Novel insights into blood markers and cardiovascular disease: Results of the Netherlands Epidemiology of Obesity study

Christen, T.

SUMMARY AND GENERAL DISCUSSION

Summary of main findings

The aim of this thesis was to unravel a selection of a multitude of potential causal pathways that may underlie the association between excess body fat and cardiovascular disease, such as adipokines, inflammation, HDL-cholesterol and postprandial triglyceride response, and cholesteryl ester transfer protein (CETP).

In the first section of this thesis, we focused on factors that have only recently been regarded as risk factors for cardiovascular disease, like inflammation and adipokines. In the second section we focused on lipids and lipoproteins that were analysed using novel methods, such as metabolomics and Mendelian randomization. Here we summarize and interpret the findings of our studies, discuss their relevance in the field of cardiovascular risk factors, and describe the used methods, their application and potential pitfalls. In this thesis, we showed that body fat and its distribution are strongly related to serum concentrations of atherogenic lipoproteins, inflammatory markers such as C-reactive protein and glycoprotein acetyls, and the adipokines leptin and adiponectin. In addition, we showed that the association between body fat (distribution) and subclinical atherosclerosis was not mediated by inflammation as measured by C-reactive protein and glycoprotein acetyls. Also we showed that the adipokines leptin and adiponectin are unlikely to be causally related to subclinical measures of cardiovascular disease. Furthermore, we showed that the postprandial triglyceride response is not associated with subclinical atherosclerosis beyond fasting triglycerides. In addition, we found that CETP may have differential effects on subclinical atherosclerosis in different subgroups based on sex and cardiometabolic risk factors, while there is no effect in the total population.

In the middle of the 20th century the interrelated risk factors smoking, hypertension and LDL-cholesterol were the strongest contributors to cardiovascular disease. [13, 255] Despite reasonably successful efforts to reduce the burden of these factors by medication and life-style interventions, the decline in incidence rate of cardiovascular disease has slowed with the emergence of (childhood) obesity and diabetes. [15, 256, 257] This implies that the relative contribution of causal mechanisms may have shifted from the more traditional risk factors LDL-cholesterol and hypertension towards the well-known risk factor diabetes and novel factors like adipokines, inflammation and lipids and lipoproteins other than LDL-cholesterol. In this thesis, we mainly focused on the novel risk factors.

Atherosclerosis was previously regarded as a purely lipid-caused disease, but lesions were found to also contain pro-inflammatory macrophages containing oxidized LDL-cholesterol, suggesting that inhibiting inflammation could reduce the rate of progression of atherosclerosis. [3] Therefore it is not surprising that

the latest development in this field was the publication of the results of the CANTOS trial that aimed to assess the effects of the anti-inflammatory drug canakinumab, which is a monoclonal antibody against interleukin-1β. [39] Indeed, treatment with canakinumab reduced the incidence of cardiovascular disease, which was offset by an increased mortality due to infectious diseases. However, the role of inflammation as an intermediate pathway between obesity and cardiovascular disease has been less studied. Therefore in **Chapter 2** we studied the promising target of inflammation as a potential mediating factor between body fat, in specific visceral fat, and carotid intima media thickness as a marker of subclinical atherosclerosis. We observed strong associations between overall and visceral body fat and markers of systemic inflammation in the general population, but these markers were not associated with intima media thickness as a marker of subclinical atherosclerosis in men and women. Due to the absence of the association between inflammation and intima media thickness, no mediation was observed. The absence of the association between inflammation and subclinical atherosclerosis was in contrast with results from previous studies, and mainly the CANTOS trial. A potential explanation may be that the NEO study population is relatively healthy as compared to other study populations, and therefore the extent of atherosclerosis was limited. Also, this study could not exclude the possibility that inflammation contributes to cardiovascular disease via effects on other factors, such as coagulation or endothelial wear. In addition, other markers than C-reactive protein and glycoprotein acetyls may be more specific for the inflammation of visceral adipose tissue.

The second class of potential novel intermediate risk factors between obesity and cardiovascular disease are adipokines. Adipokines are signal molecules that are predominantly produced by adipose tissue. [258, 259] While adiponectin is mainly produced by visceral adipose tissue and modulates insulin sensitivity and fat oxidation [260], leptin is mainly produced by subcutaneous adipose tissue and signals to the brain to regulate food intake. [123] In recent years, adiponectin and leptin have both been linked to detrimental as well as advantageous effects on cardiovascular disease outcomes, but studies doing so were of diverse quality. [30, 33, 91, 127, 144, 154, 157, 159, 261-263] Therefore, we performed three studies to investigate the effect of sex on the association between body fat and circulating adipokine concentrations, and to assess the effects of these adipokines on (sub)clinical cardiovascular disease outcomes, while adjusting for the collateral effects of body fat.

In **Chapter 3** we confirmed that women had higher adiponectin and leptin concentrations than men. Our results suggest that the sex difference in leptin concentrations can be fully explained by the difference in total body fat, whereas the sex difference in visceral fat did not completely explain differences in adi-

ponectin concentrations. Previous studies have suggested that the sex difference in adiponectin and leptin concentration exceeded the difference that could be expected due to body fat. [55, 97, 101] However, most studies did not take into account that white adipose tissue is the main producer of leptin, and in general only compared linear associations or partial correlation coefficients between BMI and leptin concentrations in women and men. Similarly, previous studies have used waist circumference as a proxy for visceral fat, which is the main determinant of adiponectin concentrations. [97] In addition, we showed that in the general population there are women with extreme leptin concentrations, without apparent symptoms except excess body fat. A genetic variant in a potential regulatory area of the *LEP* gene was overrepresented in these women.

In **Chapter 4** we studied the relation between (genetically-determined) leptin concentration and measures of subclinical cardiovascular disease in thee domains: heart function, ECG parameters, and atherosclerosis. Also, we investigated whether observed associations could be explained by confounding effects of total body fat. We also performed two-sample Mendelian randomisation studies for the effect of leptin on pulse wave velocity and coronary heart disease in summary statistics of the UK Biobank and CARDIoGRAMplusC4D. [62, 63, 134, 168, 264] Most associations between leptin and measures of subclinical cardiovascular disease were explained by potential confounding effects of total body fat, and we did not find different results in women and in men. Also, genetically-determined leptin was not related to any of the measures of cardiovascular disease in the NEO study and the UK Biobank. In addition, we found no evidence of a strong genetic effect of leptin on the risk of coronary heart disease. However, the width of confidence intervals of these estimations could indicate heterogeneity in these effects. Previous observational studies provided conflicting results regarding the direction of the effect of leptin on cardiovascular disease. Leptin was assumed to have beneficial mechanistic effects on subclinical markers of cardiometabolic health, but an overall detrimental effect was shown in meta-analysis of studies that were adjusted for BMI. [91] However, our two-sample Mendelian randomisation study indicated that it is unlikely that leptin has a strong detrimental effect on coronary heart disease. Altogether, any effects of leptin on cardiovascular may be of limited clinical relevance and heterogeneous.

In **Chapter 5**, we attempted to disentangle the adiponectin paradox, which is fuelled by contradicting studies that on the one hand associate adiponectin with non-causal beneficial effects on insulin sensitivity and diabetes, while on the other hand observational studies show that individuals with increased adiponectin are at higher risk to develop cardiovascular disease. [155, 157, 159, 170, 261] The contradiction may be explained by reverse causation or collider stratification bias. Collider stratification could have occurred because these observational

studies have been performed in individuals at high risk to develop cardiovascular disease. In observational analyses in the NEO study, we observed no association between adiponectin and MRI measures of heart function, also not in analyses stratified by sex. However, using two-sample Mendelian randomisation, we found that a genetic predisposition towards a higher NT-proBNP concentration, which is a major biomarker of deteriorated left heart function, affects adiponectin concentrations. This findings suggests that previous observations are possibly explained by reverse causation through increased NT-proBNP concentrations due to deteriorated heart function.

Section II of this thesis includes chapters that go deeper into the relations between lipid and lipoprotein profiles and subclinical atherosclerosis. Extensive efforts to improve measurement of lipid and lipoprotein subfractions by metabolomics, and to discover genetic variants that influence the lipoprotein landscape by GWAS have made it possible to look further into the mechanisms that determine lipoprotein classes, their response to a meal challenge, and their associations with atherosclerosis.

While large-scale MR studies and clinical trials with CETP inhibitors have shown that CETP does not affect the risk of cardiovascular disease through the increase of HDL-cholesterol concentrations, early genetic studies on CETP suggested effects on the risk of cardiovascular risk. [49, 176, 178] These early studies have often been performed in specific subgroups of the population, such as men at high risk of cardiovascular disease, or women with diabetes. The combination of these results suggests that CETP has heterogeneous effects on atherosclerosis and cardiovascular disease. Therefore, in **Chapter 6** we performed a MR study of CETP on subclinical atherosclerosis in subgroups of the general population using uniquely strong genetic instruments. We confirmed that in the general population of the NEO study, genetically-determined CETP was not associated with intima-media thickness as a measure of subclinical atherosclerosis. However, our results suggested differential effects in men and women, and in individuals at high and low cardiometabolic risk. These results are coherent with the findings of previous studies that reported observational and MR results in specific groups of individuals. This may suggest that CETP-inhibiting treatment could be effective in subgroups of the population, potentially through their LDL-cholesterol lowering effects.

In **Chapter 7**, we investigated the hypothesis that the postprandial triglyceride response may explain the strong predictive value of HDL-cholesterol for the risk of coronary artery disease using a multivariable two-sample Mendelian randomisation strategy. We investigated the effects of genetically-determined HDL-cholesterol, LDL-cholesterol, fasting triglycerides, and the residual post-

prandial triglycerides on coronary heart disease using data from the GLGC consortium, the NEO study and the CARDIoGRAMplusC4D consortium. Our results indicate that LDL-cholesterol and fasting triglycerides together explain the association between HDL-cholesterol and coronary artery disease. The genetic instrument used to assess explanation by residual postprandial triglycerides was too weak to make solid inferences about this mechanism.

Previous research suggested that increased postprandial serum triglyceride excursions are a stronger risk factor for the development of atherosclerosis and cardiovascular disease than fasting triglyceride concentrations because the vascular wall is exposed to high postprandial triglyceride concentrations during the majority of the day and night. However, the main limitation of previous studies that investigated the relation between postprandial hypertriglyceridemia and atherosclerosis is that fasting triglyceride concentrations have not been taken into account adequately. We showed in **Chapter 8** that the association between the postprandial triglyceride response, defined as the incremental area under the curve, and subclinical atherosclerosis was completely explained by fasting triglyceride concentration. This may be due to inadequate reflection of daily triglyceride exposure by our meal challenge, or due to the same mechanism underlying the regulation of fasting and postprandial triglyceride concentrations. Our results suggest that in subgroups of the population with a damaged or susceptible vascular wall, such as patients with diabetes or smokers, there may be an association between higher triglyceride response and increased subclinical atherosclerosis in addition to the association with fasting triglycerides. However, the sample size in our subgroups was too small to draw solid conclusions.

Causal inference in observational studies

In epidemiologic research, one can distinguish prediction and aetiology (i.e. causal research). For example, one can partly predict the risk of cardiovascular disease once a patients' HDL-cholesterol and C-reactive protein concentration is known, but an intervention to alter these factors will render futile as these are most likely to be mere predictors and not causal factors. In this thesis we aimed to estimate causal effects of several risk factors for atherosclerosis and cardiovascular disease. While there are two main limitations of using only observational studies for causal inference, we used several strategies to mitigate these limitations.

A first limitation may be that at this point in time, the Netherlands Epidemiology of Obesity (NEO) study includes only cross-sectional data. Therefore the temporal relation between the variables that are regarded as exposure and as outcomes cannot be determined. This leaves the possibility that the assumed out-

come in fact causes the assumed exposure. This phenomenon is referred to as reverse causation. In this thesis, we aimed to investigate associations in the NEO study that were unlikely to have a reverse causal relation, based on previous research. When possible, we added Mendelian randomisation analyses to reduce the risk of reverse causation.

Second, the NEO study was designed as an observational study. The exposures of interest were therefore not allocated by chance, as is the case in a randomised controlled trial, but may also be determined by other factors. Some of these factors may also be determinants of the outcomes of interest and therefore constitute confounding, provided they are not part of the causal mechanism underlying the relation between exposure and outcome (mediation). [265] To mitigate the effects of confounding factors, one can apply different strategies in the design and analysis of a study. In this thesis, we used restriction in the design of the study (**all Chapters**), stratification (**Chapters 3, 4, 6**), and statistical adjustment (**all Chapters**). Still, a main limitation of observational studies is that confounding factors can not all be measured, and if so, measured well. This poses observational studies at risk for residual confounding, which may lead to biased results. Therefore, in this thesis, Mendelian randomisation analyses are used when possible, to prevent problems with residual confounding.

Mendelian randomisation studies

In recent years, a novel methodology has been developed to mitigate the aforementioned problems of reverse causation and the potential bias as a result of residual confounding: Mendelian randomisation analysis (after Georg Mendel, the founder of modern genetics). This method makes use of the attribute of genetic variants that they are assigned randomly at conception, and inherited independently from other genetic variants.

When a genetic variant is a determinant of an exposure of interest, one can use this information to infer a causal relation between an exposure and outcome of interest, given the validity of three assumptions. [163] First, the genetic variant should be associated with the exposure of interest. Second, the genetic variant is independent of other determinants of the outcome of interest, and third, the genetic variant is only associated with the outcome of interest through the exposure of interest. The first assumption can and should be checked within the data of the study using a simple regression analysis. The second assumption may be explored, but not tested, by investigating the effects of the genetic variants on known potential confounding factors of the association between the exposure and outcome of interest. Testing the third assumption is also infeasible, because it is often impossible to distinguish whether the genetic variant influences the

outcome via the exposure of interest or via a parallel pathway. These assumptions are under pressure from several potential mechanisms, of which pleiotropy, population stratification, and linkage have the largest potential impact. [266]

Pleiotropy is the phenomenon in which a genetic variant not only affects one, but multiple factors. This may pose a problem for a successful Mendelian randomisation analysis, as the effect on the outcome cannot be assigned to one of the factors for which the genetic variant predisposes. This may violate the second assumption for Mendelian randomisation analysis, when the genetic variant affects the factor that is not of interest directly as opposed to through the factor of interest. A potential way to circumvent this problem is to perform Mendelian randomisation with multiple genetic variants for multiple risk factors, to estimate the separate effects of these risk factors instrumented by genetic risk scores on the outcome of interest. [58] This is specifically useful in situations where multiple risk factors are themselves highly intertwined, such as lipoprotein species. Therefore we aimed to perform a multivariable Mendelian randomisation analysis in **Chapter 7**.

Potentially, a genetic variant is not distributed completely at random over the population, but influenced by other factors like ancestry. This may constitute one condition of potential confounding. Often, ancestry is in some way related to the outcome of interest. This may violate the third assumption for Mendelian randomisation analysis, and lead to confounding that needs to be adjusted for. Therefore genetic analyses should be adjusted for population stratification. In our analyses, we performed our analyses adjusted for four markers of ancestry to avoid confounding by population stratification. Also, genetic variants are assumed to be inherited independently of each other. However, often genetic variants inherit with other variants more often than expected by chance, which is described by the term linkage. This may result in a form of pleiotropy, as several treats may be affected at the same time as a consequence of linked genetic variants. This would violate the second assumption for Mendelian randomisation analysis. Alternatively, linkage may be used to one's advantage when the genetic variant of interest has not been measured, but can be estimated using a linked variant. This is common practice in Mendelian randomisation analyses, as different genome analysis kits include different sets of genetic variants. Also, for many phenotypes no genetic instruments are known, because the phenotype is too difficult to measure or has a too low prevalence to perform a genome-wide association study. This limits the application of Mendelian randomisation studies in terms of rare causes of disease. For other phenotypes only weak instruments are known. The use of weak instruments requires large sample sizes. Therefore, a major strength of this thesis is that in **chapters 4, 5, and 7** we made use of publicly-available summary statistics of several large genome-wide association

studies to look up effects of genetic variants that predispose for our exposures of interest. One drawback of the use of these summary statistics is that it limits the possibilities of performing stratified analyses, e.g. to investigate effect measure modification. Also, the assumptions regarding pleiotropy and independence of the genetic variants or risk scores with potential confounders could not be tested more extensively than customary in Mendelian randomisation studies.

Mediation analyses

In **Chapter 2** a mediation study of the association between body fat and subclinical atherosclerosis is presented. Mediation is a concept in causal inference indicating that the relation between a cause and an outcome is completely or partly actuated through a mediator. [84] The use of directed acyclic graphs such as in Chapter 2, Figure 1, is a straightforward way to visualise the concept of mediation.

The Baron and Kenny method is a straightforward way of mediation analysis. Using this method, several assumptions should hold. There should be associations between exposure and mediator, and between mediator and outcome. Also, interaction between exposure and mediator should be absent. Assuming continuous and normally distributed exposure, mediator and outcome, this method consist of a linear regression analysis between exposure and outcome, performed without and with additional adjustment for the mediator. The difference between the regression coefficients represents the proportion of the association mediated by the mediator. The main limitations of this method are the sensitivity to bias due to measurement error, and strong assumptions regarding confounding. [85, 267] In most situations it is not possible to test these assumptions, but it is possible to perform sensitivity analyses using bias formulas. [85] While the Baron and Kenny method requires only simple analyses, it may give biased results in the case of interaction between exposure and mediator. More advanced methods have been developed to investigate interactions between exposure and mediator in their effects on the outcome. The most prominent models are the structural equation models by Vanderweele. [268, 269] As we did not observe an association between the mediators and the outcome in **Chapter 2**, we only used the Baron and Kenny method of mediation analysis to confirm the absence of mediation. Independent of which method is used to study mediation, it is essential to carefully consider strategies to fulfil assumptions regarding confounding and measurement error.

Clinical implications and future perspectives

A major challenge in current clinical practice is the early identification of individuals at a high risk of cardiovascular disease. This would allow for improved treatment strategies based on the characteristics of the patient, indicating the most intensive (preventive) treatment in individuals at the highest risk, while individuals with a lower risk could be treated less intensively. This strategy would reduce side effects and lower costs. To be able to identify individuals at the highest risk, studies should be performed that investigate the combined effect of risk factors. The recent setup of several large population-based cohorts and our ability to extensively genotype and phenotype participants are major opportunities to improve our risk stratification. The pursuit of extensions of such risk categorization models fits into the trend towards personalized medicine. Personalized medicine has a multitude of challenges, including but not limited to ethical, legal, practical, financial and social. However, these challenges are often reduced to a consideration of efficacy, effectiveness, and (cost-)efficiency. Major limiting factors are the economic capacity of a society and the willingness to allocate a substantial part of its budget to health. This may challenge researchers in exploring fields with a smaller chance of success, but might also lead to creative strategies and collaborations.

In this thesis we focused on the risk stratification within a study that included participants who were already at a higher risk of cardiovascular disease due to their high body mass index. In this thesis we observed heterogeneity in the associations between the different risk factors and measures of subclinical cardiovascular disease. This indicates the presence of unknown underlying mechanisms and suggests that effect modification may be present. While there are many potential explanations for heterogeneity in results of large studies, in this thesis we attempted to investigate subgroup differences as a source of variation. Therefore, when sample size allowed this, analyses were performed separately for women and men or for subgroups of the population with different baseline risk factors. Our results suggested that different risk factors may contribute to measures of subclinical cardiovascular disease to a different extent in women than in men. For example, in the Mendelian randomisation analysis of the effect of CETP on subclinical atherosclerosis the point estimates for women and men were in the opposite direction. While in large clinical trials CETP inhibition did not affect the risk of cardiovascular disease, our results were consistent with previous studies that detected effects of CETP on the risk of cardiovascular disease only in men or in women with diabetes. In addition, the cardiovascular effects of CETP inhibition are likely to differ by the genotype of the rs1967309 variant (*ADCY9*), with the AA genotype benefiting the most from treatment with dalcetrapib, and the

GG genotype being harmed by treatment, while heterozygous individuals had a slightly reduced risk of cardiovascular disease. [270] Analogously, the associations that we observed between fasting and postprandial triglyceride concentrations, and subclinical atherosclerosis may be more prominent in individuals at high risk of cardiovascular disease. This may indicate an interaction between risk factors and suggest that it may be useful to screen for additional risk factors in individuals at high risk for cardiovascular disease.

Extrapolating, we suggest that more studies should include investigations into sex differences and subgroup effects in the pathophysiology and treatment of cardiovascular disease.

However, we recognise that this will come at the cost of statistical power, and indeed larger sample sizes are needed to reliably draw conclusions about subgroup effects. This limitation will be amplified when subgroup analyses are employed in Mendelian randomisation studies, as by using genetic instruments power generally decreases. In the future, more large cohort studies, such as the UK Biobank, will be able to perform well-powered analyses using clinical and genetic data. Also, to investigate relations between the uniquely measured risk factors in the NEO study and clinical manifestations of disease, the use of follow-up data from this population is crucial. Also, future studies may investigate other underlying mechanisms that may mediate the relation between overweight and cardiovascular disease, such as sympathetic nervous system activation, coagulation and endothelial wear.

In conclusion, in this thesis we elucidated various pathways that may connect body fat distribution and cardiovascular disease. In **Section I** we showed that hs-CRP and GlycA as measures of inflammation, adiponectin, and leptin are not associated with clinical and subclinical cardiovascular disease in the general population. However, all may be relevant markers of disease risk. In addition, in the **Section II** we showed that postprandial triglyceride excursions, genetically-determined CETP and HDL-cholesterol, while not related with subclinical atherosclerosis in the general population, may be interesting targets to pursue in women and men separately, and in subgroups of individuals at high-cardiovascular risk. Therefore, there is still a remaining part of the association between overweight and cardiovascular disease that needs to be elucidated. Studying the interplay between different risk factors seems indispensable to advance our knowledge of the development of cardiovascular disease.