Metabolic risk and cardiovascular disease: insights from large biobanks with genetic epidemiological approaches

Ao, L.

CHAPTER 10

Appendices

SUMMARY

Although much progress has been made in cardiometabolic research over the last decades, at the start of this PhD project, many questions remained unanswered regarding the underlying pathophysiological mechanisms and the development of (more) effective personalized prevention strategies for cardiovascular disease (CVD). These endeavours are complicated because metabolic risk factors frequently overlap, interact, and vary between individuals and change over the life course. By applying a wide range of advanced epidemiological and analytical methods in genetic epidemiology, this thesis aimed to deepen our understanding of CVD risk from three perspectives: 1) the pathophysiological mechanisms of atherosclerosis (Part I); 2) the underlying genetics of CVD risk factors (Part II); and 3) the heterogeneity of CVD risk across populations with different metabolic characteristics (Part III).

Chapter 1 provided a brief introduction about CVD, and also introduced the overall background and aims of the studies conducted under these three perspectives, and described the main study populations used for the research included in this thesis.

Lipoprotein metabolism is highly dynamic and interrelated, which complicates causal inference and clinical interpretation when studying individual lipid components. However, compared to examining individual lipoproteins or constituents (i.e., low-density lipoprotein cholesterol [LDL-C], triglycerides), considering comprehensive lipoprotein profiles may provide additional insight in CVD risk. In **Chapter 2,** based on high-throughput 1-H nuclear magnetic resonance (^{1}H-NMR) metabolomic measures obtained from UK Biobank participants, we applied principal component analysis (PCA) to reduce 168 correlated metabolomic measures into six uncorrelated principal components (PCs) that together explained 88% of the variance. These PCs were then analysed using Cox proportional hazards models and Mendelian randomization (MR) to explore their potential causal association with CVD. The findings confirmed that PCs characterised by ApoB-containing lipoproteins were consistently associated with the risk of coronary artery disease (CAD). Notably, the study also observed that an ApoB-independent profile, characterized by a distinctive high-density lipoprotein (HDL) sub-particle distribution (decreased small HDL and increased large HDL), was associated with increased CAD risk. This study provided further evidence for a potentially important role for HDL function, rather than total HDL-C levels alone, in determining CAD risk.

Building on these findings, we observed variation in the phospholipid transfer protein *(PLTP*) encoding gene as one of the leading genetic variants responsible for this HDL-related risk profile. Given the recognised role of PLTP in lipoprotein metabolism but its unclear effects

on cardiovascular risk, in **Chapter 3**, we investigated the associations of genetically-influenced PLTP activity with ^{1}H-NMR metabolomic measures and CAD risk. Genetically-lower PLTP activity was found to be associated with increased HDL particle concentrations, smaller HDL size, and elevated triglyceride levels. However, no clear association with CAD risk was observed, possibly due to counterbalancing pro- and anti-atherogenic effects of PLTP activity. These findings highlight the pleiotropic effects of PLTP and the complex and interrelated nature of lipoprotein metabolism. Moreover, this study underscored the importance of testing potential causal associations and mechanistic hypotheses using robust methods before considering clinical intervention strategies.

In **Chapter 4** of this thesis, we investigated how genetic effects on well-known cardiovascular risk factors may vary with age, addressing an important but underexplored gap in genetic epidemiology. By conducting genome-wide interaction studies (GWIS) in over 270,000 unrelated European-ancestry participants from the UK Biobank, the study found that the majority of genetic effects on cardiometabolic risk factors remain relatively constant across adulthood, with the noted exceptions of specific genetic effects on ApoB and triglycerides. Specifically, significant variant-age interactions were found for four variants affecting ApoB (including well-characterised loci in *PCSK9*, *LDLR*, and near *APOB*) and for one variant, rs429358 tagging the well-known *APOE ε4* genetic variant, which showed interaction effects on both ApoB and triglyceride concentrations. Replication in the Estonian Biobank and Copenhagen General Population Study confirmed these findings. In conclusion, on the one hand, these findings together with prior studies suggest that most genetic effects on common cardiovascular risk factors do not change substantially over the life course, implying that the relative importance of genetic effects versus environmental influences could decrease with aging. However, on the other hand, the discovery of specific age-dependent variants highlights that future genetic risk prediction models and drug-target Mendelian randomization studies should account for potential age-specific genetic effects, especially for targeting pathways such as *PCSK9* and *LDLR*, which are already important clinical targets in lipid-lowering therapy.

Beyond the common genetic variants usually investigated in GWAS, in **Chapter 5,** we explored the effects of deoxyribonucleic acid (DNA) repeat sequences on sleep health, the latter being a recognized but underexplored cardiovascular risk factor. Based on the Netherlands Epidemiology of Obesity study (NEO) and the Netherlands Study of Depression and Anxiety (NESDA), we analysed the associations between cytosine-adenine-guanine (CAG) repeat sizes within the non-pathogenic range of three polyglutamine disease-associated genes (*HTT*, *ATXN3*, *CACNA1A*) and various sleep traits. This study identified 31 novel associations, which were differential by sex groups and depression status. Notably, CAG repeat sizes in all three genes were associated with sleep duration, with the longer allele

of *ATXN3* showing a nonlinear association with sleep duration across different population groups, except in men. Our findings also suggest that a considerable amount of variation in sleep traits could be explained by specific CAG repeats, which are not frequently studied. In the future, larger studies examining these CAG repeats in conjunction with single nucleotide polymorphisms would provide added value in exploring sleep health and its pathophysiology.

Unlike metabolomic and genomic markers, metabolic syndrome (MetS) is straightforward to measure in clinical practice and has long been recognised as a major risk factor for cardiovascular disease. However, given the intrinsic heterogeneity of the metabolic syndrome definition (>3 out of 5 symptoms), the dichotomous nature of MetS has been criticised for having limited usefulness in personalized cardiovascular risk prevention. In **Chapter 6**, we applied PCA on the five continuous MetS components in over 150,000 European-ancestry participants from the UK Biobank, and derived two uncorrelated risk profiles: one characterized by waist circumference and dyslipidaemia, and the other by hyperglycaemia. Using multivariable-adjusted Cox proportional hazard models, we observed that both risk profiles were associated with incident cardiometabolic diseases, but with effect sizes that differed by age and sex. This study emphasized the heterogenous presentation of MetS components, and the need for careful application of the dichotomous MetS definition when considering personalized prevention for cardiovascular risk.

Lipoprotein(a) [Lp(a)], which is predominantly genetically determined, has emerged in recent years as a critical risk factor for CVD. Although novel Lp(a)-lowering agents are in development, there are currently no approved drug therapies for lowering Lp(a). In addition, it remains unclear whether Lp(a)-lowering therapies provide similar benefits in all patients, whether the effect is also observed on clinical outcomes, and whether their effectiveness in CVD risk reduction depends on other common cardiovascular risk factors. In **Chapter 7,** based on data from the UK Biobank cohort, we investigated the interactions between Lp(a) and 12 common cardiovascular risk factors on CAD, calcific aortic valve stenosis (CAVS) and ischemic stroke (IS). This study confirmed the association between Lp(a) and CAD, CAVS and IS. Furthermore, we observed that Lp(a)-associated CAD risk particularly affected those having higher levels of total cholesterol, LDL-C or triglycerides. These findings suggest that the joint effects of high levels of Lp(a) and concurrent hyperlipidaemia should not be overlooked, and that controlling dyslipidaemia may be especially important for individuals with high Lp(a) levels.

To realize personalized cardiovascular risk prevention, in today's ageing society, it is particularly important to better understand the effects of ageing, especially the effects of ageing on subclinical symptoms that may be unnoticed and might develop into serious diseases over time. In **Chapter 8** of this thesis, we assessed whether levothyroxine therapy

improves lipid profiles in older adults with subclinical hypothyroidism, a population in which the benefits of treatment remain controversial. Particularly, subclinical hypothyroidism, which is biochemically characterised by elevated thyroid-stimulating hormone (TSH) concentrations in conjunction with thyroid hormone levels within the population reference range, is a common condition in older individuals. When corrected for multiple testing, pooled data from two randomized controlled trials (TRUST and IEMO80+) showed no evidence for significant treatment effects in the overall group. Although not statistically significant, directionally consistent effects on lipid lowering were observed for ApoB, total cholesterol, non-HDL-C, remnant cholesterol, LDL-C, and natural log-transformed TG (but not HDL-C). These directionally consistent effects reached nominal significance in the subgroup of participants with pretreatment TSH ≥ 10 mIU/L. Additional analyses with ^{1}H-NMR metabolomic measures further supported these observations. These findings align with current guidelines suggesting that the potential benefit of levothyroxine treatment in older adults with subclinical hypothyroidism should take into account baseline TSH levels, highlighting the importance of more tailored treatment decisions in this group.

Last, **Chapter 9** discussed the implications, limitations and future perspectives for the studies conducted in Chapter 2 through Chapter 8.