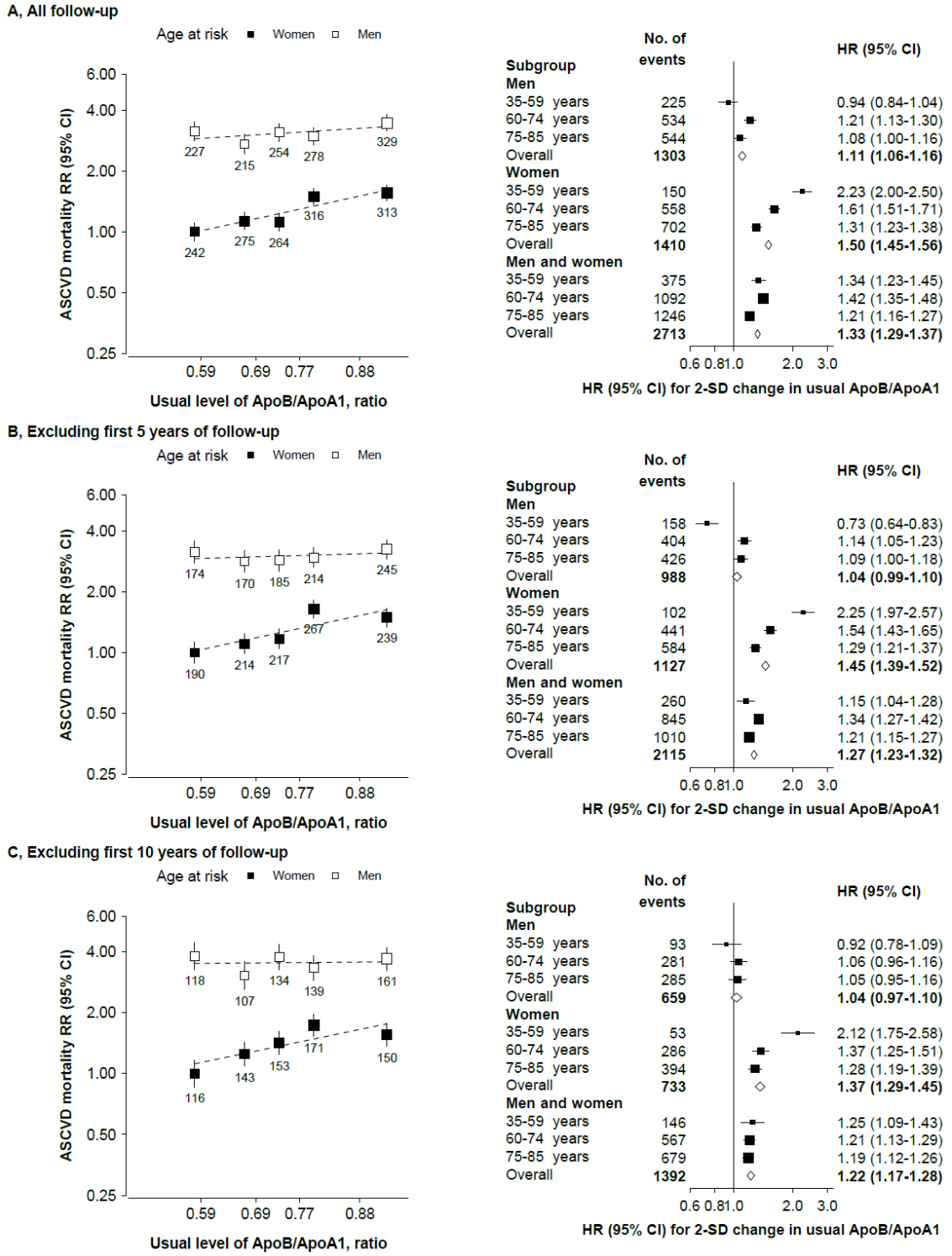
Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.S4. The association between non-HDL-C/HDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period



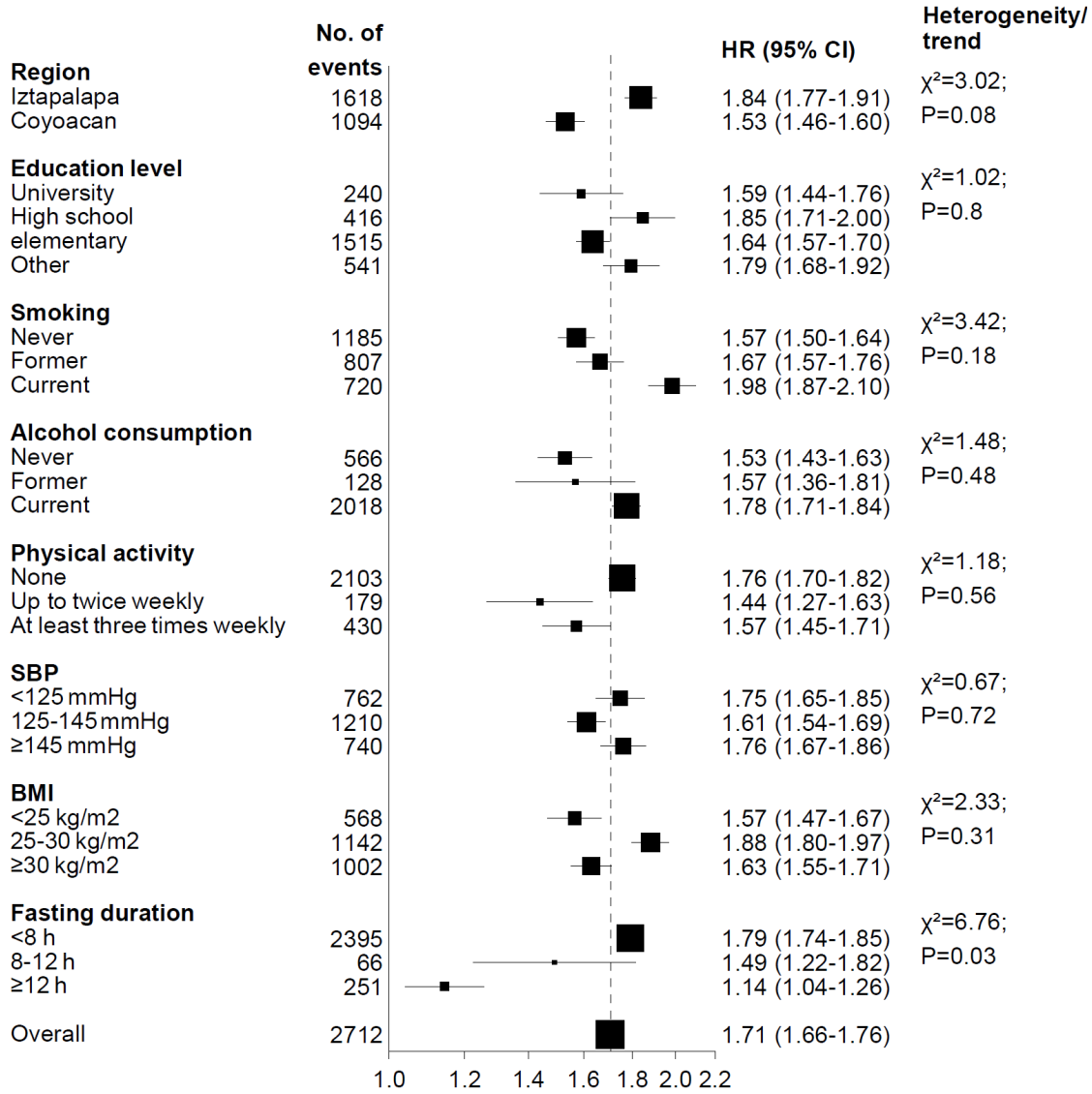
Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.S5. The association between ApoB/ApoA1 and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period



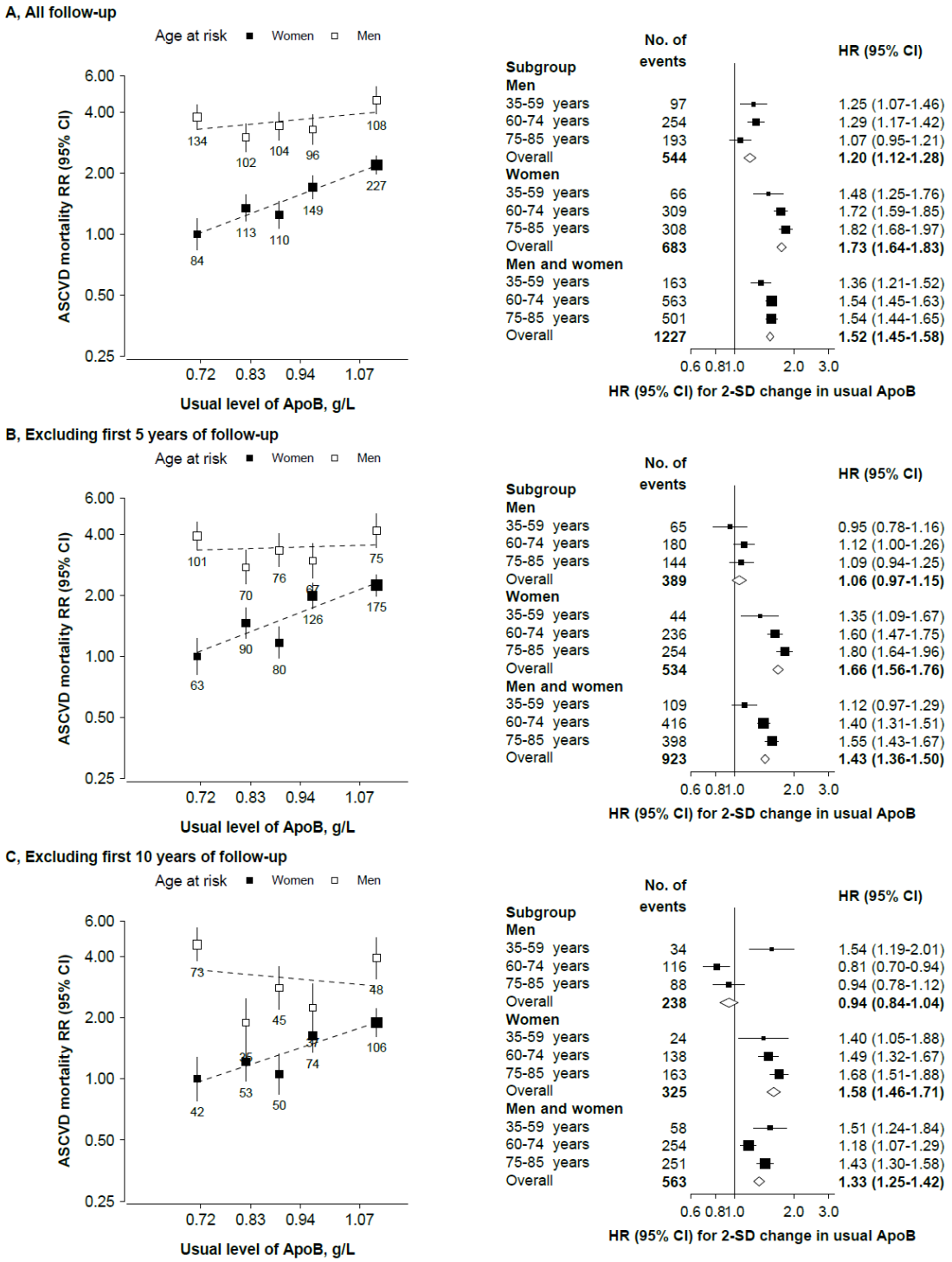
Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.S6. Subgroup analyses for the association between ApoB given LDL-C in participants without prior chronic disease at baseline during full follow-up period



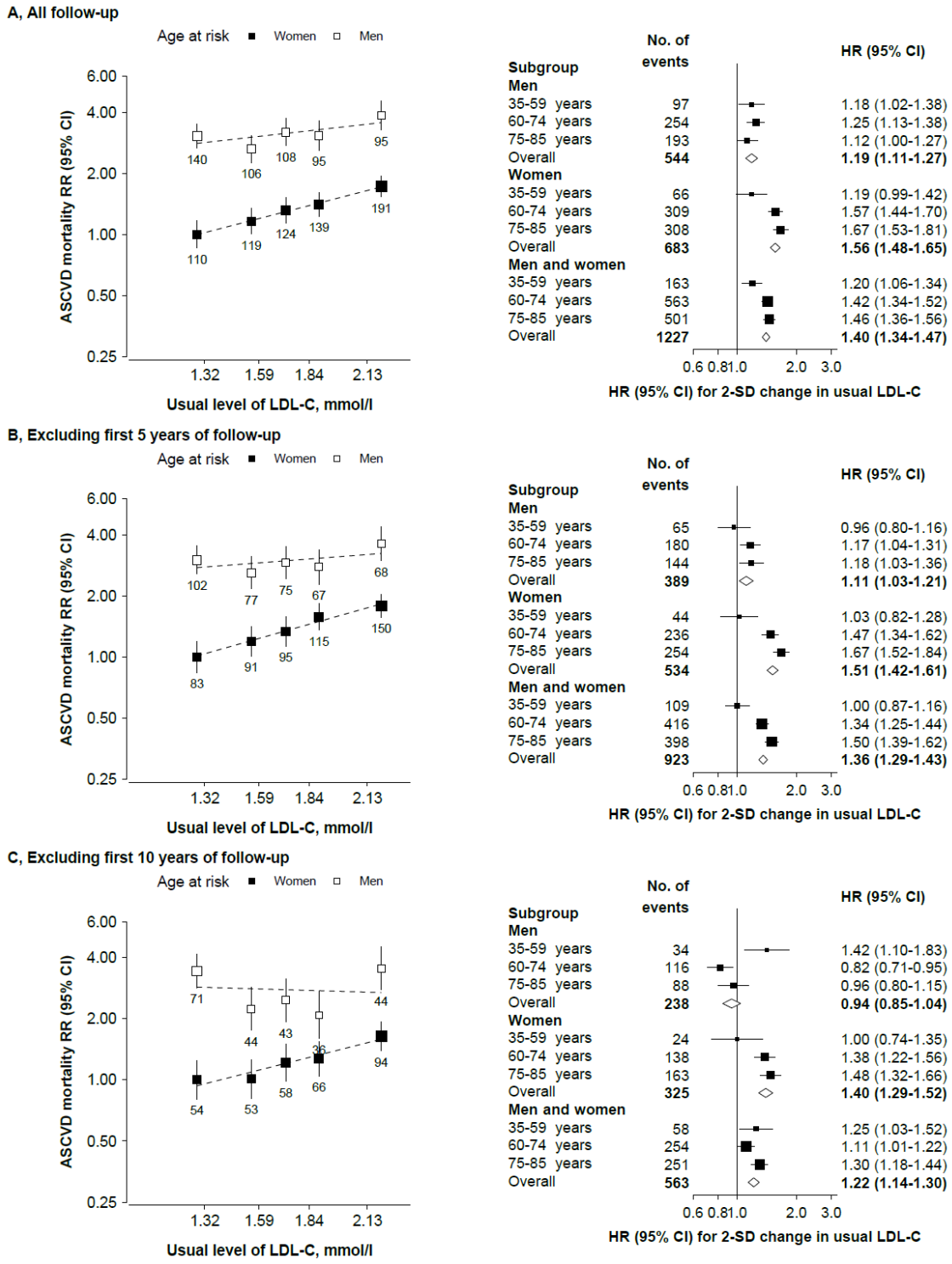
Analyses excluded those with chronic diseases at baseline (except diabetes) with mutual adjustment for age, sex, region, education level, smoking, alcohol consumption, physical activity, SBP, BMI, and fasting duration. HRs are estimated separately by levels of each subgroup shown (with mutual adjustment for other confounders). Each square was inversely proportional to the variance of the log of the hazard ratio.

Figure 6.S7. The association between ApoB and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period in participants without diabetes at baseline



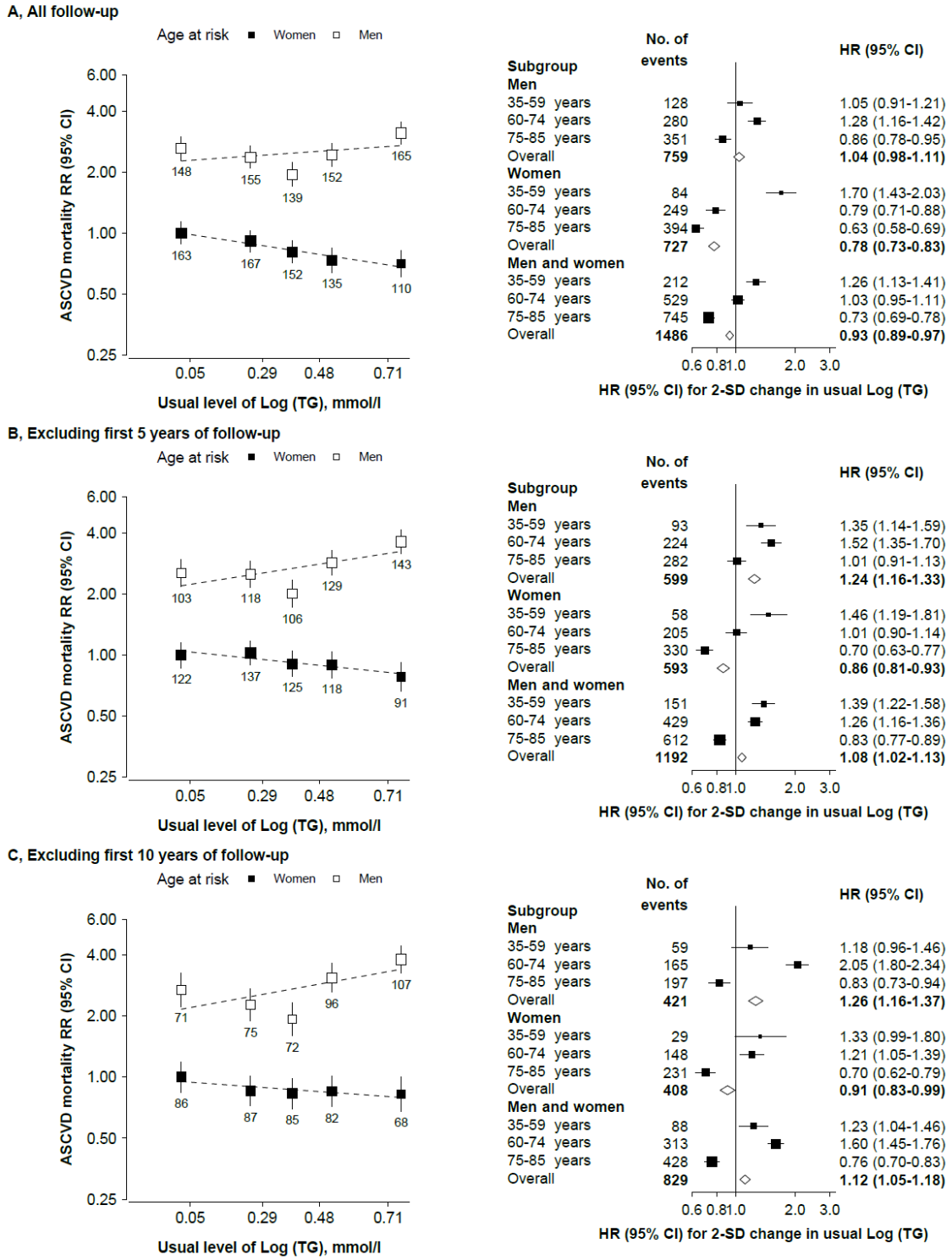
Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.S8. The association between LDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period in participants without diabetes at baseline



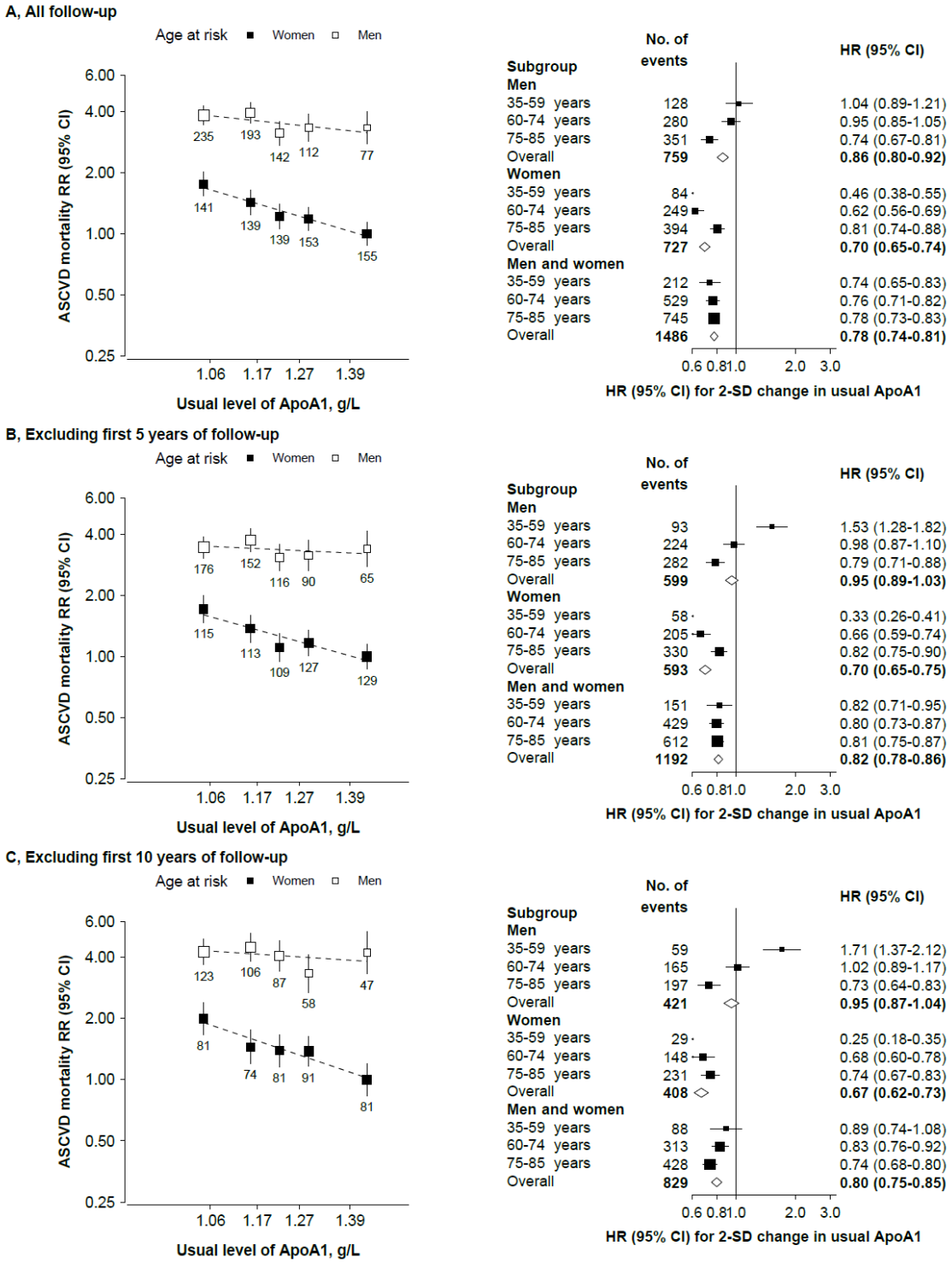
Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.S9. The association between TG and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period in participants without diabetes at baseline



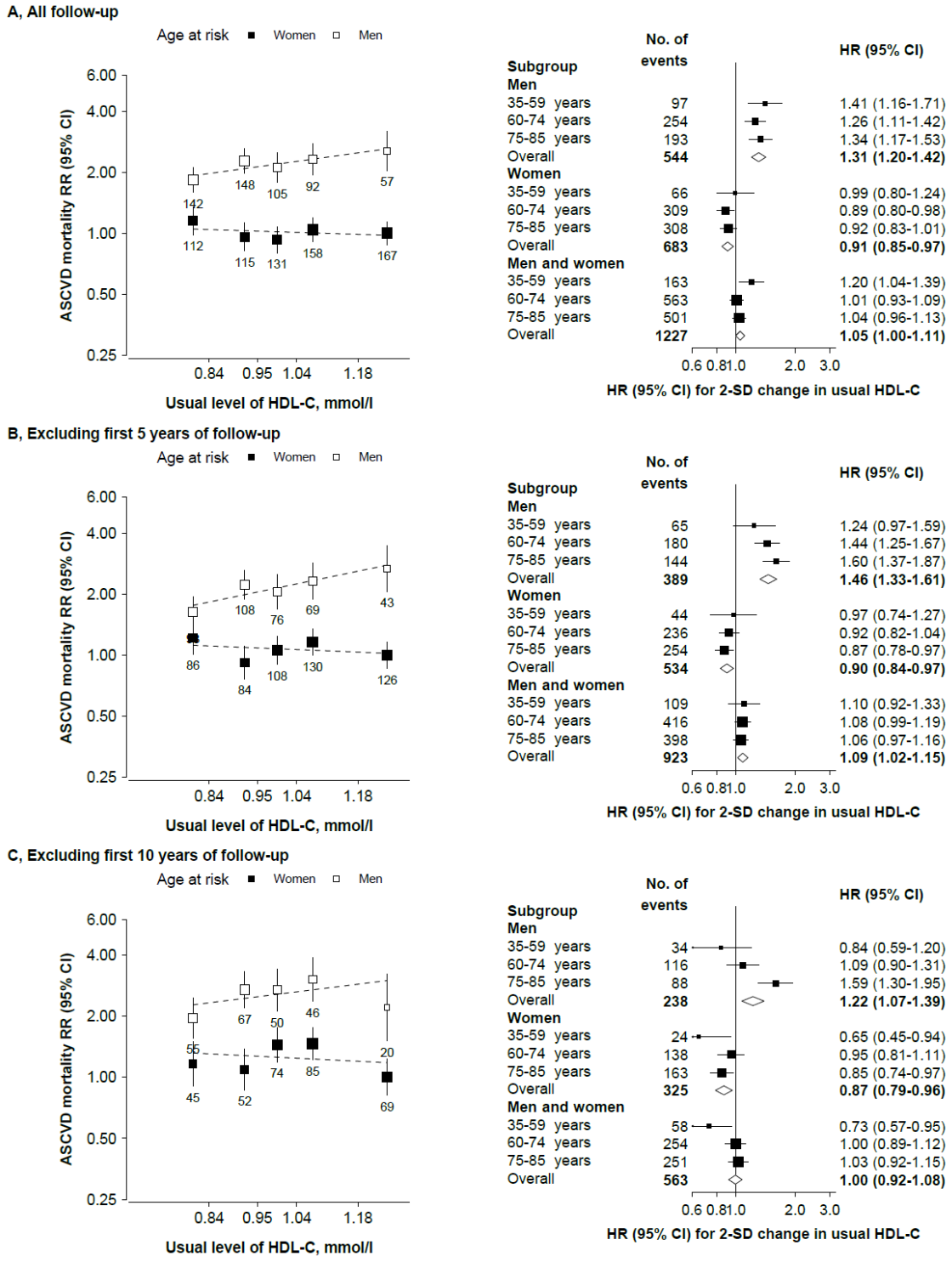
Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.S10. The association between ApoA1 and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period in participants without diabetes at baseline



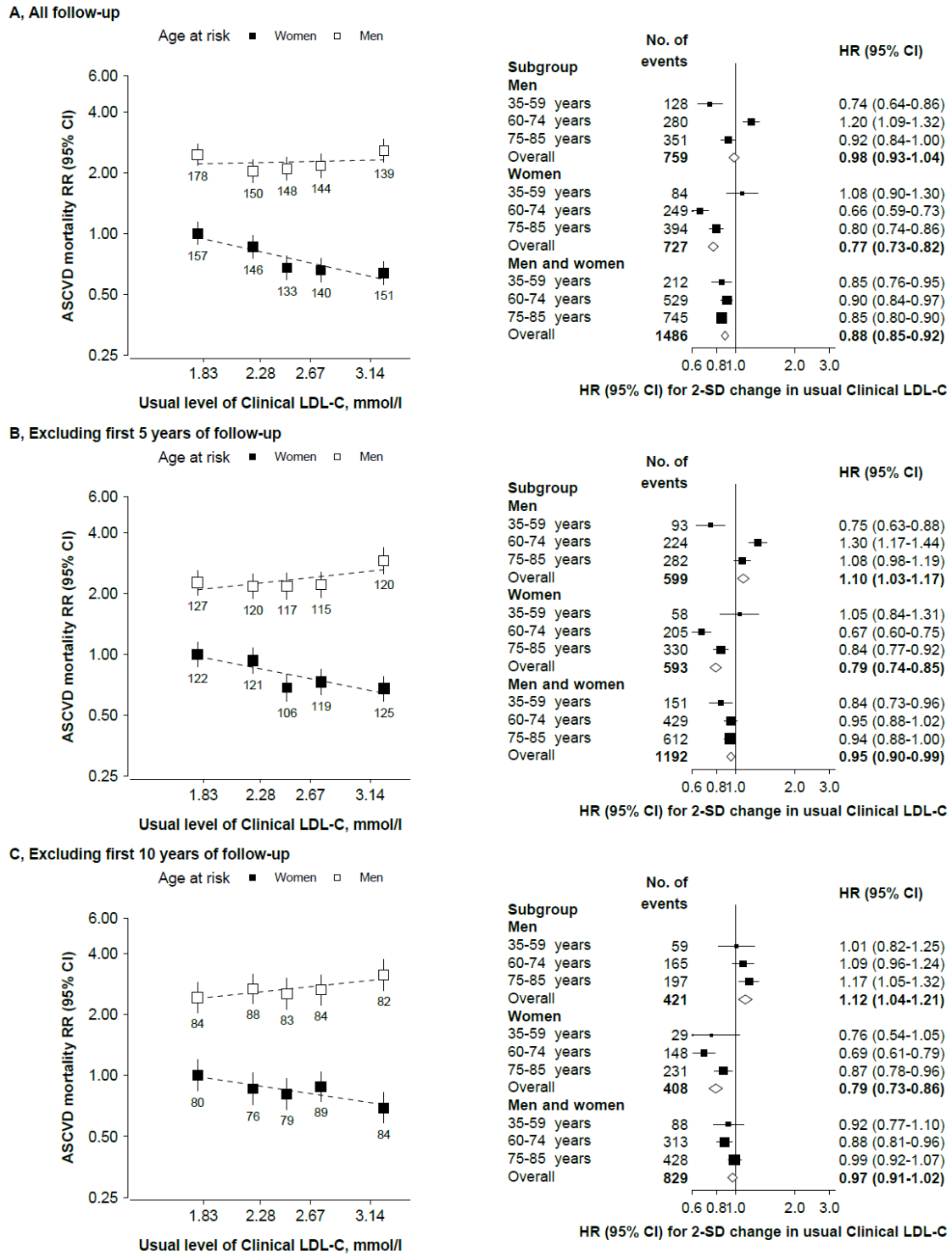
Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.S11. The association between HDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period in participants without diabetes at baseline



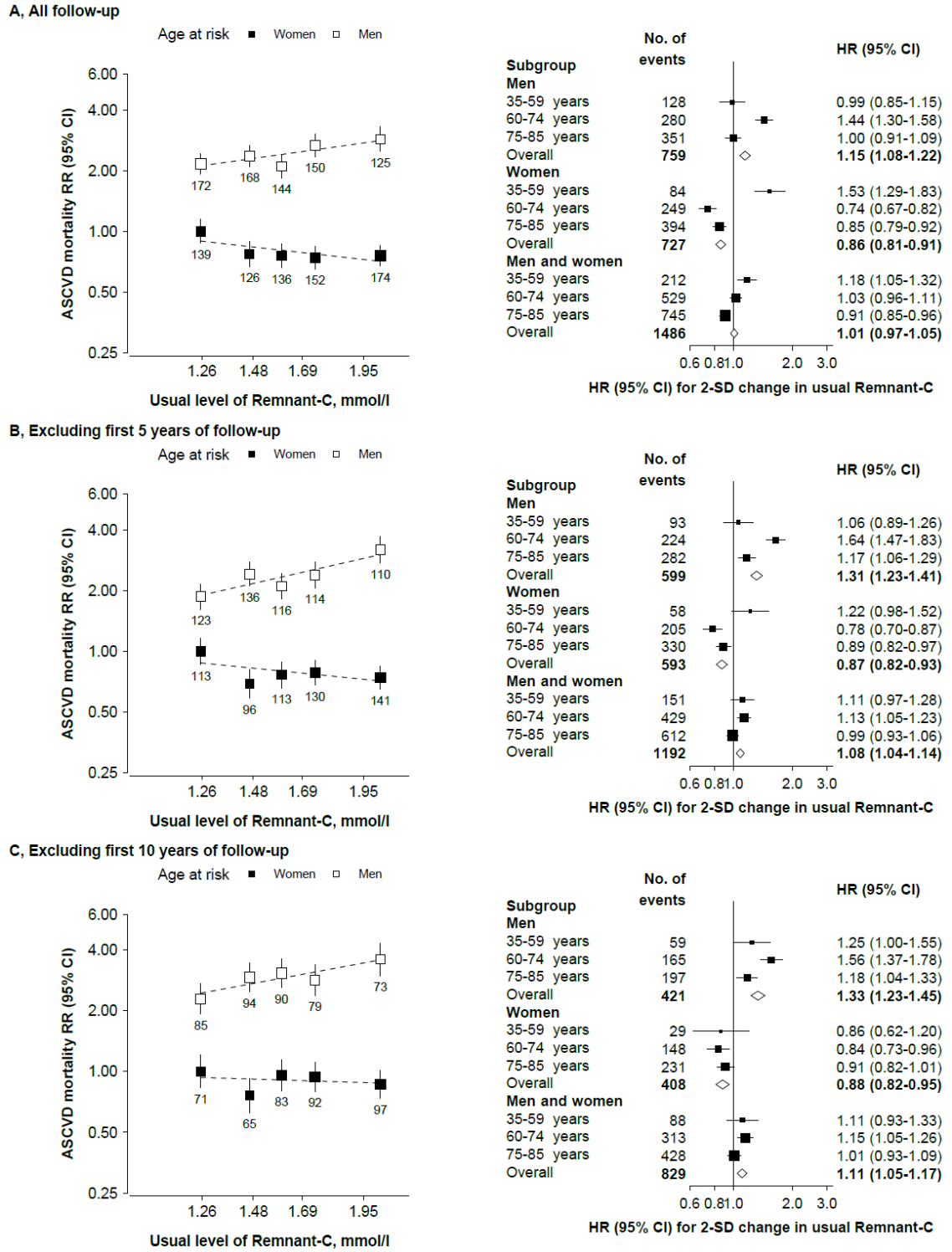
Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.S12. The association between clinical LDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period in participants without diabetes at baseline



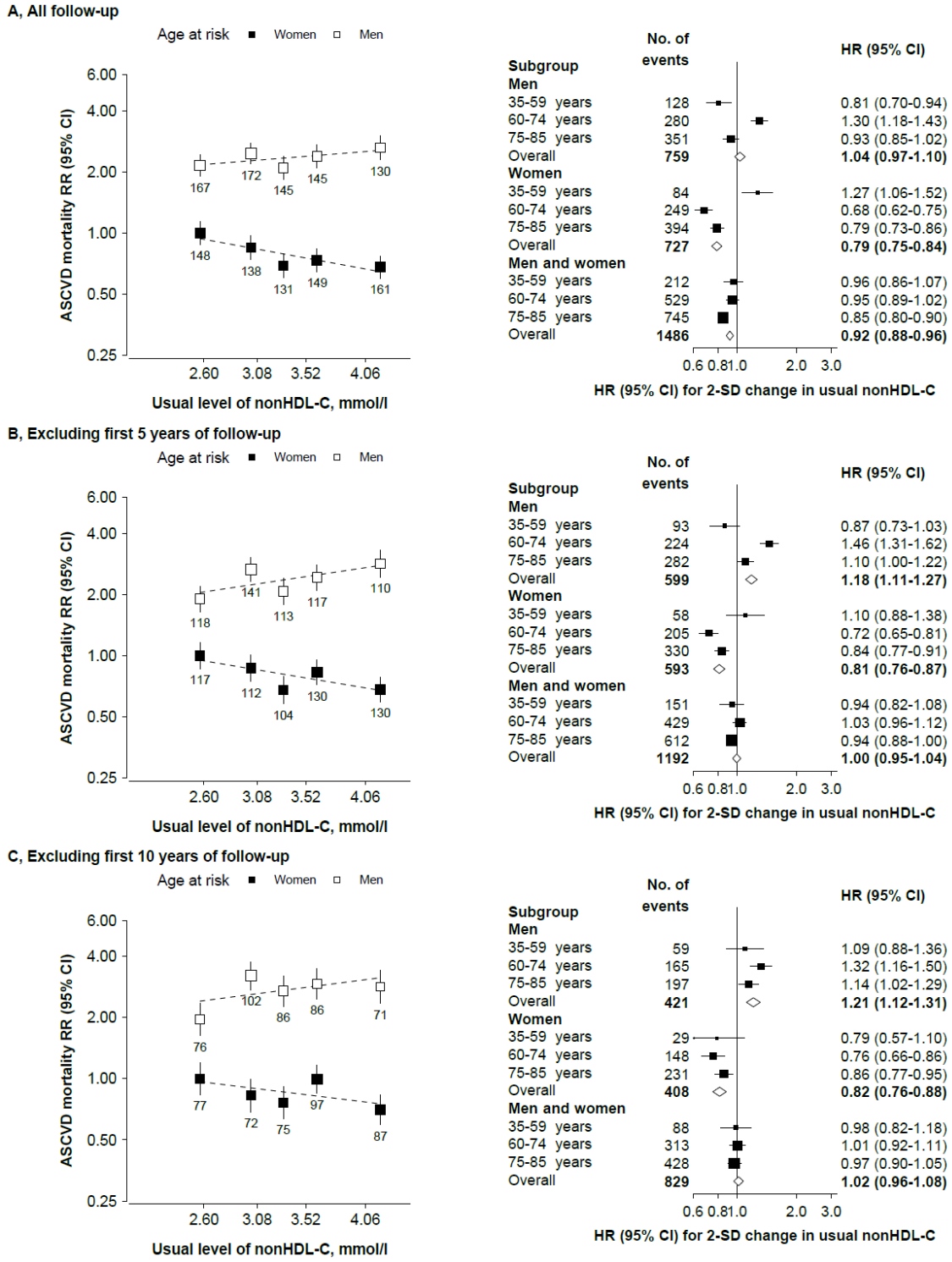
Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.S13. The association between remnant-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period in participants without diabetes at baseline



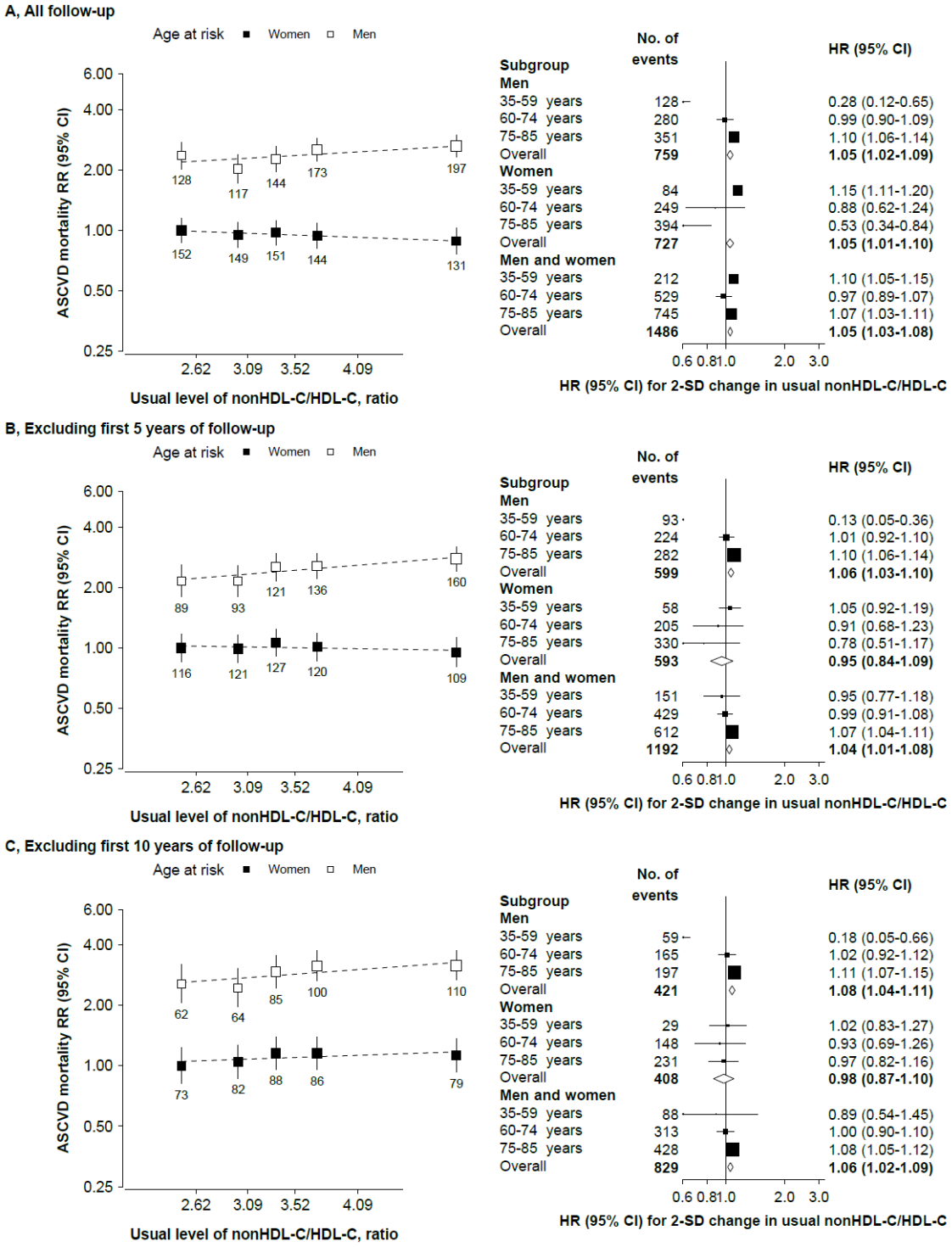
Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.S14. The association between non-HDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period in participants without diabetes at baseline



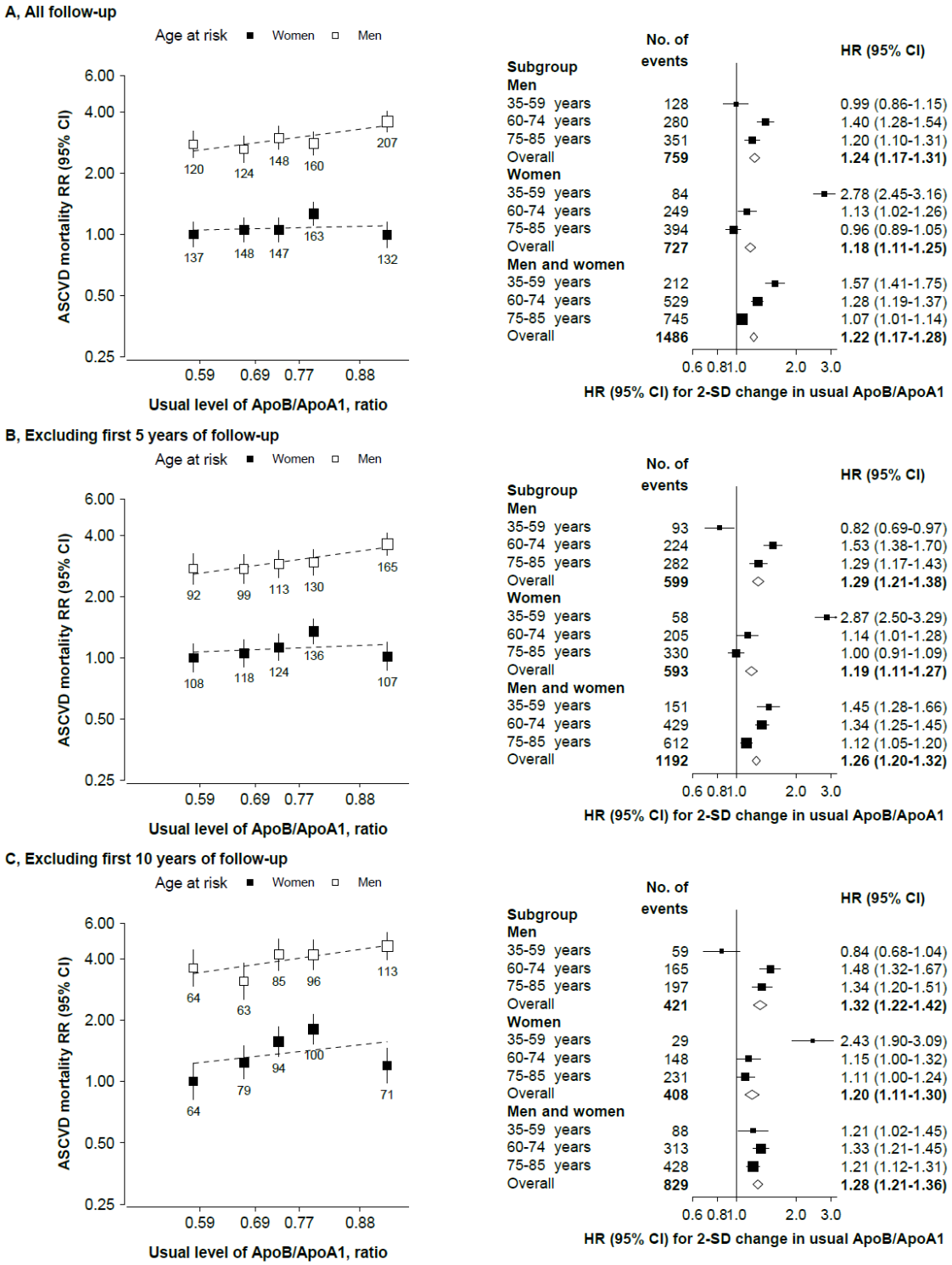
Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.S15. The association between non-HDL-C/HDL-C and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period in participants without diabetes at baseline



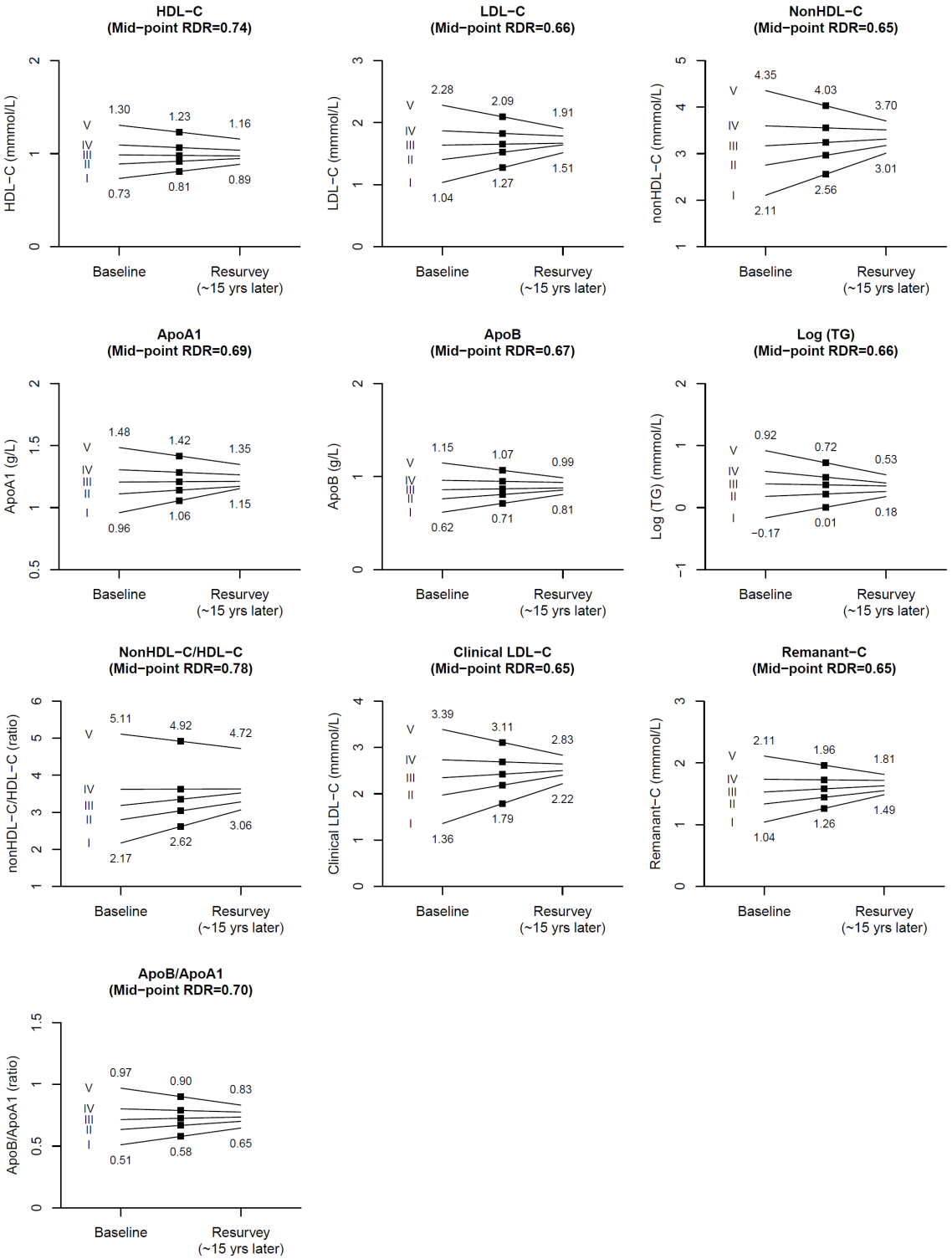
Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.S16. The association between ApoB/ApoA1 and ASCVD mortality by age-at risk groups and sex, stratified by follow-up period in participants without diabetes at baseline



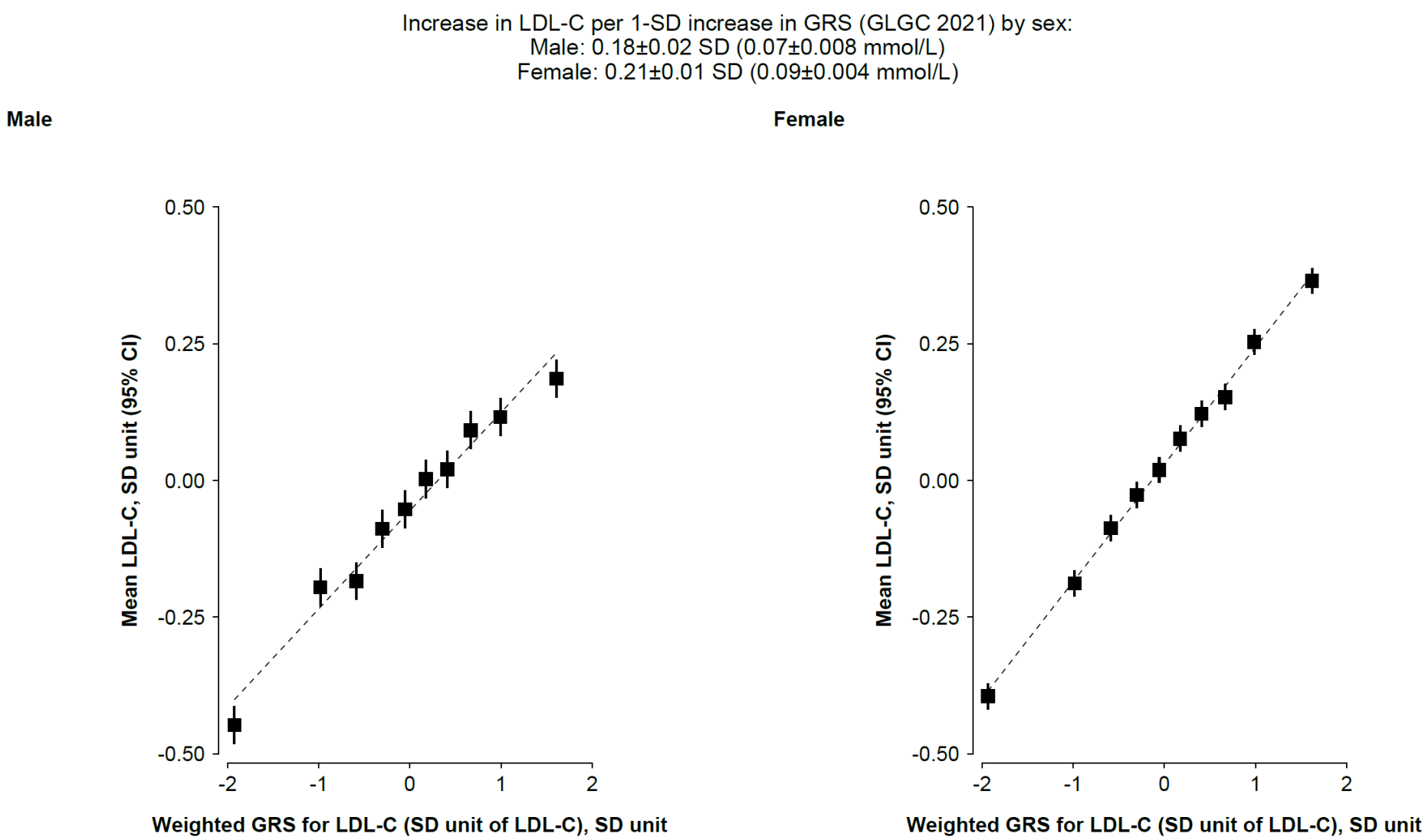
Analyses excluded those with chronic diseases at baseline (except diabetes) with adjustment for age, sex, region, and education. Hazard ratios on the left are plotted on a floating absolute scale of risk. The slopes of the age-specific lines on the left are given on the right, subdivided by sex. Each square (left or right) was inversely proportional to the variance of the log of the hazard ratio. The age-at-risk groups were divided as follows: [35, 59], [60, 74], and [75, 85).

Figure 6.S17. Extent of regression to the mean in 10 major composite lipid-related traits between baseline and resurvey, for each fifth of the baseline distribution



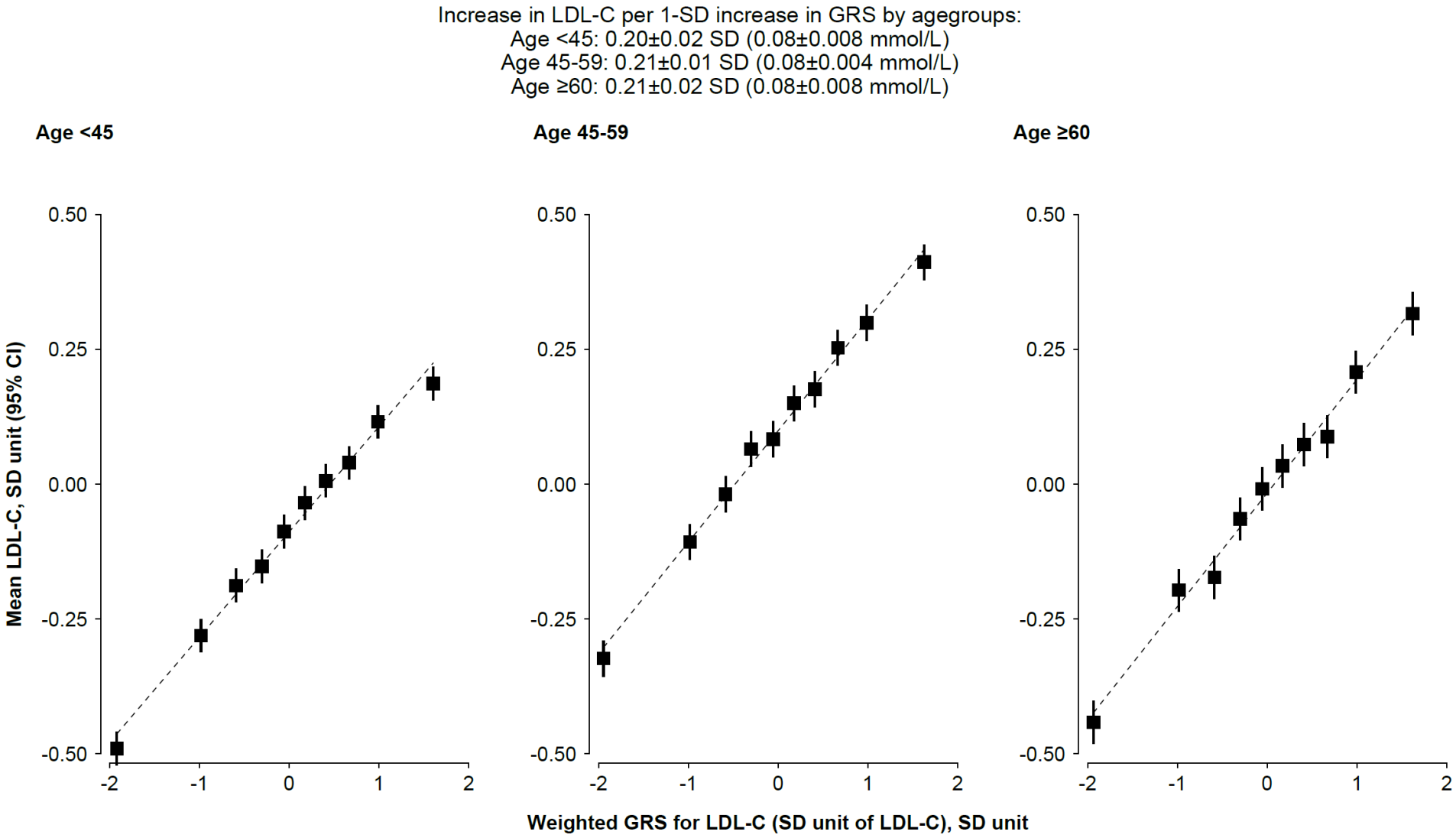
Analysis restricted to participants aged 35−84 and without prior chronic disease (chronic kidney disease, ischemic heart disease, stroke, cirrhosis, cancer, or emphysema) at recruitment, with complete data on analysis covariates and with complete, plausible, lipid−related NMR data at both recruitment and resurvey. The regression dilution ratio (RDR) is calculated as the ratio of the range of mid−point to baseline means; for example for HDL−C this equates to (1.23−0.81)/(1.30−0.73)=0.74

Figure 7.S1 Mean baseline NMR-measured LDL-C by GRS for LDL-C using SNPs from GLGC 2021 in male and female



Adjusted for age, study region, and ancestry groups. Squares represent mean LDL-C level with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals.

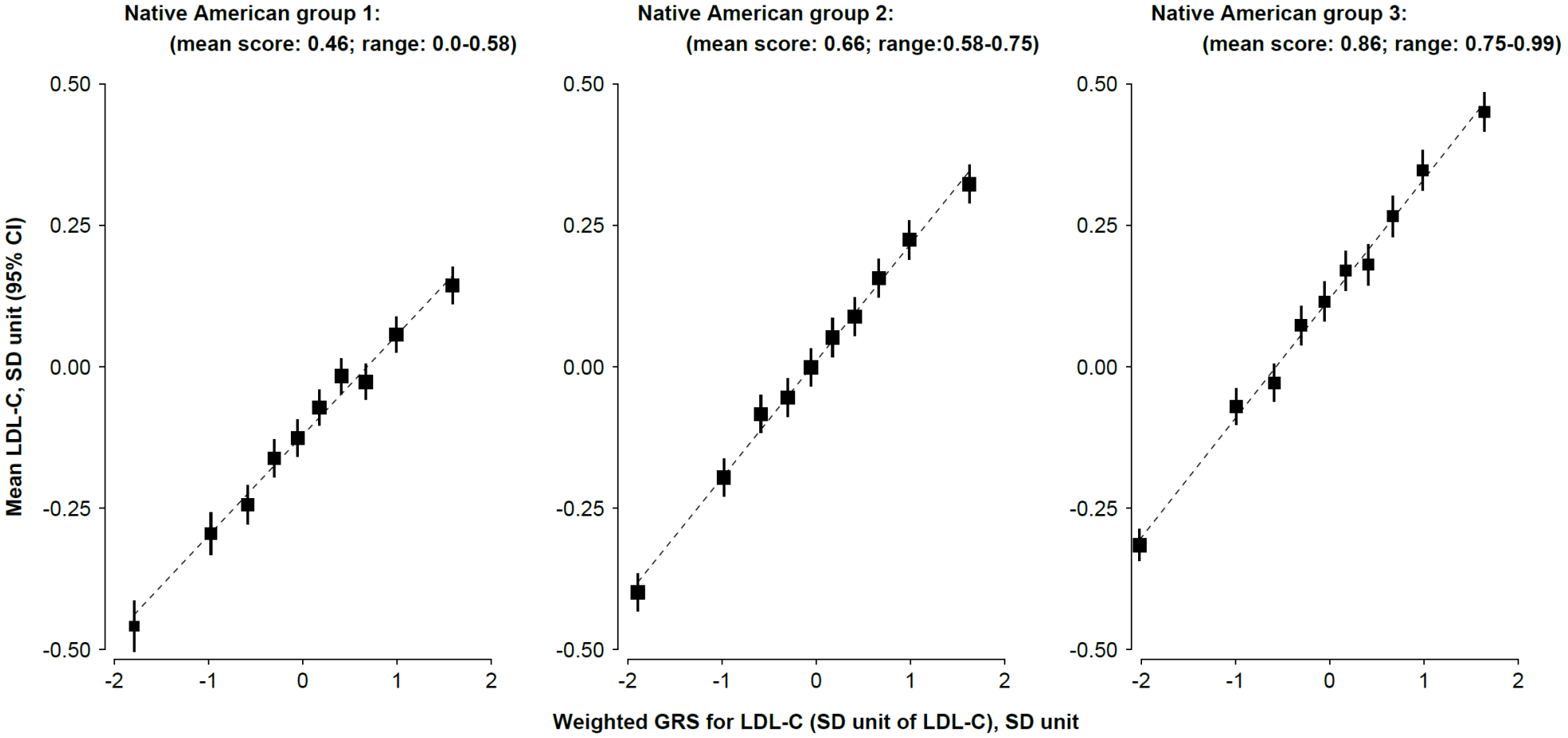
Figure 7.S2 Mean baseline NMR-measured LDL-C by GRS for LDL-C using SNPs from GLGC 2021 by age groups



Adjusted for sex, study region, and ancestry groups. Squares represent mean LDL-C level with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals.

Figure 7.S3 Mean baseline NMR-measured LDL-C by GRS for LDL-C using SNPs from GLGC 2021 by ancestry groups

Increase in LDL-C per 1-SD increase in GRS by ancestry groups:
Ancestry group 1: 0.18±0.01 SD (0.09±0.005 mmol/L)
Ancestry group 2: 0.21±0.01 SD (0.11±0.005 mmol/L)
Ancestry group 3: 0.21±0.01 SD (0.11±0.005 mmol/L)



Adjusted for age, sex, and study region. Squares represent mean LDL-C level with area inversely proportional to the variance.
Vertical lines represent the corresponding 95% confidence intervals.
The three groups are defined by the tertile of the Native American ancestry score.

Figure 7.S4 Mean baseline NMR-measured LDL-C by GRS for LDL-C using SNPs from GLGC 2021 by study region

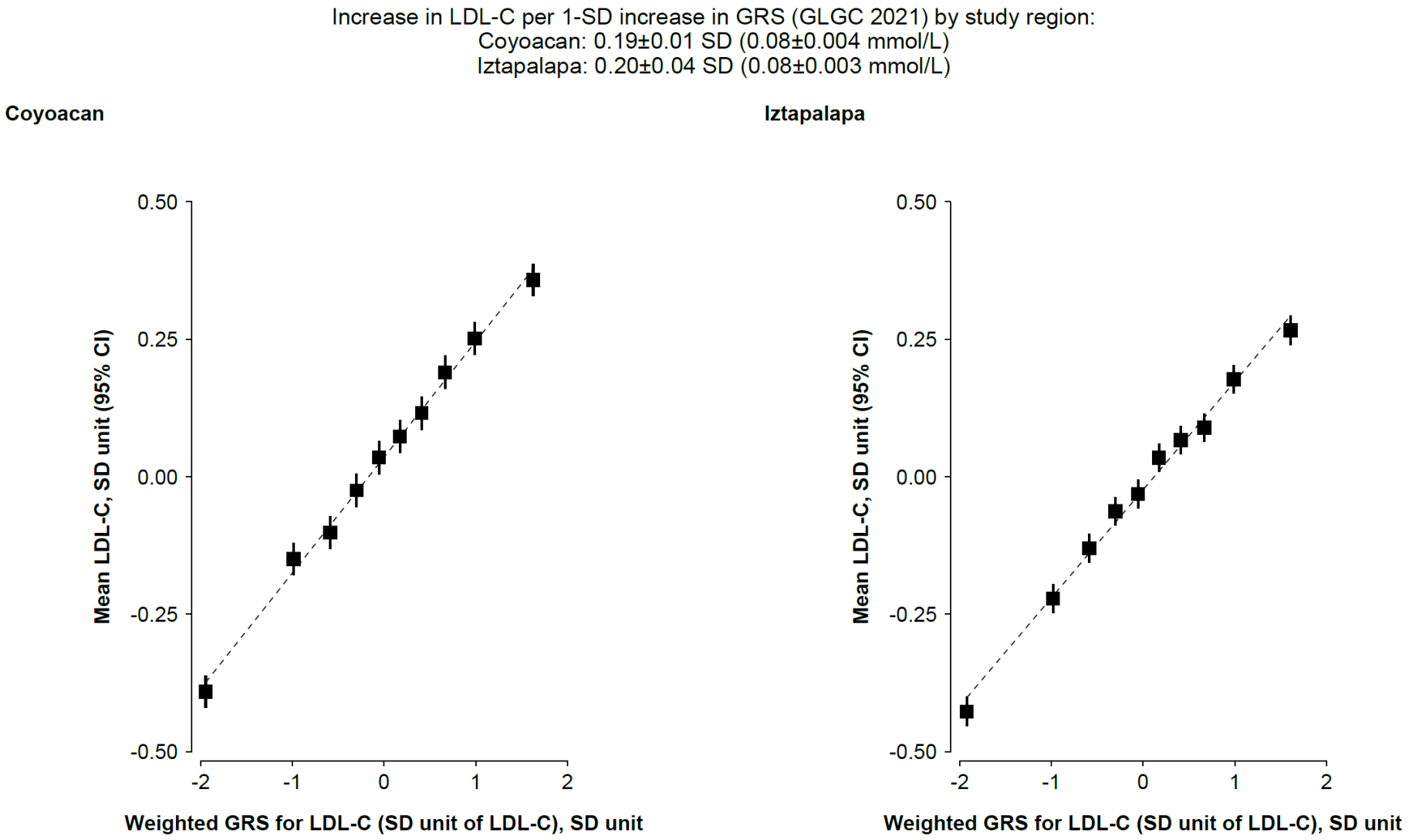
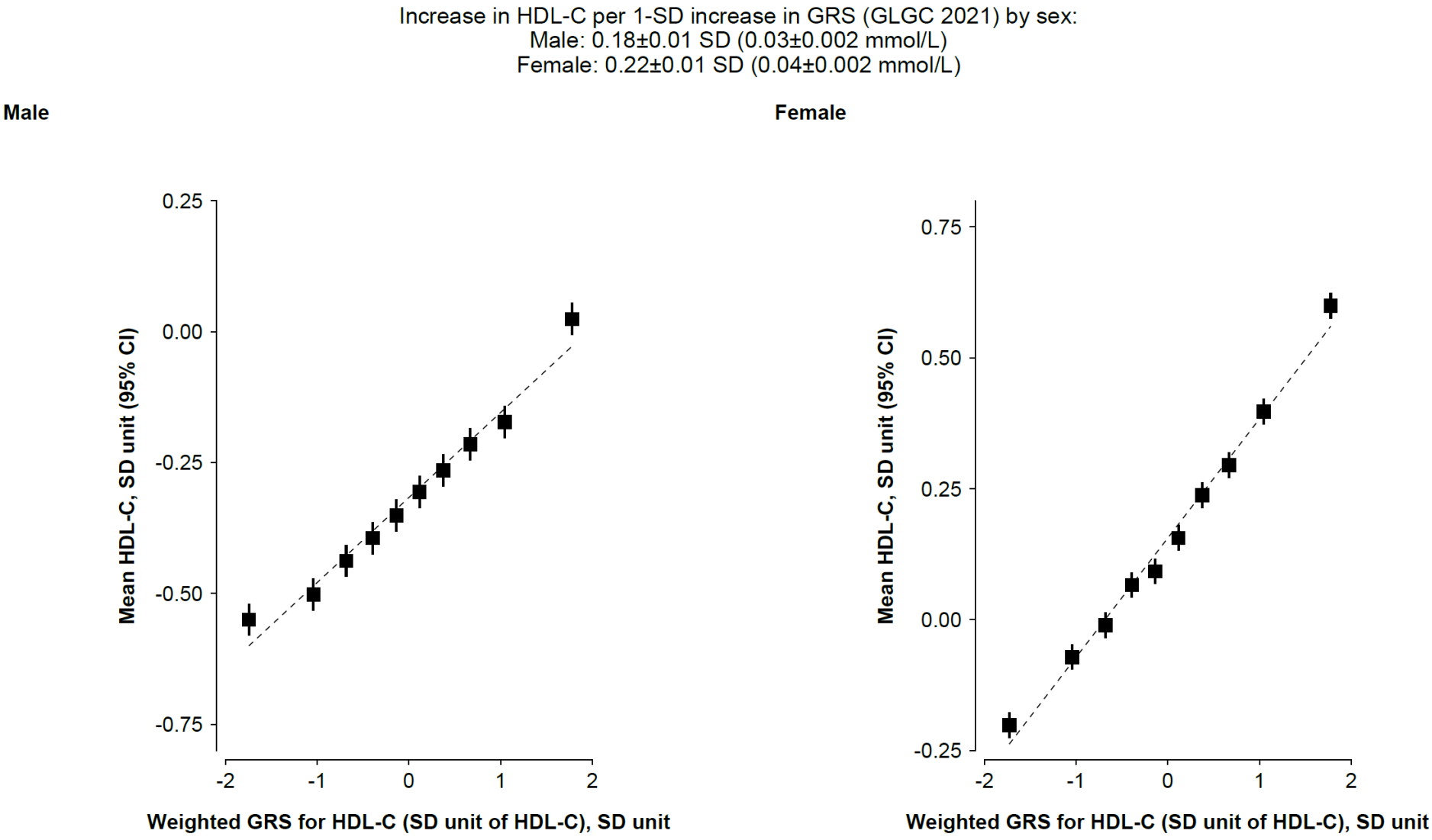
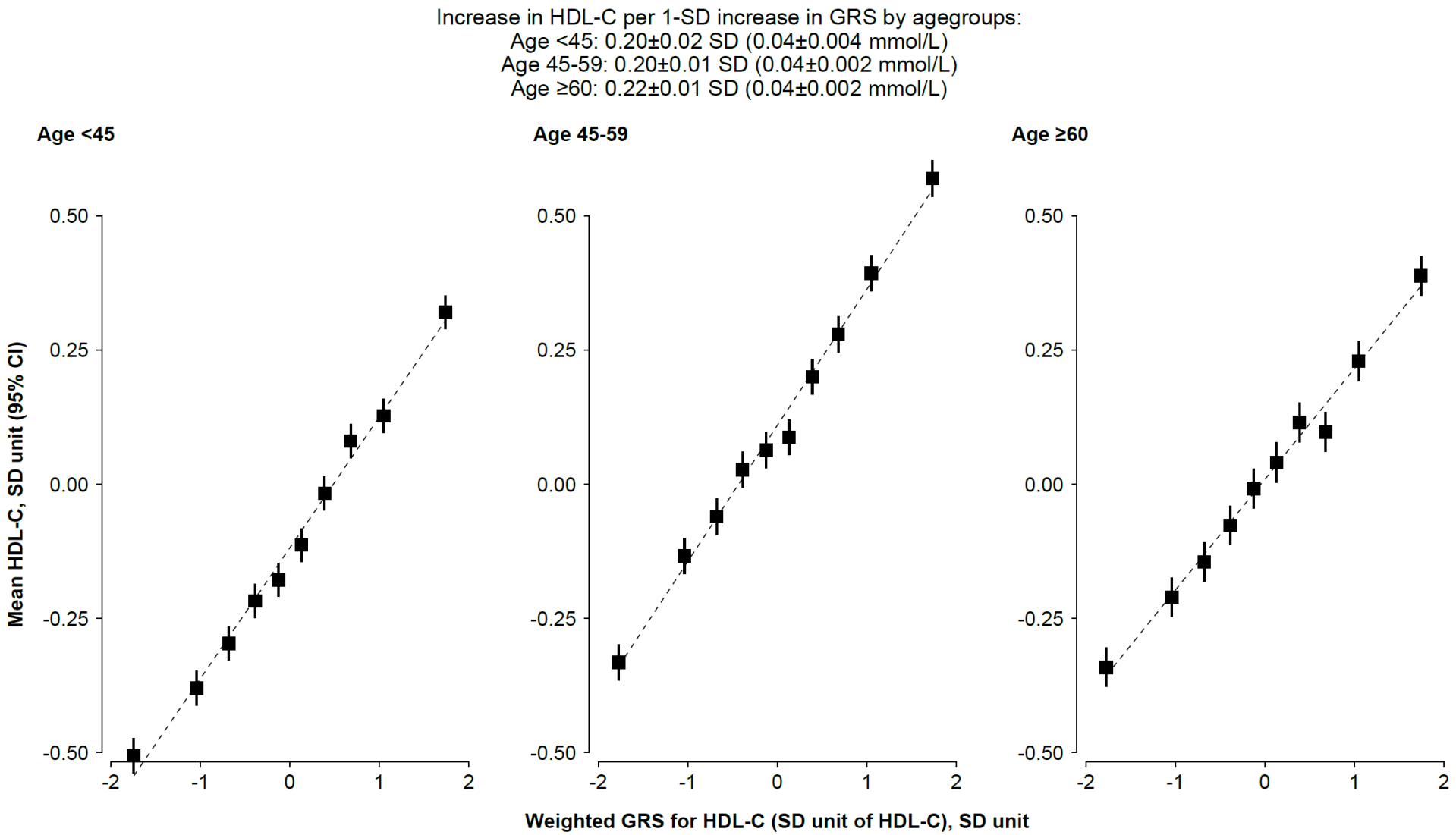


Figure 7.S5 Mean baseline NMR-measured HDL-C by GRS for HDL-C using SNPs from GLGC 2021 in male and female



Adjusted for age, study region, and ancestry score. Squares represent mean HDL-C level with area inversely proportional to the variance.
Vertical lines represent the corresponding 95% confidence intervals.

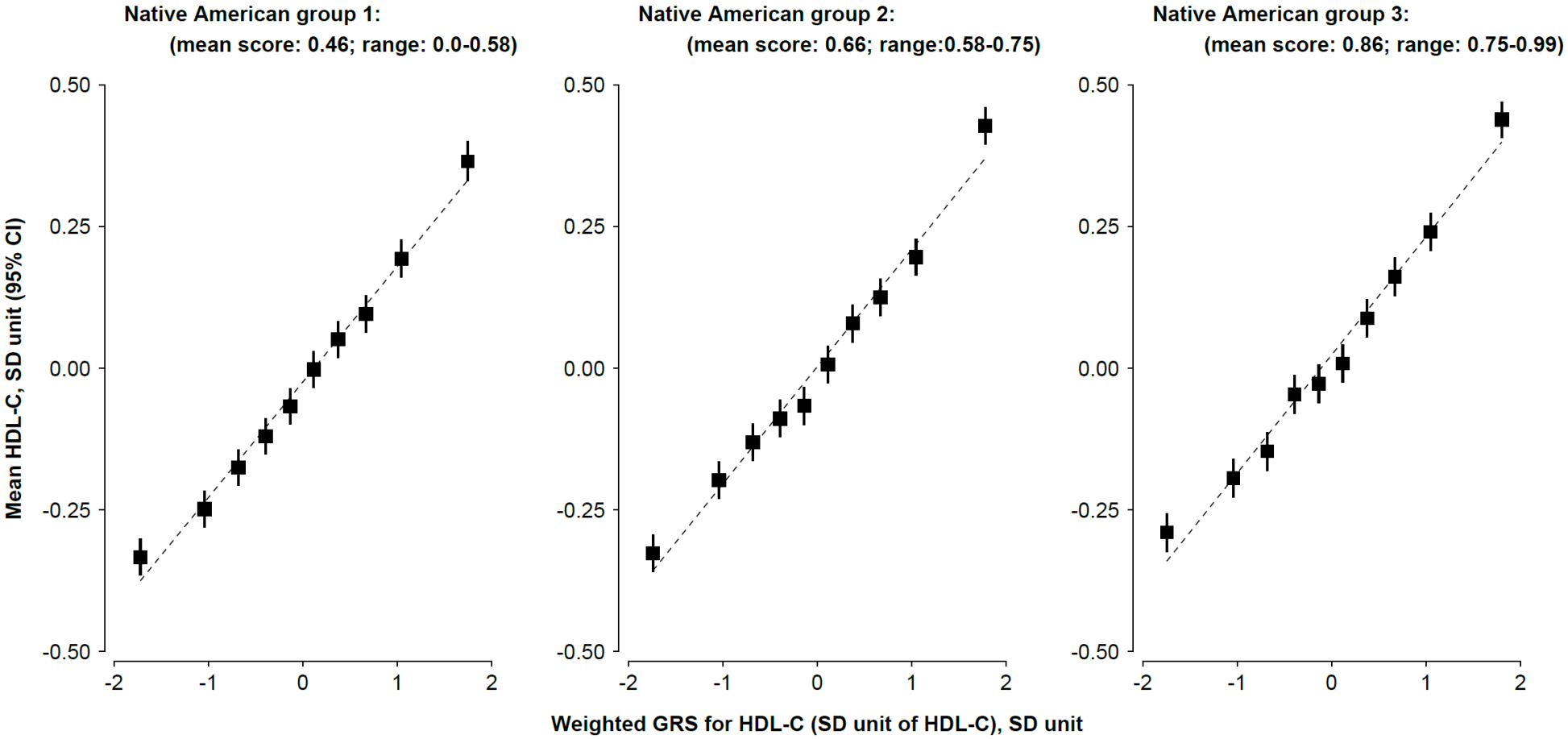
Figure 7.S6 Mean baseline NMR-measured HDL-C by GRS for HDL-C using SNPs from GLGC 2021 by age groups



Adjusted for sex, study region, and ancestry score. Squares represent mean HDL-C level with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals.

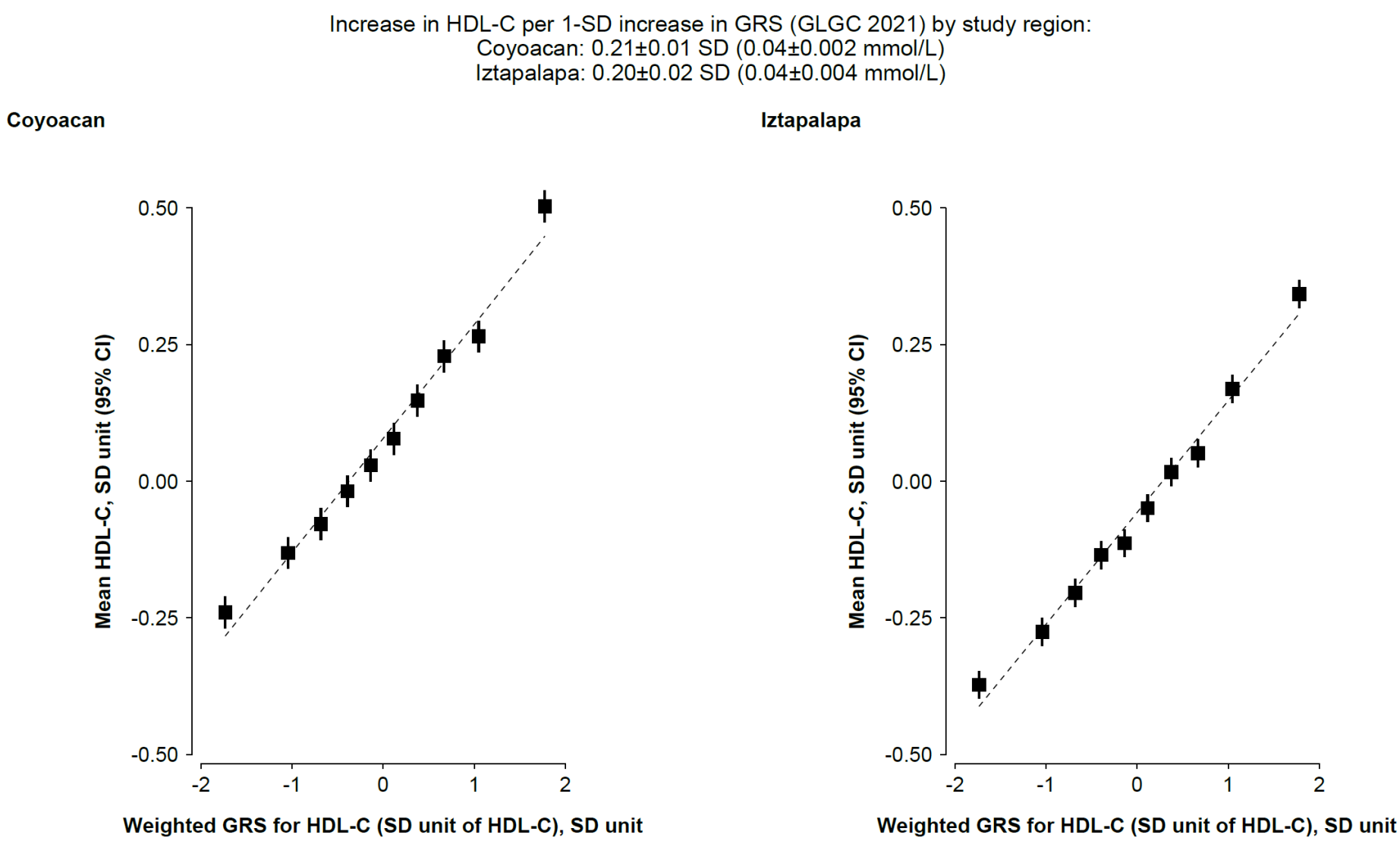
Figure 7.S7 Mean baseline NMR-measured HDL-C by GRS for HDL-C using SNPs from GLGC 2021 by ancestry groups

Increase in HDL-C per 1-SD increase in GRS by ancestry groups:
Ancestry group 1: 0.20±0.02 SD (0.04±0.004 mmol/L)
Ancestry group 2: 0.21±0.01 SD (0.04±0.002 mmol/L)
Ancestry group 3: 0.21±0.01 SD (0.04±0.002 mmol/L)



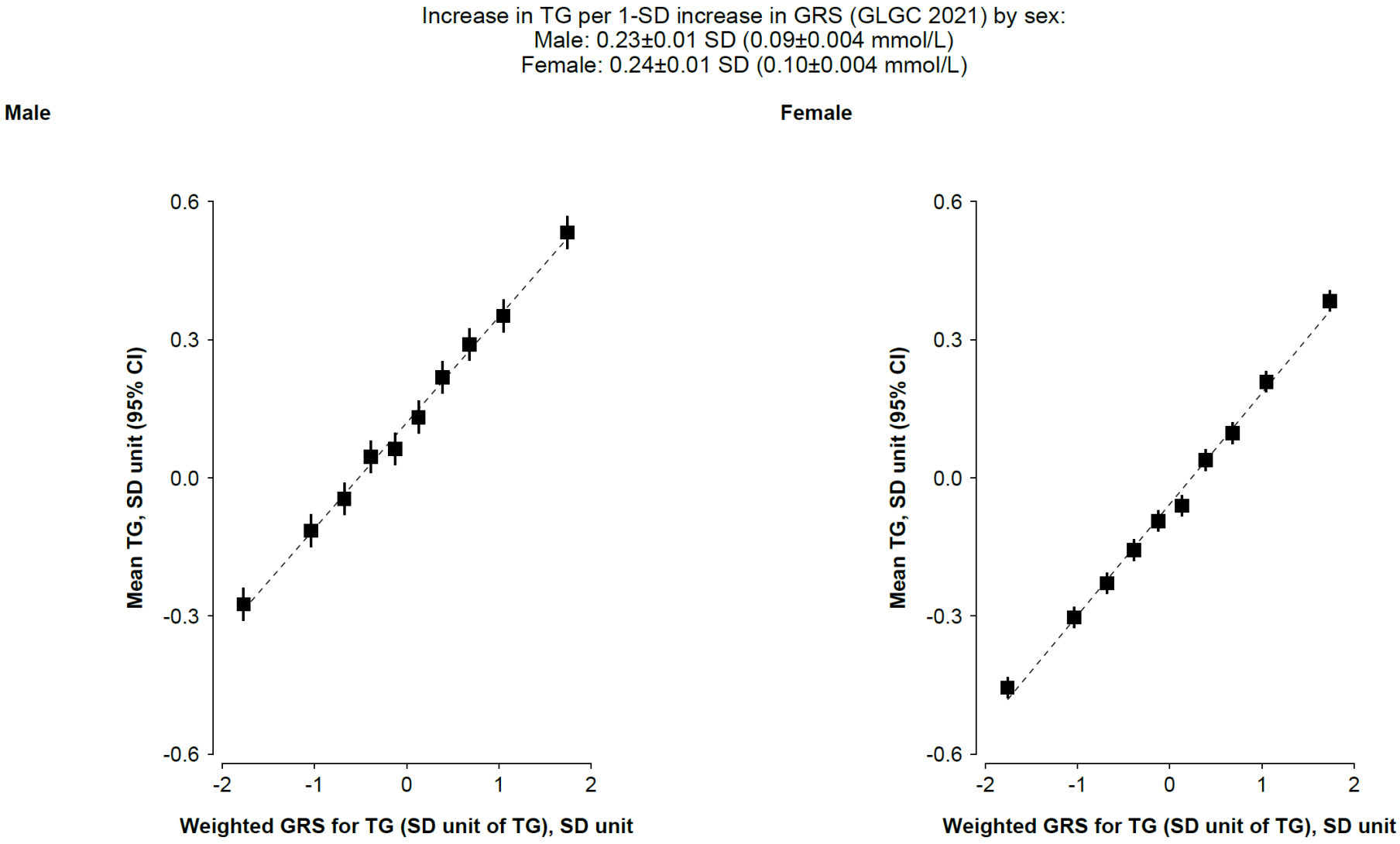
Adjusted for age, sex, and study region. Squares represent mean LDL-C level with area inversely proportional to the variance.
Vertical lines represent the corresponding 95% confidence intervals.
The three groups are defined by the tertile of the Native American ancestry score.

Figure 7.S8 Mean baseline NMR-measured HDL-C by GRS for HDL-C using SNPs from GLGC 2021 by study region



Adjusted for age, sex, and ancestry score. Squares represent mean HDL-C level with area inversely proportional to the variance.
Vertical lines represent the corresponding 95% confidence intervals.

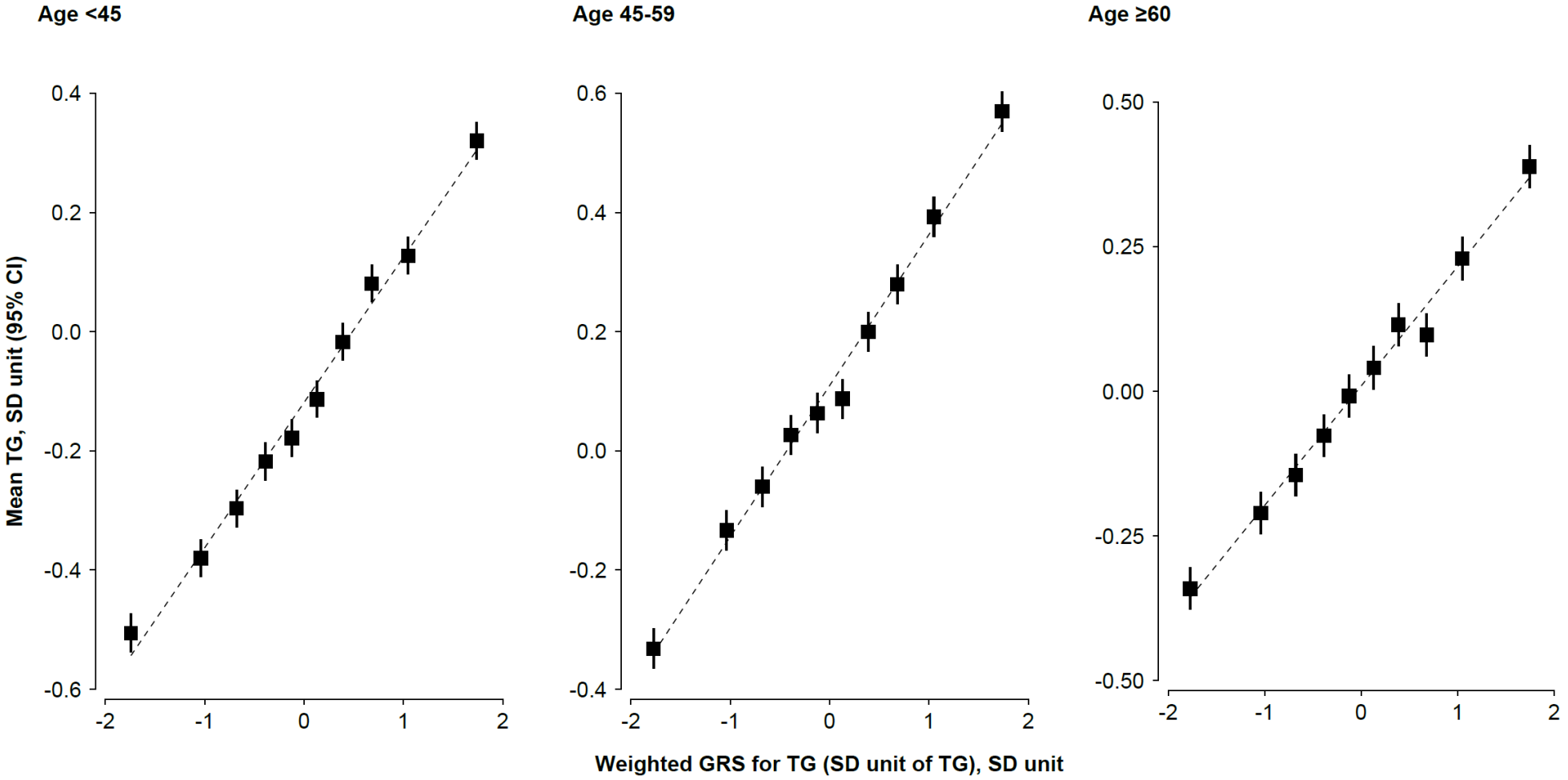
Figure 7.S9 Mean baseline NMR-measured TG by GRS for TG using SNPs from GLGC 2021 in male and female



TG levels were natural-log transformed.
Adjusted for age, study region, and ancestry score. Squares represent mean TG level with area inversely proportional to the variance.
Vertical lines represent the corresponding 95% confidence intervals.

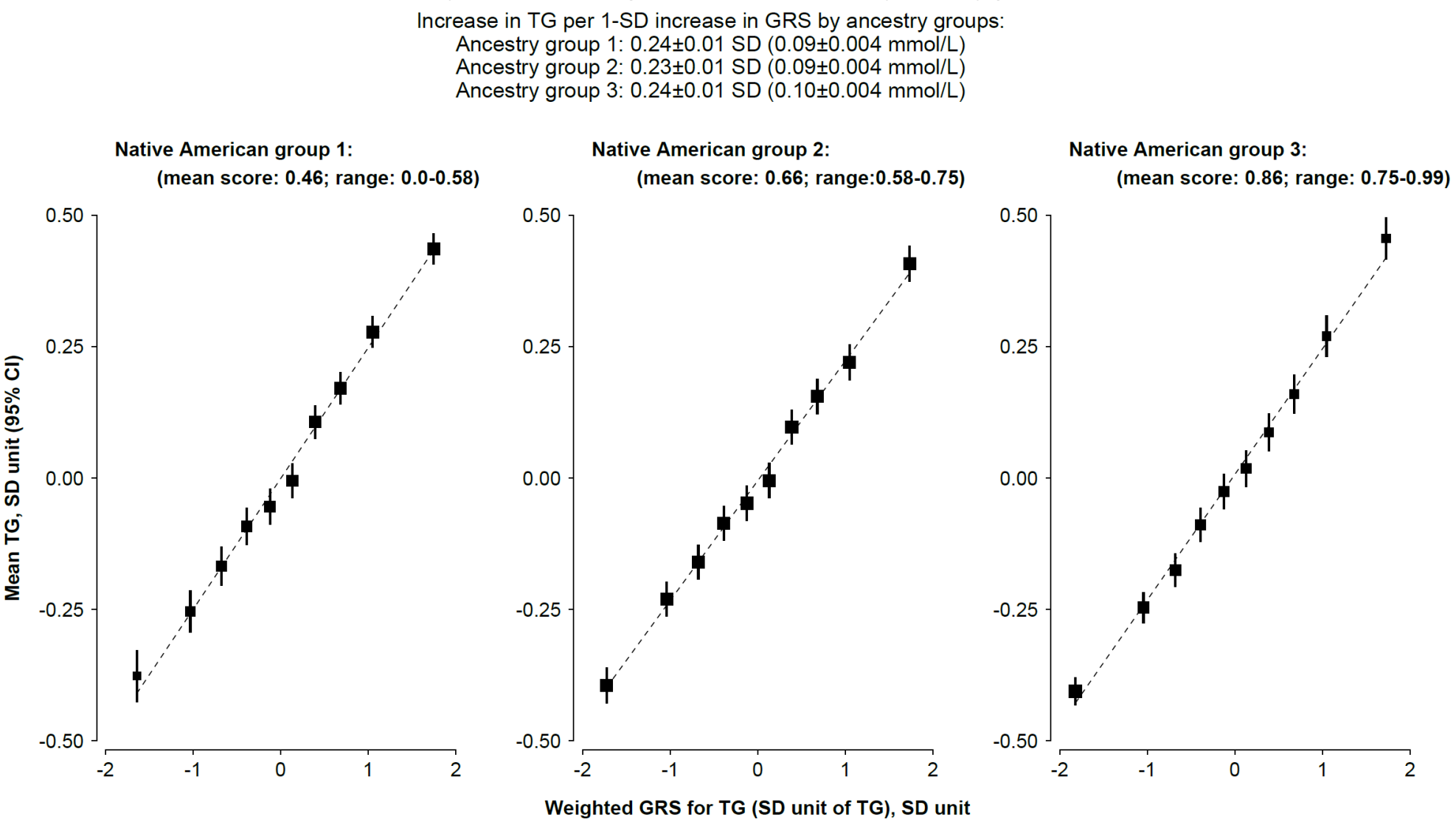
Figure 7.S10 Mean baseline NMR-measured TG by GRS for TG using SNPs from GLGC 2021 by age groups

Increase in TG per 1-SD increase in GRS by agegroups:
Age <45: 0.24±0.01 SD (0.09±0.004 mmol/L)
Age 45-59: 0.25±0.01 SD (0.10±0.004 mmol/L)
Age ≥60: 0.21±0.01 SD (0.08±0.004 mmol/L)



Adjusted for sex, study region, and ancestry score. Squares represent mean HDL-C level with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals.

Figure 7.S11 Mean baseline NMR-measured TG by GRS for TG using SNPs from GLGC 2021 by ancestry groups



Adjusted for age, sex, and study region. Squares represent mean LDL-C level with area inversely proportional to the variance.
Vertical lines represent the corresponding 95% confidence intervals.
The three groups are defined by the tertile of the Native American ancestry score.

Figure 7.S12 Mean baseline NMR-measured TG by GRS for TG using SNPs from GLGC 2021 by study region

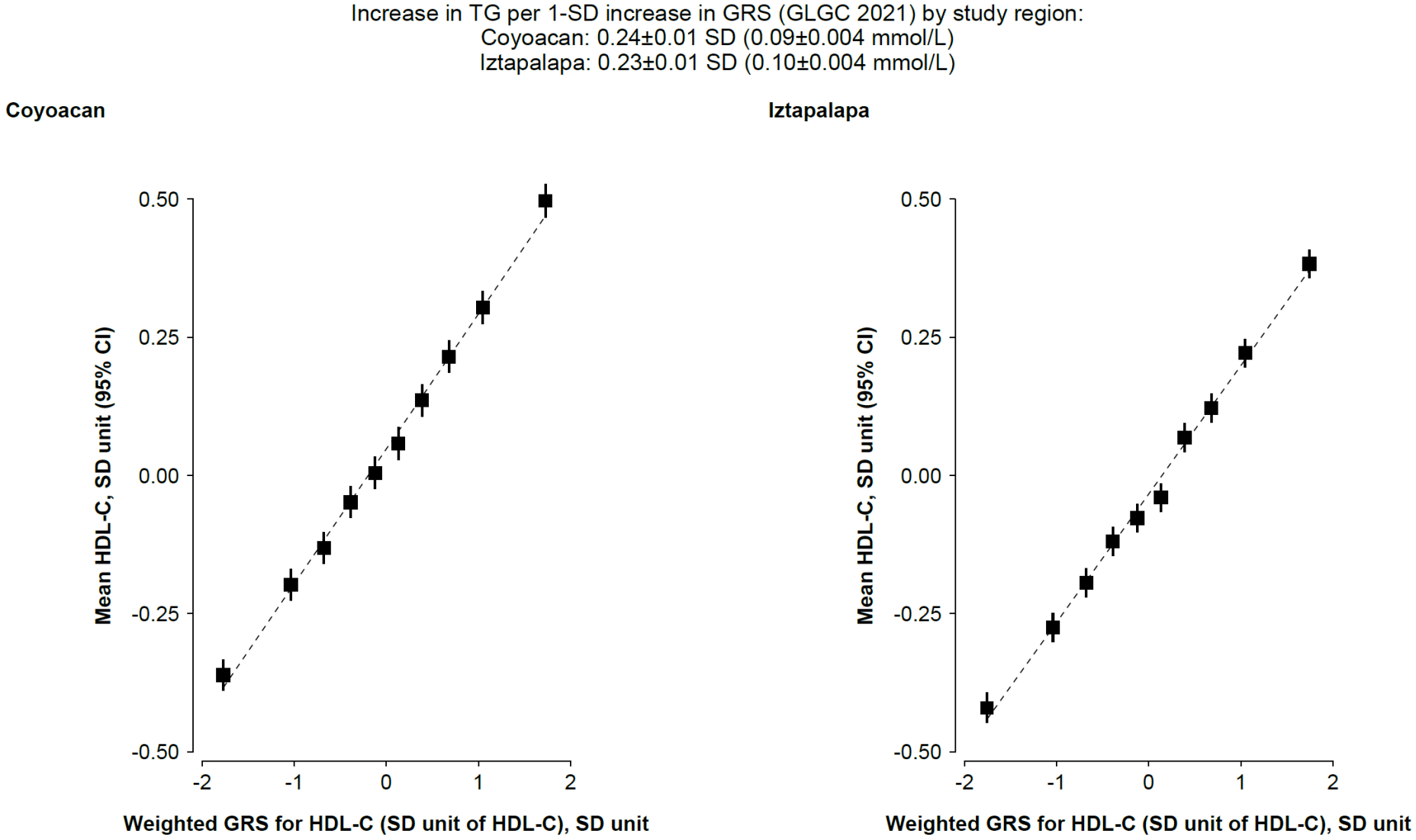
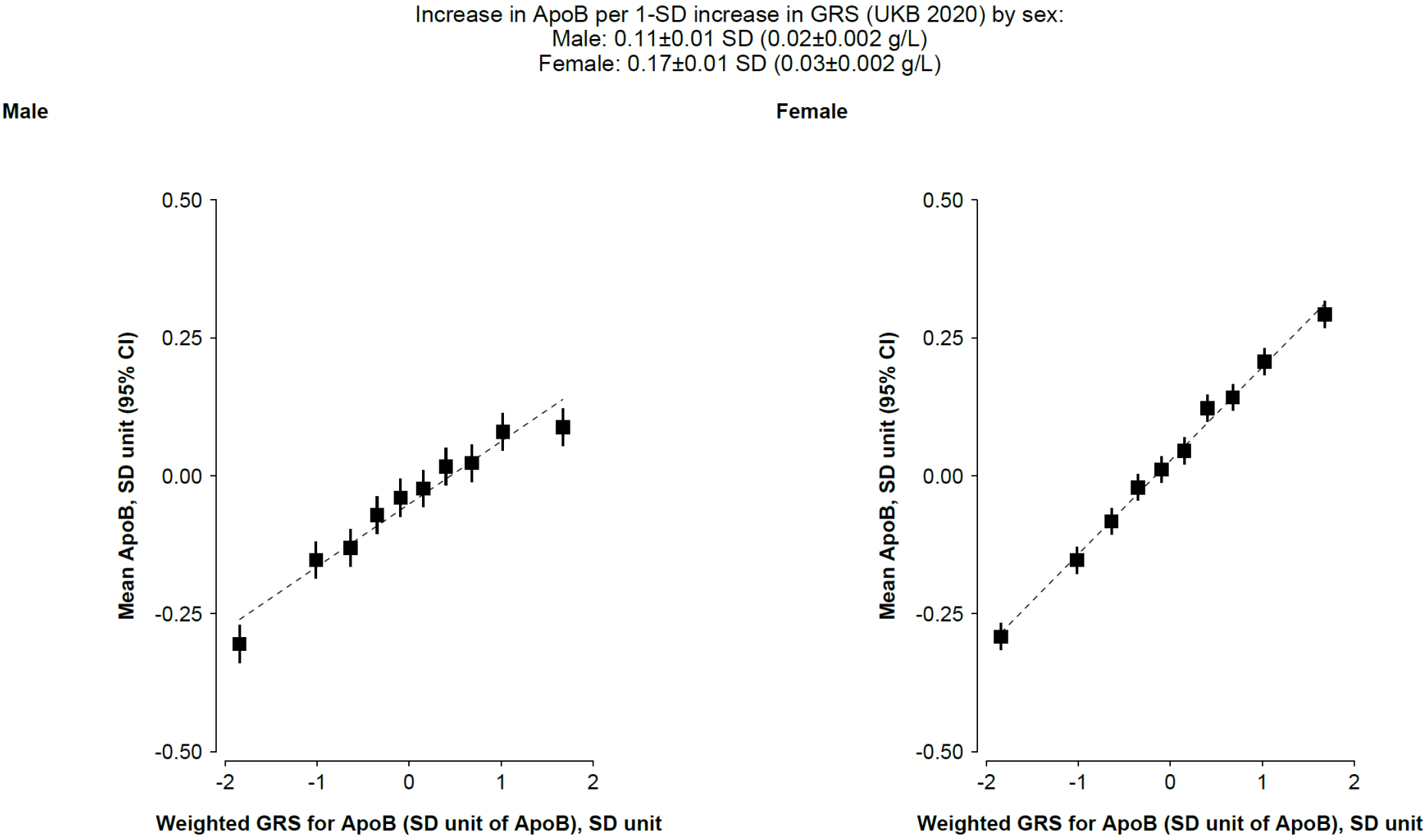
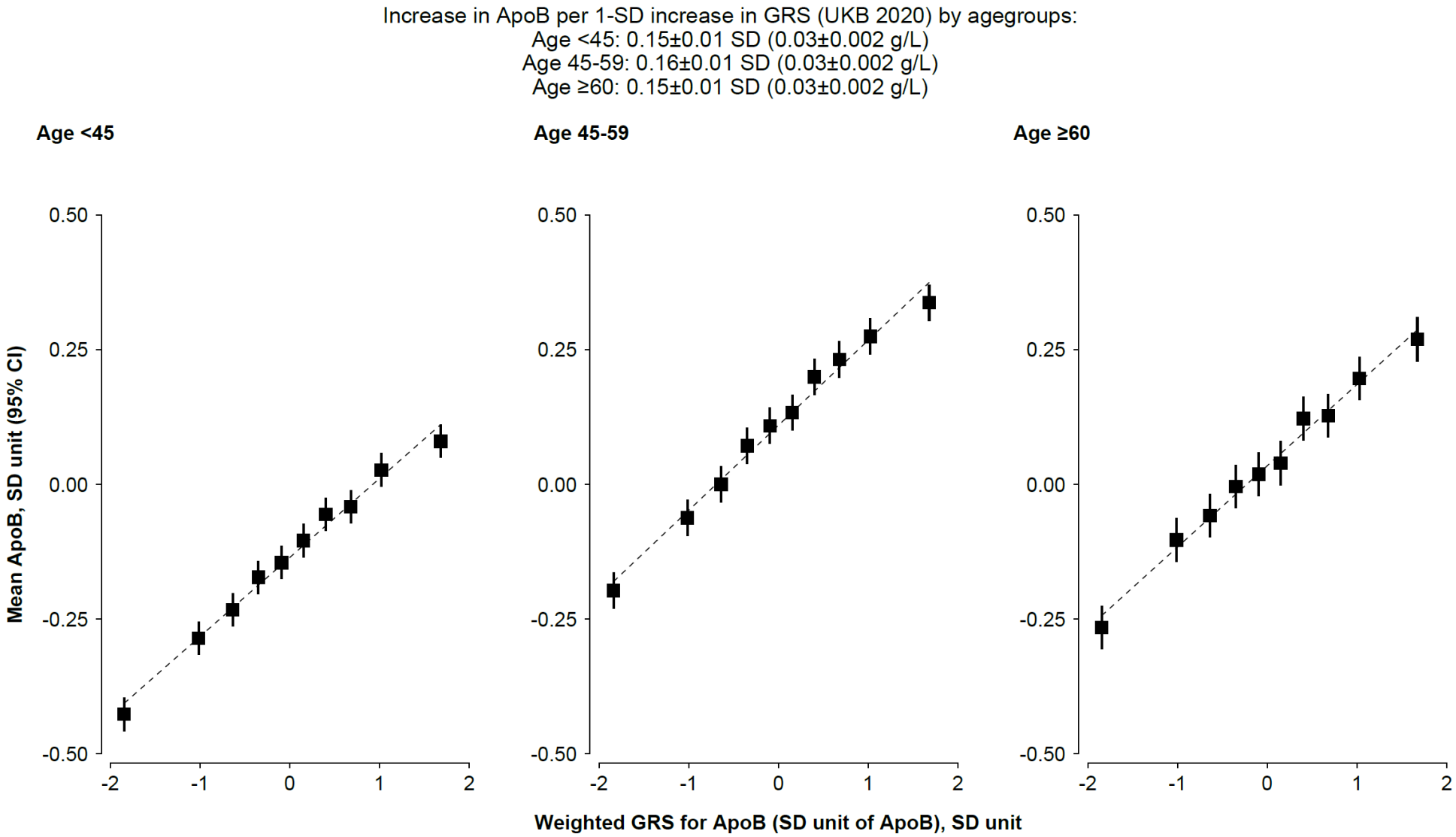


Figure 7.S13 Mean baseline NMR-measured ApoB by GRS for ApoB using SNPs from UKB 2020 in male and female



Adjusted for age, study region, and ancestry score. Squares represent mean ApoB level with area inversely proportional to the variance.
Vertical lines represent the corresponding 95% confidence intervals.

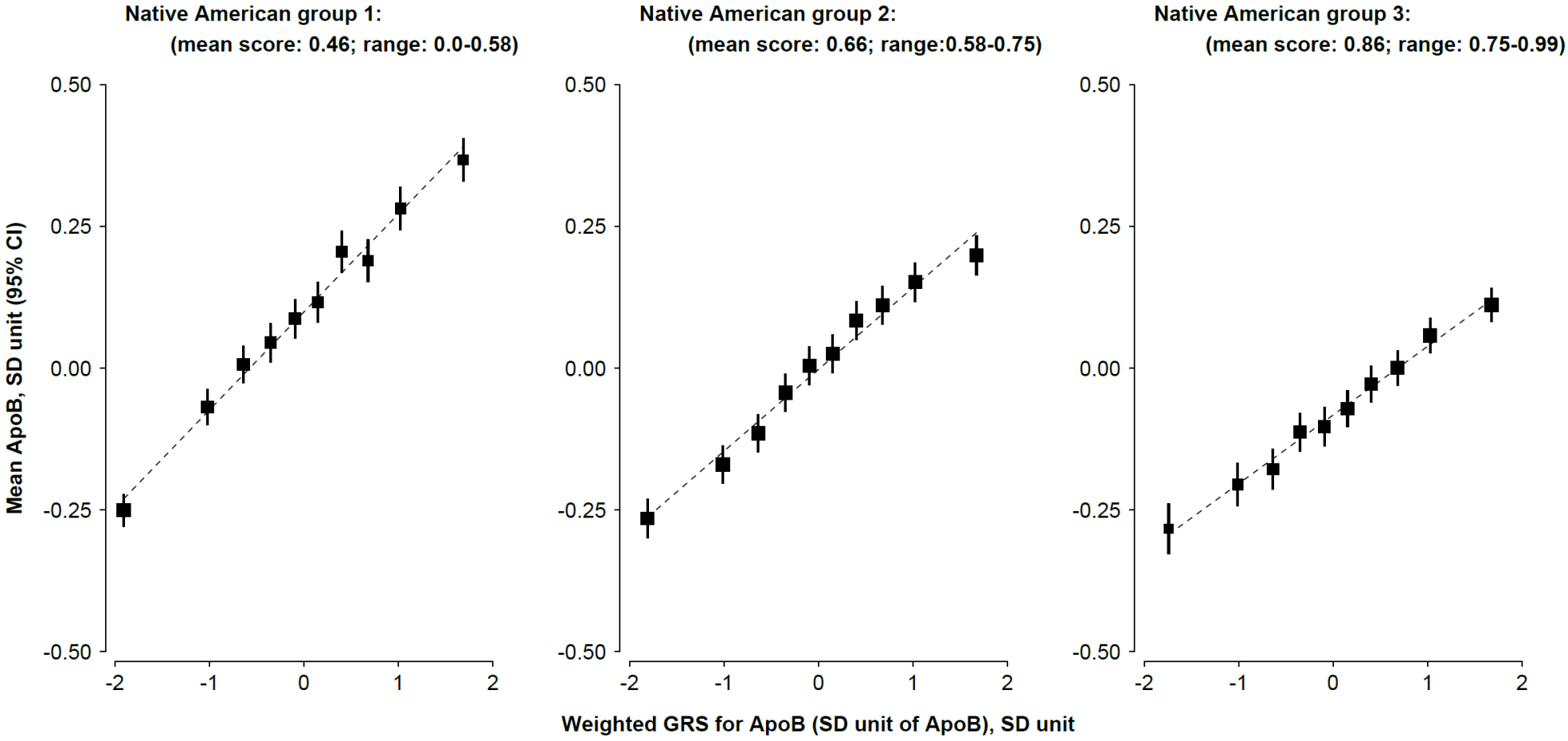
Figure 7.S14 Mean baseline NMR-measured ApoB by GRS for ApoB using SNPs from UKB 2020 by age groups



Adjusted for sex, study region, and ancestry score. Squares represent mean ApoB level with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals.

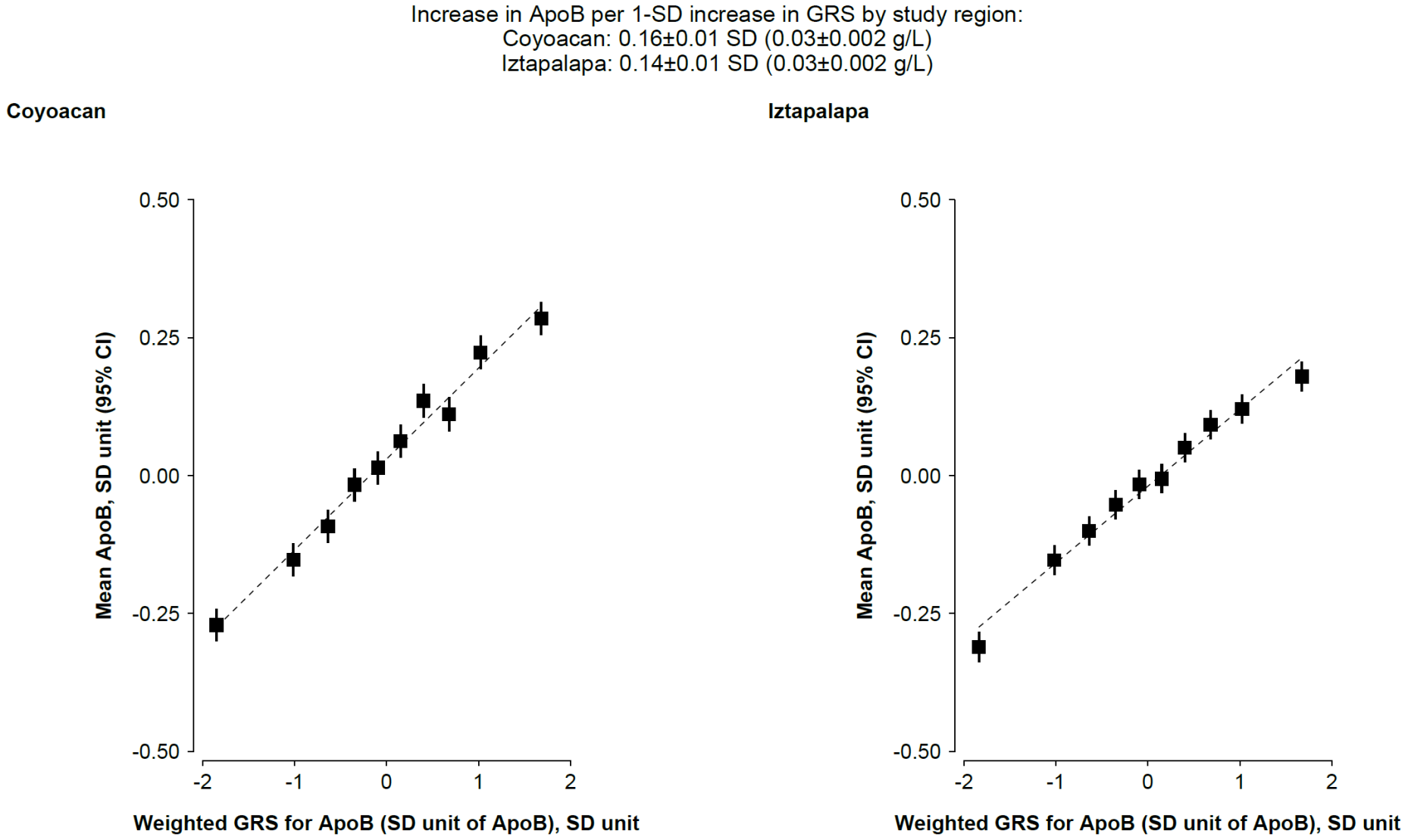
Figure 7.S15 Mean baseline NMR-measured ApoB by GRS for ApoB using SNPs from UKB 2020 by ancestry groups

Increase in ApoB per 1-SD increase in GRS by ancestry groups:
Ancestry group 1: 0.17±0.01 SD (0.03±0.002 g/L)
Ancestry group 2: 0.14±0.01 SD (0.03±0.002 g/L)
Ancestry group 3: 0.12±0.01 SD (0.02±0.002 g/L)



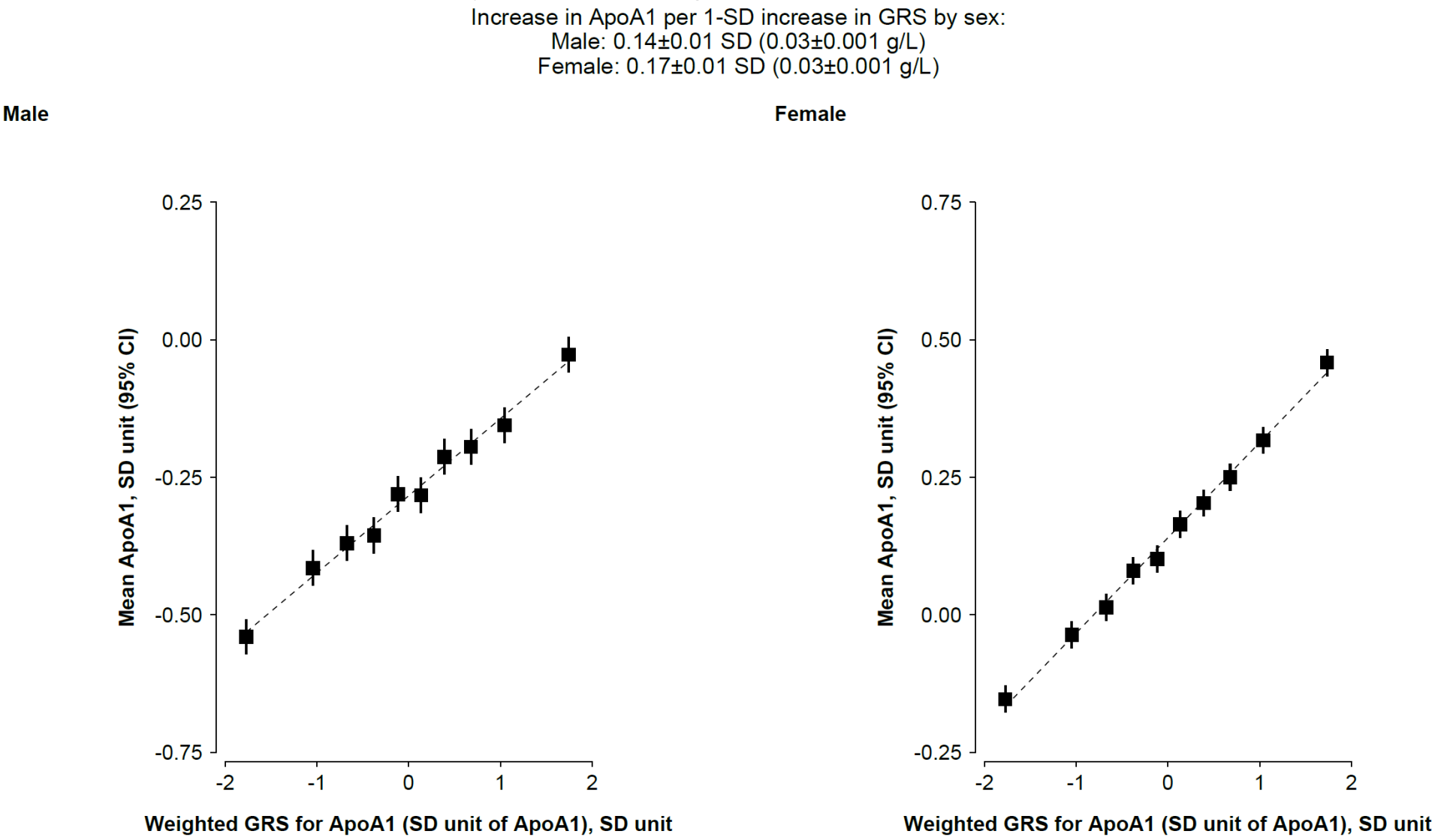
Adjusted for age, sex, and study region. Squares represent mean LDL-C level with area inversely proportional to the variance.
Vertical lines represent the corresponding 95% confidence intervals.
The three groups are defined by the tertile of the Native American ancestry score.

Figure 7.S16 Mean baseline NMR-measured ApoB by GRS for ApoB using SNPs from UKB 2020 by study region



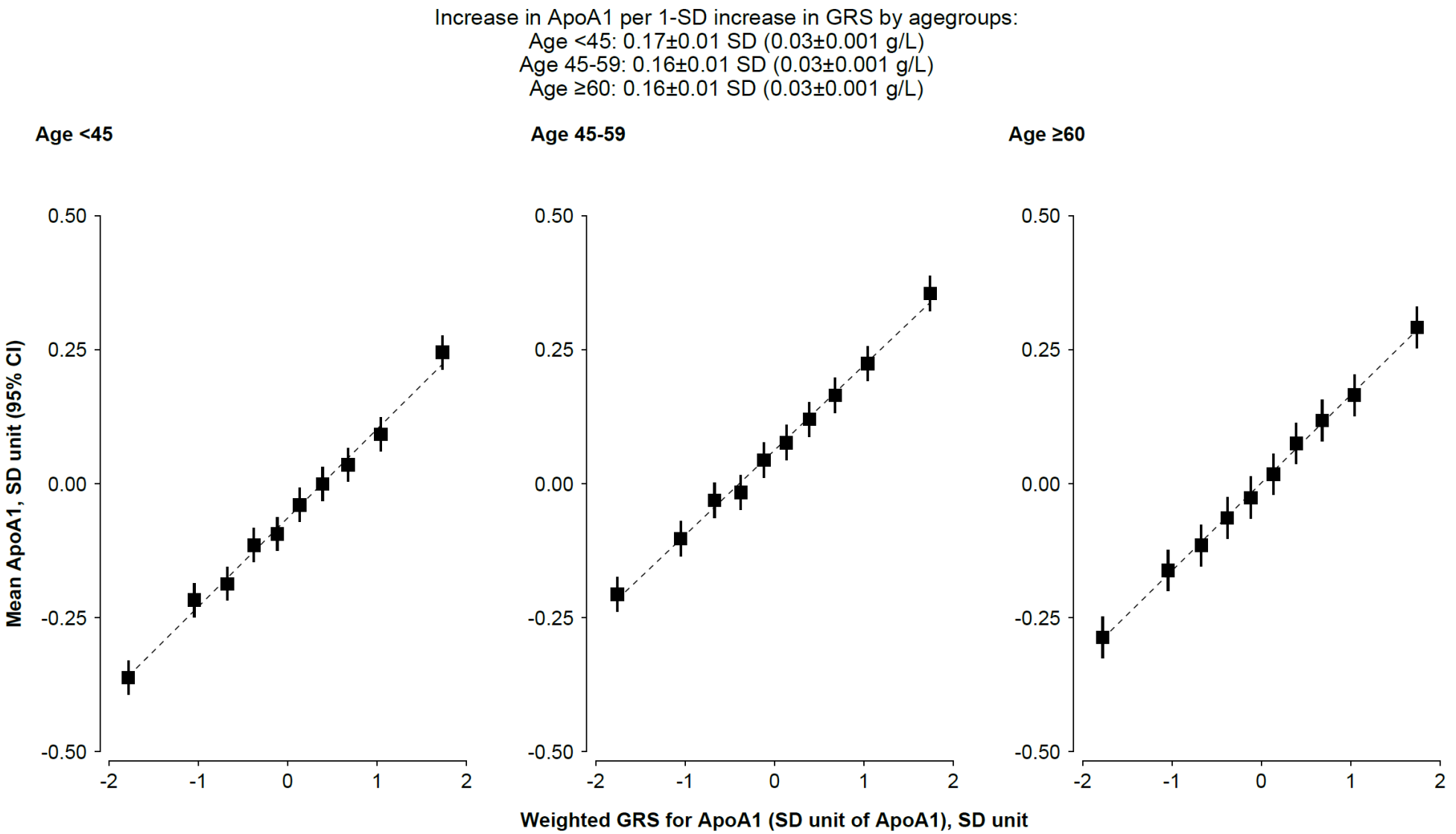
Adjusted for age, sex, and ancestry score. Squares represent mean ApoB level with area inversely proportional to the variance.
Vertical lines represent the corresponding 95% confidence intervals.

Figure 7.S17 Mean baseline NMR-measured ApoA1 by GRS for ApoA1 using SNPs from UKB 2020 in male and female



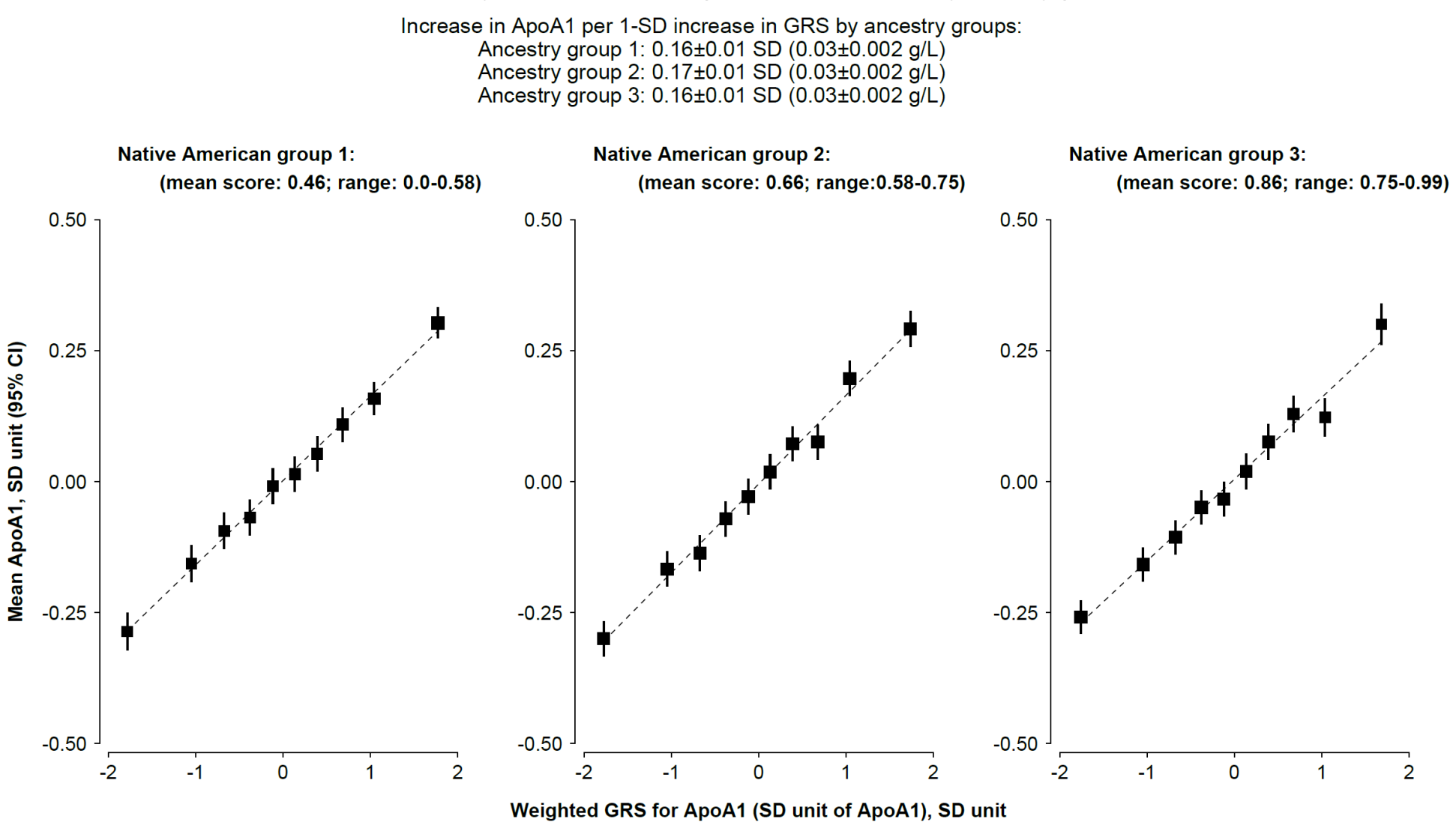
Adjusted for age, study region, and ancestry score. Squares represent mean ApoA1 level with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals.

Figure 7.S18 Mean baseline NMR-measured ApoA1 by GRS for ApoA1 using SNPs from UKB 2020 by age groups



Adjusted for sex, study region, and ancestry score. Squares represent mean ApoA1 level with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals.

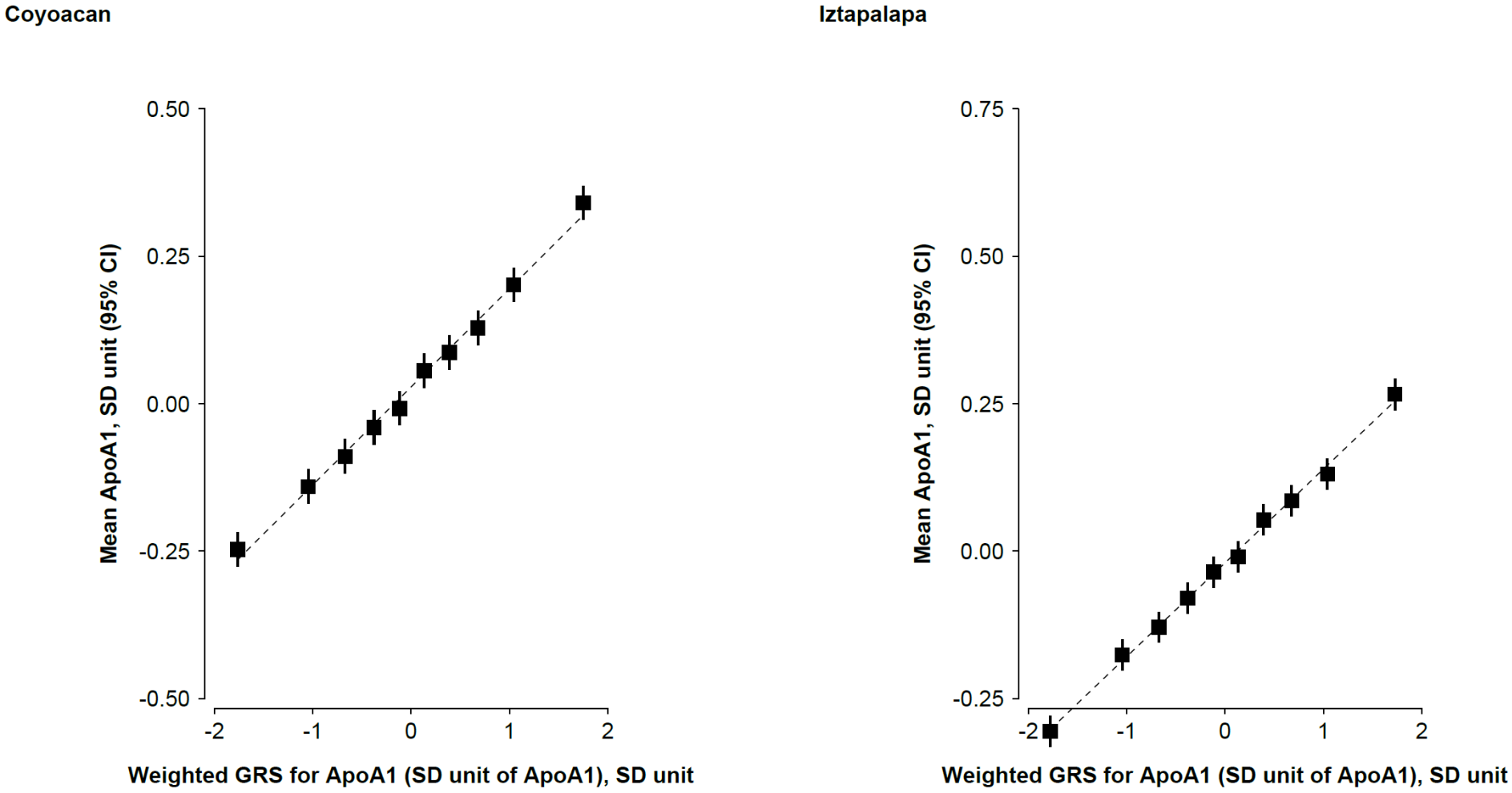
Figure 7.S19 Mean baseline NMR-measured ApoA1 by GRS for ApoA1 using SNPs from UKB 2020 by ancestry groups



Adjusted for age, sex, and study region. Squares represent mean LDL-C level with area inversely proportional to the variance.
Vertical lines represent the corresponding 95% confidence intervals.
The three groups are defined by the tertile of the Native American ancestry score.

Figure 7.S20 Mean baseline NMR-measured ApoA1 by GRS for ApoA1 using SNPs from UKB 2020 by study region

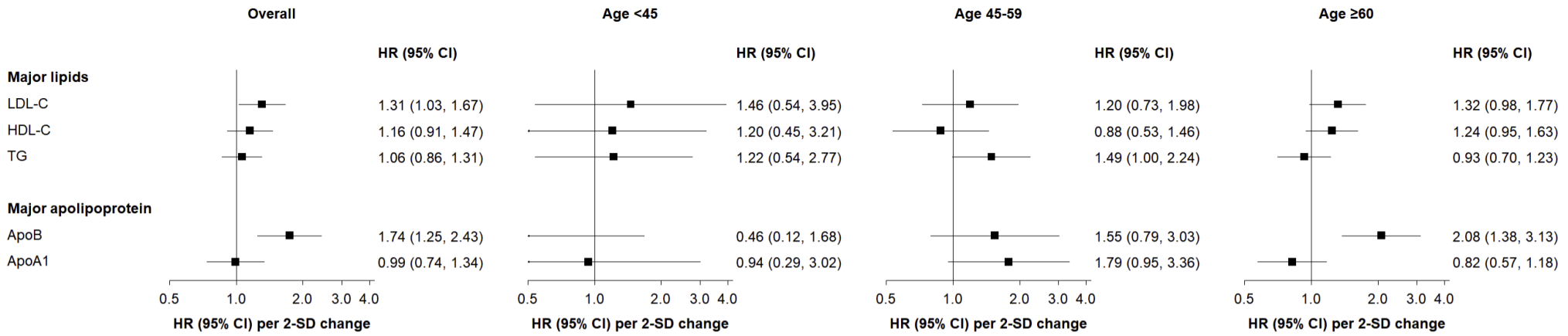
Increase in ApoA1 per 1-SD increase in GRS by study region:
Coyoacan: 0.17±0.01 SD (0.03±0.002 g/L)
Iztapalapa: 0.16±0.01 SD (0.03±0.002 g/L)



Adjusted for age, sex, and ancestry score. Squares represent mean ApoA1 level with area inversely proportional to the variance.
Vertical lines represent the corresponding 95% confidence intervals.

Figure 7.S21 One-sample univariable and multivariable MR of major lipids and apolipoproteins and risk of ASCVD by age groups

a



b

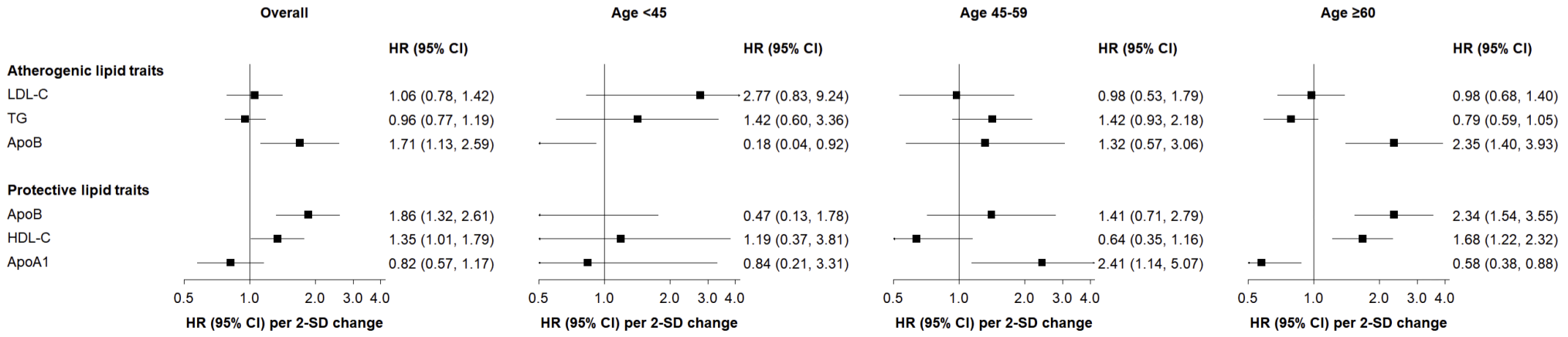
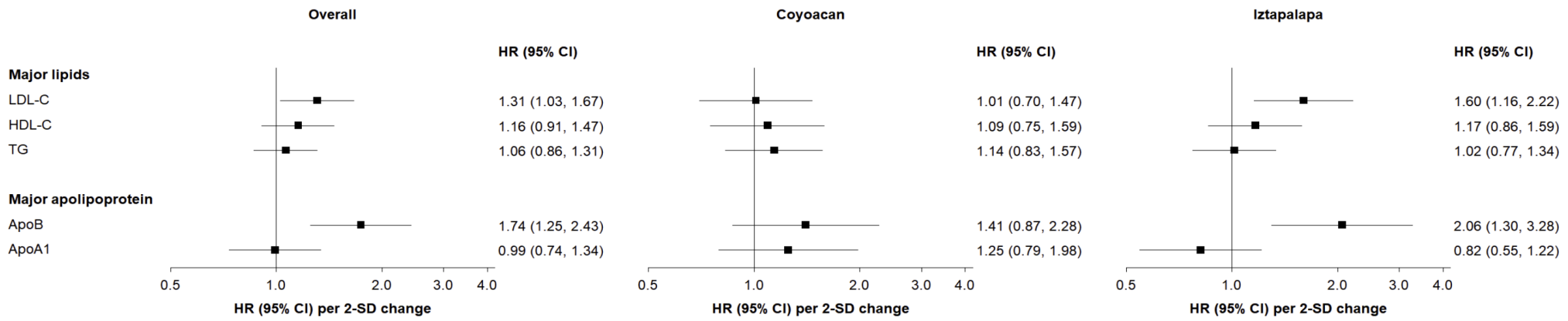


Figure 7.S22 One-sample univariable and multivariable MR of major lipids and apolipoproteins and risk of ASCVD by regions

a



b

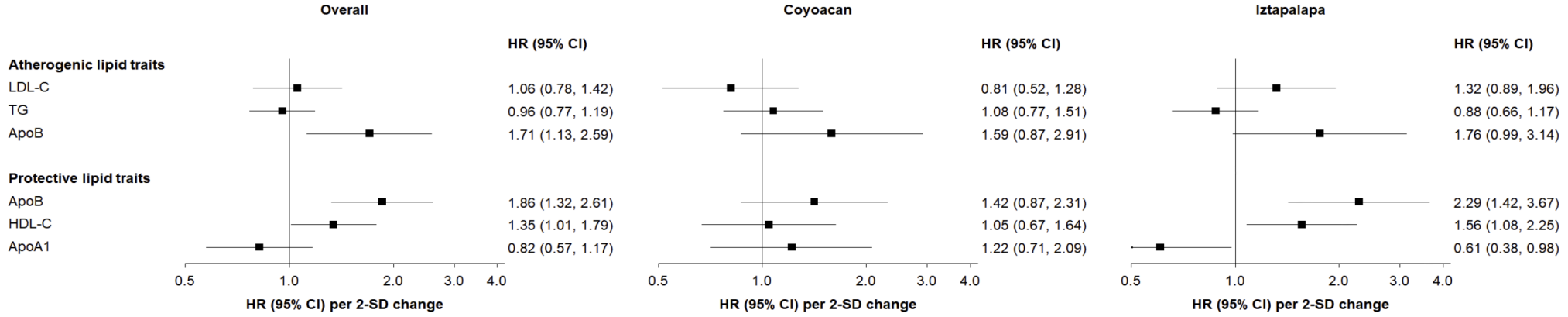
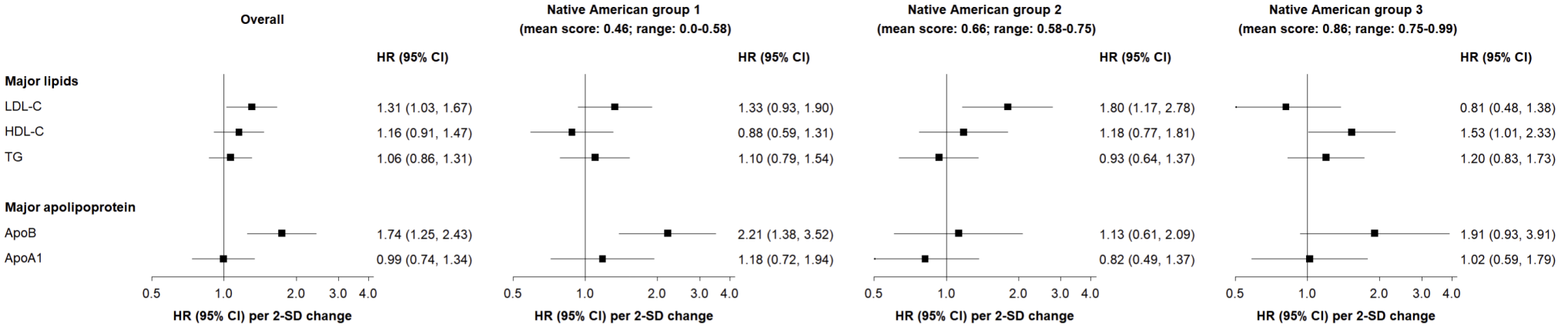


Figure 7.S23 One-sample univariable and multivariable MR of major lipids and apolipoproteins and risk of ASCVD by ancestry groups

a



b

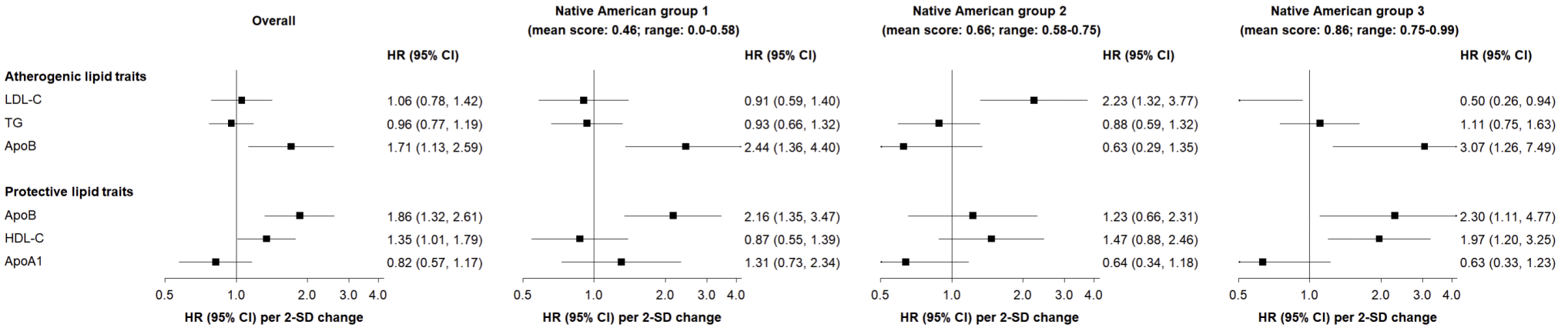
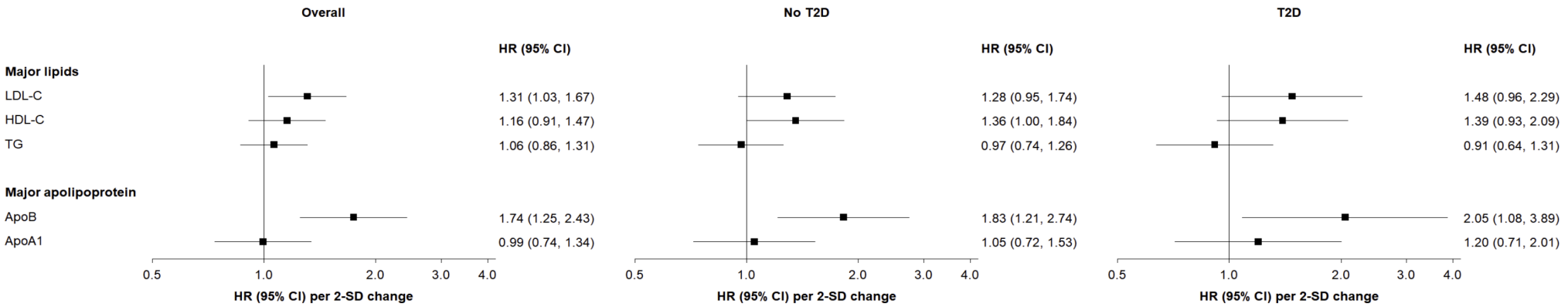


Figure 7.S24 One-sample univariable and multivariable MR of major lipids and apolipoproteins and risk of ASCVD by diabetes

a



b

