

Genetic determinants of cholesterol and energy metabolism : implications for cardiometabolic health Blauw, L.L.

Citation

Blauw, L. L. (2018, September 20). *Genetic determinants of cholesterol and energy metabolism : implications for cardiometabolic health*. Retrieved from https://hdl.handle.net/1887/65600

Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

The handle <http://hdl.handle.net/1887/65600> holds various files of this Leiden University dissertation.

Author: Blauw, L.L. **Title**: Genetic determinants of cholesterol and energy metabolism : Implications for cardiometabolic health **Issue Date**: 2018-09-20

Circulating CETP causally decreases large HDL and increases small VLDL without affecting LDL

Lisanne L. Blauw, Raymond Noordam, Sebastian Soidinsalo, C. Alexander Blauw, Ruifang Li-Gao, Renée de Mutsert, Jimmy F.P. Berbée, Yanan Wang, Diana van Heemst, Frits R. Rosendaal, J. Wouter Jukema, Dennis O. Mook-Kanamori, Peter Würtz, Ko Willems van Dijk, Patrick C.N. Rensen

Submitted

Abstract

Introduction

According to current dogma, cholesteryl ester transfer protein (CETP) decreases high-density lipoprotein (HDL)-cholesterol (C) and increases low-density lipoprotein (LDL)-C. HDL-C is measured directly, whereas LDL-C is commonly inferred from the Friedewald formula without direct measurement of specific lipoprotein subclasses. In depth insight in the causal effects of CETP on the lipoprotein profile may assist in understanding outcomes of recent CETP inhibitor trials. Therefore, we aimed to assess associations between CETP and the detailed lipoprotein profile, using Mendelian randomization and observational analyses.

Methods

Using nuclear magnetic resonance spectroscopy, we quantified 159 metabolic measures in 5,672 participants of the Netherlands Epidemiology of Obesity (NEO) study. We used three previously identified SNPs (rs247616, rs12720922 and rs1968905) that determine CETP concentrations as genetic instruments to estimate the causal effects on the metabolic measures, and reported effect estimates as the Ln-transformed SD difference in metabolic measure per unit $(\mu q/mL)$ increase in CETP. Findings were replicated in an independent dataset (MAGNETIC NMR GWAS; n≈20,000). Additionally, observational associations between ELISA-measured CETP concentration and the metabolic measures were assessed.

Results

Higher CETP concentrations were casually associated with less large HDL components (largest effect XL-HDL-C, $P=6\times10^{-22}$), and with more small VLDL components (largest effect S-VLDL cholesteryl esters, $P=6\times10^{-6}$), while no association was observed with any LDL subclasses or constituents. This pattern of lipidomic associations was replicated. In contrast, observationally, measured CETP concentration predominantly associated with more VLDL, IDL and LDL components.

Conclusion

Based on Mendelian randomization, we conclude that CETP is an important determinant of HDL and VLDL concentration and composition. Since HDL does not causally affect cardiovascular disease risk, we speculate that the CETP inhibitor anacetrapib decreased cardiovascular disease risk through specific reduction of small VLDL rather than LDL. We propose that the contrast between genetic and observational associations may be explained by the higher capacity of VLDL, IDL and LDL subclasses to carry CETP, thereby concealing causal associations with HDL.

Introduction

Cholesteryl ester transfer protein (CETP) is able to transfer cholesteryl esters from highdensity lipoproteins (HDL) to apolipoprotein (Apo) B containing triglyceride-rich lipoproteins, mainly very-low-density lipoproteins (VLDL). In exchange, triglycerides are transferred from VLDL to triglyceride-poor particles, which are both HDL and low-density lipoprotein (LDL).^[1] As such, CETP facilitates an atherogenic lipoprotein profile, as has been extensively studied in both humans and in mice transgenic for human CETP.^[2,3]

Despite the promising results of preclinical studies, ^[4,5] the clinical trials with the initial CETP inhibitors torcetrapib, dalcetrapib, and evacetrapib, were terminated: torcetrapib had offtarget effects on blood pressure and caused an increase in cardiovascular events, [6] and both dalcetrapib and evacetrapib lacked efficacy in reducing cardiovascular events on top of statin therapy. $[7,8]$ All these CETP inhibitors caused a large increase in HDL-C, accompanied by a modest or no decrease in LDL-C. Although a high HDL-C concentration was previously proposed to decrease the risk of cardiovascular disease (CVD) based on observational studies, ^[9] Mendelian randomization showed that higher HDL-C concentrations do not lower the risk of myocardial infarction, $[10-13]$ which implies that the association between HDL-C and CVD is not causal. This may provide one of the explanations for the lack of efficacy of the three initial CETP inhibitors. $[6-8]$

Interestingly, the fourth clinical trial with the CETP inhibitor anacetrapib (REVEAL) did meet its primary endpoint by showing a 9% relative risk reduction in major coronary events.^[14] The reduction in coronary events by anacetrapib, albeit limited, was attributed to a reduction in LDL-C and a corresponding reduction in ApoB-containing lipoprotein particles.^[14,15] However, the specific lipoprotein subclasses affected by CETP have not been determined. Specific lipoprotein fractions have previously been associated with increased CVD risk, [^{16]} and more in-depth insight in the causal effects of CETP on the circulating lipoprotein profile may therefore assist in understanding CETP inhibitor trial outcomes. This is also of specific importance in the light of the dal-GenE randomized phase III clinical trial (NCT number: NCT02525939), which is currently being performed to confirm whether dalcetrapib is able to reduce CVD risk in patients with the AA genotype of the rs1967309 polymorphism in the *ADCY9* gene.[7,17]

We recently performed a genome-wide association study (GWAS) on serum CETP concentration in the Netherlands Epidemiology of Obesity (NEO) study, and showed that CETP concentration has a strong genetic component.^[18] Notably, three independent single nucleotide polymorphisms (SNPs), all mapped to the *CETP* gene, together explained 16.4% of the total variation in serum CETP concentration. With the use of these SNPs as genetic instruments in Mendelian randomization, the causal effects of circulating CETP on lipoprotein subclasses can be determined. Besides, by using a Mendelian randomization approach, the effects of these genetic variants can mimic CETP inhibition, and we may infer about the mechanisms underlying the outcomes of CETP inhibitor trials.^[19,20]

In the present study, we aimed to assess the causal effects of CETP concentration on 159 circulating metabolic measures, primarily lipoprotein subclasses, using a Mendelian randomization approach in a cohort of the Dutch general population.^[21] We also compared the causal effect estimates with observational associations between serum CETP concentration and these measures of lipid metabolism.

Materials and Methods

Study design and population

The present study is embedded in the Netherlands Epidemiology of Obesity (NEO) study, a population-based prospective cohort study of men and women aged 45 to 65 years. From the greater area of Leiden, The Netherlands, all inhabitants with a self-reported body mass index (BMI) of 27 kg/ $m²$ or higher were eligible to participate. In addition, inhabitants from one nearby municipality (Leiderdorp, The Netherlands) in the same age group were invited to participate regardless of their BMI, forming a reference population for BMI distribution. In total, 6,671 participants were included from September 2008 until September 2012. Participants visited the NEO study center for extensive physical examination. Venous blood samples were obtained from the antecubital vein after a 10 hour overnight fast. Research nurses recorded current medication use by means of a medication inventory. Prior to the study visit, participants completed questionnaires at home with respect to demographic, lifestyle, and clinical information. For the present study, we excluded participants with missing data on serum CETP concentration, metabolic profiling or genotype. Therefore, the present study population consists of 5,672 individuals. The NEO study was approved by the medical ethics committee of the Leiden University Medical Center (LUMC), and all participants gave their written informed consent. Detailed information about the study design and data collection has been described elsewhere.^[22]

Genotyping and imputation

DNA was isolated from venous blood samples. Genotyping was performed in participants form European ancestry, using the Illumina HumanCoreExome-24 BeadChip (Illumina Inc., San Diego, California, United States of America). Subsequently, genotypes were imputed to the 1000 Genome Project reference panel (v3 2011)^[23] using IMPUTE $(v2.2)$ software.^[24] From the whole-genome data, we extracted the three independent genetic variants that have been previously identified in relationship to CETP concentration in the NEO study population, $[18]$ notably rs247616 (directly genotyped; coding allele (C) frequency 0.67), rs12720922 (imputation quality 0.98; coding allele frequency (A) 0.17) and rs1968905 (imputation quality 0.85; coding allele frequency (G) 0.82). The CETPincreasing alleles are rs247616-C, rs12720922-A and rs1968905-G. Based on these three polymorphisms, we calculated a weighted genetic score per participant. The genetic score was constructed as the sum of the number of CETP-increasing alleles weighted by their effect size on CETP concentration, as previously described.^[18]

Serum CETP concentration and routine-lipid profile

After centrifugation, aliquots of serum were stored at -80°C. From 11 April until 16 July 2014 CETP concentrations were measured with enzyme-linked immune sorbent assay (ELISA) kits according to the manufacturer's instructions (DAIICHI CETP ELISA, Alpco, Salem, USA; coefficient of variation (CV) 11.7%) in serum that had undergone one previous freeze-thaw cycle. Fasting serum total cholesterol and triglycerides concentrations were measured with enzymatic colorimetric assays (Roche Modular P800 Analyzer, Roche Diagnostics, Mannheim, Germany) and fasting serum HDL-C concentrations with third generation homogenous HDL-C methods (Roche Modular P800 Analyzer, Roche Diagnostics, Mannheim, Germany). Fasting LDL-C concentrations were calculated using the Friedewald equation.^[2]

NMR-based metabolic biomarker profiling

A high-throughput proton NMR metabolomics platform^[21] (Nightingale Health Ltd., Helsinki, Finland) was used to quantify 159 lipid and metabolite measures. The NMR spectroscopy was conducted at the Medical Research Council Integrative Epidemiology Unit (MRC IEU) at the University of Bristol, Bristol, United Kingdom, and processed by Nightingale's biomarker quantification algorithms (version 2014). This method provides quantification of lipoprotein subclass profiling with lipid concentrations within 14 lipoprotein subclasses. The 14 subclass sizes were defined as follows: extremely large VLDL with particle diameters from 75 nm upwards and a possible contribution of chylomicrons, five VLDL subclasses (average particle diameters of 64.0 nm, 53.6 nm, 44.5 nm, 36.8 nm, and 31.3 nm), IDL (28.6 nm), three LDL subclasses (25.5 nm, 23.0 nm, and 18.7 nm), and four HDL subclasses (14.3 nm, 12.1 nm, 10.9 nm, and 8.7 nm). Within the lipoprotein subclasses the following components were quantified: total cholesterol, total lipids, phospholipids, free cholesterol, cholesteryl esters, and triglycerides. The mean size for VLDL, LDL and HDL particles was calculated by weighting the corresponding subclass diameters with their particle concentrations. Furthermore, 58 metabolic measures were determined that belong to

classes of apolipoproteins, cholesterol, fatty acids, glycerides, phospholipids, amino acids, fluid balance, glycolysis-related metabolites, inflammation, and ketone bodies. Details of the experimentation and applications of the NMR metabolomics platform have been described previously, $[21]$ as well as CVs for the metabolic biomarkers. $[25]$ A full list of the measured biomarkers is included in Supplementary table 6.C.1.

Statistical analyses

For all analyses, metabolic measures were natural-log transformed to obtain normal distributions. For comparison of the strengths of the associations between the different metabolic measures, standardized z-scores were composed. Consequently, the outcome variable of all analyses was the natural-log transformed SD difference in metabolic measure. When the concentration of a metabolic measure for an individual was below the detection limit, a value of half of the minimum concentration of that metabolic measure in the total population was imputed.

First, we performed a Mendelian randomization analysis to determine the causal associations between serum CETP concentration and the 159 metabolic measures. The genetic score was used as determinant in a linear regression analysis to assess the effect of a one unit $(\mu q/mL)$ increase in serum CETP concentration on the metabolic measures. Second, the observational associations between serum CETP concentration and the 159 metabolic measures were determined with linear regression analyses. Participants with a serum CETP concentration beyond four SD from the mean were excluded (n=1). All linear regression analyses were adjusted for age and sex. Beta coefficients, SE and P-values from linear regression analyses were reported. In addition, a linear regression analysis adjusted for age and sex was performed to determine the associations of serum CETP concentration with routinely measured HDL-C concentration and calculated Friedewald LDL-C concentration.

The metabolic biomarkers used for the present study are correlated with each other, and therefore, conventional correction for multiple testing (e.g. Bonferroni) is too stringent. To obtain the number of independent metabolic biomarkers, we used the method as described by Li et al.,^[26] which takes the correlation between the different metabolic biomarkers into account. Based on this method, we found 37 independent metabolic markers. For this reason, associations were considered to be statistically significant in case the P-value was below 0.00134 (i.e. 0.05/37).

All results were based on analyses weighted towards the reference BMI distribution of the general Dutch population, and therefore apply to a population-based study without oversampling of individuals with overweight or obesity (see Supplement 6.A). Analyses

were performed using STATA Statistical Software version 12.0 (Statacorp, College Station, Texas, USA) and R version 3.4.0 (The R Project, https://www.r-project.org/). Figures were designed with Python version 2.7.6 (Python Software Foundation, https://www.python.org/).

Replication

We aimed to replicate the findings from the Mendelian randomization analyses in the NEO study in an independent population. For that purpose, we used publically-available summary statistics from the MAGNETIC NMR GWAS dataset, ^[25] which comprises the additive (per-allele) beta coefficients with accompanying standard errors of the associations between genome-wide SNPs and 123 metabolic measures, of which 111 overlapped with the 159 metabolic measures that were quantified in the NEO study. This GWAS meta-analysis included data of approximately 20,000 individuals from 14 datasets that were derived from cohorts of European ancestry. The 123 metabolic measures were quantified by a prior version of the same high-throughput proton NMR metabolomics platform^[21] (Nightingale Health Ltd., Helsinki, Finland) as used in the NEO study. The 123 metabolic measures were quantified by a prior version of the same high-throughput proton NMR metabolomics platform^[21] (Nightingale Health Ltd., Helsinki, Finland) as used in the NEO study. In these metabolomics data, 22 principal components were identified, ^[25] and therefore associations were considered to be statistically significant in case the P-value was below 0.00227 (i.e. 0.05/22).

For the Mendelian randomization analyses, we combined the individual genetic variants for CETP concentration to estimate the causal effect of CETP concentration on the metabolic measures. Analogous to pooling estimates from different observational studies with conventional meta-analysis using inverse-variance weighting, we weighted this combined effect estimate of the *CETP* SNPs on the metabolic measures by the inverse of the variance for each individual additive (per-allele) effect on the metabolic measures, and incorporated the individuals additive effects of the genetic instruments on CETP concentration. Additionally, we determined the correlation (R2) between the effect estimates from the NEO study and the MAGNETIC NMR GWAS dataset. A strong correlation indicates high consistency in the overall association profile of higher CETP concentrations with the lipid and metabolite measures.

Results

Population characteristics

Characteristics of the total study population are summarized in Table 6.1. The mean (SD) age was 56 (6) years. Mean (SD) concentration of CETP was 2.47 (0.65) μ g/mL, of LDL-C 3.56 (0.96) mmol/L, and of and HDL-C 1.57 (0.46) mmol/L.

Table 6.1: Characteristics of the discovery and replication cohort from the Netherlands Epidemiology of Obesity (NEO) study.

n=5,672^b. Results are based on analyses weighted towards the reference BMI distribution of the general Dutch population, and presented as mean (SD) or percentage.

^b Missing data: n=4 for ethnicity, n=56 for educational level, n=4 for smoking, n=21 for cardiovascular disease, n=8 for total cholesterol concentration, n=10 for LDL-cholesterol concentration, n=9 for HDL-cholesterol concentration and n=9 for triglycerides concentration.

CETP, cholesteryl ester transfer protein; HDL, high-density lipoprotein; LDL, low-density lipoprotein.

Causally, CETP is negatively associated with large HDL components and positively associated with small VLDL components

Figure 6.1 and Supplementary table 6.C.2 show the results from the Mendelian randomization analyses of CETP and 159 circulating metabolic measures. The association with the CETP genetic score was statistically significant for 46 metabolic measures (P<0.00134). CETP concentration most strongly affected very large, large and medium HDL subclasses. With a 1 μ g/mL increase in CETP, all components of these lipoprotein subclasses de-

a High educational level: higher secondary education (according to Dutch educational system), higher vocational education, university, PhD.

creased. The only exception was the triglyceride content in medium HDL, for which the effect size was positive (effect (SE): 0.272 (0.055), $P=9\times10^{-7}$). In line with this, triglycerides in small HDL also showed a positive beta (effect (SE): 0.293 (0.055), $P=1\times10^{-9}$). In accordance with the decrease in larger HDL subclasses with higher CETP concentration, higher CETP was also associated with a smaller HDL diameter and less ApoA1. Overall, the largest effects was found for cholesterol in very large HDL (effect (SE): -0.517 (0.053), $P=6\times10^{-22}$). When comparing the lipoprotein components within the very large, large and medium HDL subclasses, the cholesterol components, i.e. cholesteryl esters and free cholesterol, consistently showed the largest effect sizes.

Remarkably, CETP concentration did not associate with any of the LDL subclass components, while a higher CETP concentration associated with more small and very small VLDL. The largest increasing effect was found for cholesteryl esters in small VLDL (effect (SE): 0.276 (0.061), $P=6\times10^{-6}$). There were no pronounced differences in the effect sizes between the various components within VLDL subclasses.

The results from the independent replication dataset can be found in Supplementary figure 6.B.1 and Supplementary table 6.C.3. The pattern of lipidomic associations was consistent in the replication dataset (Figure 6.2), also when evaluated for the three SNPs of the genetic score separately (Supplementary figures 6.B.2, 6.B.3 and 6.B.4).

Observationally, CETP concentration is predominantly positively associated with very small VLDL, IDL and LDL components

Figure 6.3 and Supplementary table 6.C.4 show the observational associations between measured serum CETP concentration and the 159 circulating metabolic measures. A higher CETP concentration was positively associated with all VLDL, IDL and LDL subclasses, of which very small VLDL, IDL, and all three LDL subclasses showed the strongest associations. This is in line with the associations between serum CETP concentration and the routine-lipid measurements (age- and sex- adjusted), since we observed a large effect on Friedewald LDL-C concentration per unit increase in CETP concentration (effect size 0.46 mmol/L; 95% CI 0.40, 0.52), compared to the relatively small effect on HDL-C concentration (effect size -0.05 mmol/L; 95% CI -0.07, -0.02). Crude associations are shown in Supplementary figure 6.B.5.

The largest effect sizes for observational associations were found for free cholesterol in very small VLDL (effect (SE): 0.476 (0.030), P=2×10⁻⁵⁵) and phospholipids in very small VLDL (effect (SE): 0.472 (0.030), $P=2\times10^{-56}$). When comparing the lipoprotein components (i.e. cholesteryl esters, free cholesterol, phospholipids and triglycerides) within the very small VLDL, IDL and LDL subclasses, the shell components, i.e. phospholipids and free cholesterol, quite consistently (apart from small LDL) showed the largest effect sizes, whereas the effect for the core component triglycerides was, consistently, the smallest. Concerning the HDL subclasses, higher CETP concentration was associated with a higher concentration of small HDL particles, whereas no associations were found with concentrations of very large, large and medium HDL particles.

Overall, observational and genetic associations were hardly consistent. Only for VLDL effect directions were similar (i.e. positive) between genetic and observational associations.

Figure 6.1: Causal associations between CETP concentration and 159 circulating metabolic measures, which were assessed with a Mendelian randomization approach based on a CETP genetic score, in the Netherlands Epidemiology of Obesity (NEO) study (n=5,672). The genetic score for Mendelian randomization is based on the CETP SNPs rs247616, rs12720922 and rs1968905, as previously determined with a genome-wide association study on serum CETP concentration.^[18] Bar heights represent the magnitude of the beta coefficient from linear regression, which is expressed as the SD difference in metabolic measure per 1 μ g/mL increase in CETP concentration. Red bars indicate positive betas and blue bars indicate negative betas. The transparency of the bars indicates the level of statistical significance. A P-value<0.00134 is regarded statistical significant, as represented by the black dots. Full names and descriptions of metabolic measures are listed in Supplementary table 6.C.1. Results are based on analyses weighted towards the reference BMI distribution of the general Dutch population.

Effect size NEO study

Figure 6.2: The effects sizes of the causal associations between CETP concentration and metabolic measures in the Netherlands Epidemiology of Obesity (NEO) study are strongly replicated in independent European populations (MAGNETIC NMR GWAS dataset), as shown by the high correlation between the beta coefficients from both cohorts. Beta coefficient from linear regression are expressed as the SD difference in metabolic measure per 1 μ g/mL increase in CETP concentration. NEO study results are based on analyses weighted towards the reference BMI distribution of the general Dutch population.

Figure 6.3: Observational associations between serum CETP concentration and 159 circulating metabolic measures, in the Netherlands Epidemiology of Obesity (NEO) study (n=5,672). Bar heights represent the magnitude of the beta coefficient from linear regression, which is expressed as the SD difference in metabolic measure per 1 μ g/mL increase in serum CETP. Red bars indicate positive betas and blue bars indicate negative betas. The transparency of the bars indicates the level of statistical significance. A P-value<0.00134 is regarded statistical significant, as represented by the black dots. Full names and descriptions of metabolic measures are listed in Supplementary table 6.C.1. Results are based on analyses weighted towards the reference BMI distribution of the general Dutch population.

Discussion

In this study, we determined the causal effects of serum CETP on the circulating lipoprotein profile, and contrasted the results to observational associations. We used three recently identified CETP SNPs (i.e. rs247616, rs12720922 and rs1968905) that together explain 16.4% of the total variation in serum CETP concentration as genetic instruments in Mendelian randomization, to enable inferences of causality on 159 circulating metabolic measures, and mimic the effects of CETP inhibition. Higher circulating CETP concentrations were causally most strongly associated with lower concentrations of very large, large and medium-sized HDL components, a smaller overall HDL diameter, less ApoA1, and more small VLDL components, while there was no association with LDL components. In contrast, observationally, measured serum CETP concentration predominantly associated with more VLDL, IDL and LDL components. We propose that this contrast between genetic and observational associations may be explained by the higher capacity of VLDL, IDL and LDL subclasses to carry CETP, thereby concealing causal associations with HDL.

Based on the principle of Mendelian randomization, the association between the CETP genetic score and lipoprotein subclasses can be interpreted as causal. The results from our Mendelian randomization analyses are fully consistent with the mechanism by which CETP transfers cholesteryl esters from HDL towards ApoB containing triglyceride-rich lipoproteins, mainly VLDL, and mediates the transfer of triglycerides from VLDL towards triglyceride-poor particles, such as HDL .^[1,7,27] CETP accelerates the clearance of HDL from blood by enrichment of HDL with triglycerides, which makes this HDL a preferred substrate for hepatic triglyceride lipase.^[2] Catabolism of triglyceride-rich HDL by hepatic triglyceride lipase leads to the formation of very small remnant HDL that is cleared by the liver and kidneys.^[2,13,28–31] Indeed, in the present study, the CETP genetic score was associated with a smaller HDL diameter, and with a higher triglyceride content specifically in the medium and small HDL particles. Also, our results show that the CETP mainly affects the HDL pool, which is in accordance with the lipoprotein profile observed in CETP-deficient individuals, who have markedly increased amounts of large HDL particles, [32-34] while effects on LDL and VLDL subclasses are less pronounced.^[34]

Based on the principle of Mendelian randomization, our results may also shed light on the mechanisms that underlie the effectiveness of CETP inhibition.^[19,20] Therefore, it is highly interesting to discuss the present study results in light of the results from the RE-VEAL trial with the CETP inhibitor anacetrapib, which showed a reduction in major coronary events.^[35] Our findings implicate that CETP inhibition causes a relatively large increase in HDL components, which is predominantly caused by an increase in large and mediumsized HDL particles, in addition to a more modest reduction in VLDL components, which

is mainly caused by a decrease in small and extra small VLDL particles. Of note, LDL concentration and composition were not affected by CETP, which was unexpected based on the current dogma that CETP increases LDL-C. However, it should be realized that LDL-C is generally not measured directly but calculated from the Friedewald formula, [2] which may well misclassify cholesterol contained within small VLDL subclasses as LDL-C. It was previously reported that small and very small VLDL particle concentrations (reflecting triglyceride-rich lipoprotein remnants) and cholesterol in VLDL (reflecting remnant cholesterol) are among the lipoprotein components that most strongly associate with an increased CVD risk.^[16,36,37] Thus, the effects of CETP on small VLDL particles may well explain the positive outcome of the latest CETP inhibition trial with anacetrapib.^[14]

We compared the genetic associations with non-causal observational associations between serum CETP concentration measured with ELISA and all 159 metabolic measures. Observationally, associations were markedly different, as CETP concentration was predominantly positively associated with very small VLDL, IDL and LDL components, whereas no negative associations with HDL components were observed. This indicates that confounding factors conceal the true causal association. Although we showed previously that CETP concentration is strongly genetically determined, $[18]$ still the majority of the total variation in CETP concentration between individuals is caused by environmental factors. It should be realized that, just like apolipoproteins, CETP is an amphiphilic protein that binds to lipid surfaces of lipoproteins. It is thus conceivable that the capacity of lipoproteins to carry CETP is one of the factors that may explain the marked differences between the observational and causal associations.

The distribution of CETP over circulating lipoproteins has been extensively studied in the past, however, with diverging conclusions. The earliest studies proposed that CETP is predominantly bound to HDL particles.^[24,38,39] However, thereafter, it was shown that CETP also avidly binds to LDL and VLDL, although the binding to these particles might be more labile.^[40] It should be noted that the distribution of CETP over the various lipoprotein subclasses has commonly been determined after separating lipoproteins with ultracentrifugation, thereby presumably disrupting the binding within the least stable CETP-lipoprotein complexes. In addition, storage of plasma samples may also alter the distribution of CETP over lipoproteins. Indeed, marked effects of storage on the lipoprotein distribution have been observed for ApoC3 and ApoE following fast protein liquid chromatography.^[41] Results from previous studies may therefore not accurately represent the distribution of CETP over lipoproteins within the circulation. Remarkably, in the present study, observational serum CETP concentration showed the strongest associations with the shell components (i.e. phospholipids and free cholesterol) of very small VLDL, IDL, and LDL subclasses. This may indicate that CETP is mostly carried on the surface of these particles in the circulation, although future studies are needed to substantiate this hypothesis.

The main strength of the present study is the use of a strong genetic instrument in Mendelian randomization to draw conclusions on the causal effects of CETP concentration on the circulating lipoprotein profile. In addition, sufficient statistical power for the analyses was provided, as we had genetic data and NMR-metabolomic profiles available of 5,672 individuals, which we were able to replicate in an independent dataset (n≈20,000). However, the study populations were from European ancestry and results may therefore not be generalizable to other populations. Interestingly, a recent study in a Chinese population, which used a different CETP genetic score (i.e. rs3764261, rs1800775, rs708272, rs9939224, and rs2303790), showed comparable results.^[42] This indicates consistency of the results among different ancestries.

Based on Mendelian randomization, we conclude that CETP is an important determinant of HDL concentration and composition, without affecting LDL concentration and composition. Our finding challenges the current dogma that CETP increases LDL-C, mainly based on indirect LDL-C estimation using the Friedewald formula.^[2] Instead, by directly assessing the lipoprotein profile by NMR, we now show that CETP specifically increases small VLDL fractions that likely represent VLDL remnants. Therefore, we speculate that the CETP inhibitor anacetrapib attenuated cardiovascular disease risk through specific reduction of remnant cholesterol rather than LDL-C.

Acknowledgements

We express our gratitude to all individuals who participate in the NEO study. We are grateful to all participating general practitioners for inviting eligible participants. We furthermore thank Pat van Beelen and all research nurses for collecting the data, Petra Noordijk and her team for laboratory management and DNA isolation, and Ingeborg de Jonge for all data management of the NEO study. We sincerely thank Chris van der Bent for performing the serum CETP concentration measurements. The genotyping in the NEO study was supported by the Centre National de Génotypage (Paris, France), headed by Jean-Francois Deleuze.

The NEO study was supported by the participating Departments, the Division and the Board of Directors of the Leiden University Medical Center, and by the Leiden University, Research Profile Area 'Vascular and Regenerative Medicine'. L.L. Blauw was supported by a grant from the Board of Directors of the Leiden University Medical Center. Y. Wang is supported by the Dutch Science Organization [ZonMW-VENI Grant 91617027]. D. van Heemst was supported by the European Commission funded project HUMAN [Health-2013-INNOVATION-1-602757]. D. Mook-Kanamori is supported by the Dutch Science Organization [ZonMW-VENI Grant 91614023]. P. Würtz is supported by the Academy of Finland [312476 and 312477] and the Novo Nordisk Foundation [15998]. This work is further supported by the Netherlands CardioVascular Research Initiative: the Dutch Heart Foundation, Dutch Federation of University Medical Centers, the Netherlands Organisation for Health Research and Development and the Royal Netherlands Academy of Sciences for the GENIUS project 'Generating the best evidence-based pharmaceutical targets for atherosclerosis' [CVON2011-9].

Conflict of Interest: S. Soidinsalo and P. Würtz are employees of Nightingale Health Ltd, a company offering NMR-based metabolite profiling. P. Würtz is shareholder of Nightingale Health Ltd.

References

- [1] A. R. Tall. Plasma cholesteryl ester transfer protein. *J Lipid Res*, 34(8):1255–74, 1993.
- [2] M. J. Chapman, W. Le Goff, and M. Guerin et al. Cholesteryl ester transfer protein: at the heart of the action of lipid-modulating therapy with statins, fibrates, niacin, and cholesteryl ester transfer protein inhibitors. *Eur Heart J*, 31(2):149–64, 2010.
- [3] M. Westerterp, C. C. van der Hoogt, and W. de Haan et al. Cholesteryl ester transfer protein decreases highdensity lipoprotein and severely aggravates atherosclerosis in APOE*3-Leiden mice. *Arterioscler Thromb Vasc Biol*, 26(11):2552–9, 2006.
- [4] L. A. Morehouse, E. D. Sugarman, and P. A. Bourassa et al. Inhibition of CETP activity by torcetrapib reduces susceptibility to diet-induced atherosclerosis in New Zealand White rabbits. *J Lipid Res*, 48(6):1263–72, 2007.
- [5] H. Okamoto, F. Yonemori, and K. Wakitani et al. A cholesteryl ester transfer protein inhibitor attenuates atherosclerosis in rabbits. *Nature*, 406(6792):203–7, 2000.
- [6] P. J. Barter, M. Caulfield, and M. Eriksson et al. Effects of torcetrapib in patients at high risk for coronary events. *N Engl J Med*, 357(21):2109–22, 2007.
- [7] D. Lei, M. Rames, and X. Zhang et al. Insights into the tunnel mechanism of cholesteryl ester transfer protein through all-atom molecular dynamics simulations. *J Biol Chem*, 291(27):14034–44, 2016.
- [8] L. L. Blauw, R. de Mutsert, and H. J. Lamb et al. Serum CETP concentration is not associated with measures of body fat: The NEO study. *Atherosclerosis*, 246:267–73, 2016.
- [9] D. J. Gordon, J. L. Probstfield, and R. J. Garrison et al. High-density lipoprotein cholesterol and cardiovascular disease. Four prospective American studies. *Circulation*, 79(1):8–15, 1989.
- [10] R. Frikke-Schmidt, B. G. Nordestgaard, and M. C. Stene 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(21):2524–32, 2008.
- [11] C. L. Haase, A. Tybjaerg-Hansen, and P. Grande et al. Genetically elevated apolipoprotein A-I, high-density lipoprotein cholesterol levels, and risk of ischemic heart disease. *J Clin Endocrinol Metab*, 95(12):E500–10, 2010.
- [12] T. H. Johannsen, P. R. Kamstrup, and R. V. Andersen et al. Hepatic lipase, genetically elevated high-density lipoprotein, and risk of ischemic cardiovascular disease. *J Clin Endocrinol Metab*, 94(4):1264–73, 2009.
- [13] B. F. Voight, G. M. Peloso, and M. Orho-Melander et al. Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. *Lancet*, 380(9841):572–80, 2012.
- [14] L. Bowman, J. C. Hopewell, and F. Chen et al. Effects of anacetrapib in patients with atherosclerotic vascular disease. *N Engl J Med*, 377(13):1217–1227, 2017.
- [15] B. A. Ference, J. J. P. Kastelein, and H. N. Ginsberg et al. Association of Genetic Variants Related to CETP Inhibitors and Statins With Lipoprotein Levels and Cardiovascular Risk. *JAMA*, 318(10):947–956, 2017.
- [16] P. Wurtz, A. S. Havulinna, and P. Soininen et al. Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts. *Circulation*, 131(9):774–85, 2015.
- [17] J. C. Tardif, D. Rhainds, and M. Brodeur et al. Genotype-dependent effects of dalcetrapib on cholesterol efflux and inflammation: Concordance with clinical outcomes. *Circ Cardiovasc Genet*, 9(4):340–8, 2016.
- [18] L. L. Blauw, R. Li-Gao, and R. Noordam et al. CETP (Cholesteryl Ester Transfer Protein) Concentration: A Genome-Wide Association Study Followed by Mendelian Randomization on Coronary Artery Disease. *Circ Genom Precis Med*, 11(5):e002034, 2018.
- [19] B. A. Ference. How to use mendelian randomization to anticipate the results of randomized trials. *Eur Heart J*, 2017.
- [20] G. Thanassoulis and C. J. O'Donnell. Mendelian randomization: nature's randomized trial in the post-genome era. *JAMA*, 301(22):2386–8, 2009.
- [21] P. Soininen, A. J. Kangas, and P. Würtz et al. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. *Circ Cardiovasc Genet*, 8(1):192–206, 2015.
- [22] R. de Mutsert, M. den Heijer, and T. J. Rabelink et al. The Netherlands Epidemiology of Obesity (NEO) study: study design and data collection. *Eur J Epidemiol*, 28(6):513–23, 2013.
- [23] G. R. Abecasis, A. Auton, and L. D. Brooks et al. An integrated map of genetic variation from 1,092 human genomes. *Nature*, 491(7422):56–65, 2012.
- [24] J. E. Groener, A. J. Van Rozen, and D. W. Erkelens. Cholesteryl ester transfer activity. localization and role in distribution of cholesteryl ester among lipoproteins in man. *Atherosclerosis*, 50(3):261–71, 1984.
- [25] J. Kettunen, A. Demirkan, and P. Wurtz et al. Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. *Nat Commun*, 7:11122, 2016.
- [26] J. Li and L. Ji. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity (Edinb)*, 95(3):221–7, 2005.
- [27] X. Qiu, A. Mistry, and M. J. Ammirati et al. Crystal structure of cholesteryl ester transfer protein reveals a long tunnel and four bound lipid molecules. *Nat Struct Mol Biol*, 14(2):106–13, 2007.
- [28] M. A. Charles and J. P. Kane. New molecular insights into CETP structure and function: a review. *J Lipid Res*, 53(8):1451–8, 2012.
- [29] K. Guendouzi, X. Collet, and B. Perret et al. Remnant high density lipoprotein2 particles produced by hepatic lipase display high-affinity binding and increased endocytosis into a human hepatoma cell line (HEPG2). *Biochemistry*, 37(42):14974–80, 1998.
- [30] H. H. Newnham and P. J. Barter. Synergistic effects of lipid transfers and hepatic lipase in the formation of very small high-density lipoproteins during incubation of human plasma. *Biochim Biophys Acta*, 1044(1):57–64, 1990.
- [31] K. Rye, M. A. Clay, and P. J. Barter. Remodelling of high density lipoproteins by plasma factors. *Atherosclerosis*, 145(2):227–238, 1999.
- [32] T. Arai, T. Tsukada, and T. Murase et al. Particle size analysis of high density lipoproteins in patients with genetic cholesteryl ester transfer protein deficiency. *Clin Chim Acta*, 301(1-2):103–17, 2000.
- [33] M. L. Brown, A. Inazu, and C. B. Hesler et al. Molecular basis of lipid transfer protein deficiency in a family with increased high-density lipoproteins. *Nature*, 342(6248):448–51, 1989.
- [34] J. Koizumi, A. Inazu, and K. Yagi et al. Serum lipoprotein lipid concentration and composition in homozygous and heterozygous patients with cholesteryl ester transfer protein deficiency. *Atherosclerosis*, 90(2):189–196, 1991.
- [35] V. A. Eyvazian and W. H. Frishman. Evacetrapib: Another CETP Inhibitor for Dyslipidemia with No Clinical Benefit. *Cardiol Rev*, 25(2):43–52, 2017.
- [36] A. Varbo, M. Benn, and B.G. Nordestgaard. Remnant cholesterol as a cause of ischemic heart disease: Evidence, definition, measurement, atherogenicity, high risk patients, and present and future treatment. *Pharmacol Ther*, 141(3):358–367, 2014.
- [37] A. Varbo, M. Benn, and A. Tybjærg-Hansen et al. Remnant cholesterol as a causal risk factor for ischemic heart disease. *J Am Coll Cardiol*, 61(4):427–436, 2013.
- [38] P. J. Barter, G. J. Hopkins, and L. Gorjatschko et al. A unified model of esterified cholesterol exchanges between human plasma lipoproteins. *Atherosclerosis*, 44(1):27–40, 1982.
- [39] N. M. Pattnaik and D. B. Zilversmit. Interaction of cholesteryl ester exchange protein with human plasma lipoproteins and phospholipid vesicles. *J Biol Chem*, 254(8):2782–6, 1979.
- [40] R. E. Morton. Binding of plasma-derived lipid transfer protein to lipoprotein substrates. the role of binding in the lipid transfer process. *J Biol Chem*, 260(23):12593–9, 1985.
- [41] J. S. Cohn, C. Rodriguez, and H. Jacques et al. Storage of human plasma samples leads to alterations in the lipoprotein distribution of apoC-III and apoE. *J Lipid Res*, 45(8):1572–9, 2004.
- [42] I. Y. Millwood, D. A. Bennett, and M. V. Holmes et al. Association of CETP gene variants with risk for vascular and nonvascular diseases among chinese adults. *JAMA Cardiology*, 2017.
- [43] E. L. Korn and B. I. Graubard. Epidemiologic studies utilizing surveys: accounting for the sampling design. *Am J Public Health*, 81(9):1166–73, 1991.
- [44] T. Lumley. Analysis of complex survey samples. 2004.

Appendix

6.A. Expanded methods

BMI-weighted analyses

In the NEO study, individuals with a BMI of 27 kg/m² or higher were oversampled. Men and women aged between 45 and 65 years with a self-reported BMI of 27 kg/ $m²$ or higher were eligible to participate. From one nearby municipality (Leiderdorp, The Netherlands) all inhabitants aged between 45 and 65 years were invited to participate regardless of their BMI, in order to obtain a reference distribution for BMI. To correctly represent baseline associations in the general population^[43], adjustments for the oversampling of individuals with a BMI ≥27 kg/m² were made in the analyses. This was done by weighting all participants towards the BMI distribution of participants from the Leiderdorp municipality^[44], whose BMI distribution was similar to the BMI distribution of the general Dutch population in the age range of 45-65 years^[8]. In practice, this means that participants with a lower BMI were assigned larger weights than participants with a higher BMI in the analyses. All results are based on weighted analyses, and therefore apply to a population-based study without oversampling of individuals with overweight or obesity. As a consequence, the weighted characteristics of the population are expressed in proportions instead of absolute numbers.

Figure 6.B.1: Replication of the causal associations between serum CETP concentration and 111 circulating metabolic measures, which were assessed with a Mendelian randomization approach based on a CETP genetic score, using summary statistics from the MAGNETIC NMR GWAS dataset (n≈20,000). The genetic score for Mendelian randomization is based on the CETP SNPs rs247616, rs12720922 and rs1968905, as previously determined with a genome-wide association study on serum CETP concentration.[18] Bar heights represent the magnitude of the beta coefficient from linear regression, which is expressed as the SD difference in metabolic measure per 1 µg/mL increase in CETP concentration. Red bars indicate positive betas and blue bars indicate negative betas. The transparency of the bars indicates the level of statistical significance. Names of unmeasured metabolic measures are grey. A P-value<0.00227 is regarded statistical significant, as represented by the black dots. Full names and descriptions of metabolic measures are listed in Supplementary table 6.C.1.

Figure 6.B.2: Causal associations between the CETP SNPs rs247616, rs12720922 and rs1968905 and 159 circulating metabolic measures, which were assessed with a Mendelian randomization approach, in the Netherlands Epidemiology of Obesity (NEO) study (n=5,672). Dots represent the beta coefficient from linear regression, which is expressed as the SD difference in metabolic measure per additional CETP-increasing allele. Bars represent 95% confidence intervals. Alleles that are associated with an increase in serum CETP concentration are rs247616-C, rs12720922-A and rs1968905-G. Full names and descriptions of metabolic measures are listed in Supplementary table 6.C.1.

Figure 6.B.3: Replication of the causal associations between the CETP SNPs rs247616, rs12720922 and rs1968905 and 112 circulating metabolic measures, which were assessed with a Mendelian randomization approach, using summary statistics from the MAGNETIC NMR GWAS dataset (n≈20,000). Dots represent the beta coefficient from linear regression, which is expressed as the SD difference in metabolic measure per additional CETP-increasing allele. Bars represent 95% confidence intervals. Alleles that are associated with an increase in serum CETP concentration are rs247616-C, rs12720922-A and rs1968905-G. Full names and descriptions of metabolic measures are listed in Supplementary table 6.C.1.

Figure 6.B.4: Correlation between the effect sizes (i.e. beta coefficients) for the different CETP genetic variants included in the genetic score, i.e. (**a**) rs247616 vs rs12720922 and (**b**) rs1968905 vs rs247616, in the MAGNETIC NMR GWAS dataset (n≈20,000). Beta coefficient from linear regression are expressed as the SD difference in metabolic measure per 1 μ g/mL increase in CETP concentration.

Figure 6.B.5: Crude correlations of serum CETP concentration with (**a**) calculated Friedewald LDL cholesterol concentration and (**b**) routinely measured HDL cholesterol concentration (clinical chemistry assay), in the Netherlands Epidemiology of Obesity (NEO) study (n=5,672). Results are based on analyses weighted towards the reference BMI distribution of the general Dutch population.

6.C. Supplementary tables

Table 6.C.1: Description of 159 metabolic measures that were measured with a high-throughput proton nuclear magnetic resonance metabolomics platform (Nightingale Health Ltd., Helsinki, Finland).

Lipoprotein particle size VLDL particle size LDL particle size HDL particle size *Apolipoproteins* ApoB/ApoA1 *Cholesterols* Esterified cholesterol Free cholesterol Total cholesterol in HDL2 Total cholesterol in HDL3 Total cholesterol in HDL Total cholesterol in LDL Remnant cholesterol Serum total cholesterol Total cholesterol in VLDL *Fatty acids* Conjugated linoleic acid Conjugated linoleic acid/FA Docosahexaenoic acid Docosahexaenoic acid/FA Fatty acid chain length Omega-3/FA Omega-6/FA Linoleic acid/FA Total fatty acids Degree of unsaturation *Glycerides* Diacylglycerol Diacylglycerol/FA Triglycerides in HDL

The genetic score for Mendelian randomization is based on the CETP SNPs rs247616, rs12720922 and rs1968905. A P-value<0.00134 is regarded statistical significant. Full names and descriptions of metabolic measures are listed in Supplementary table 6.C.1. Results are based on analyses weighted towards the reference BMI distribution of the general Dutch population.

b Remnant cholesterol is defined as non-HDL and non-LDL-cholesterol.

Table 6.C.2: Causal associations between CETP concentration and 159 circulating metabolic measures, which were assessed with a Mendelian randomization approach based on a CETP genetic score, in the Netherlands Epidemiology of Obesity (NEO) study (n=5,672).

a The HDL2 and HDL3 fractions are defined as those HDL particles within the density range of 1.063-1.125 g/mL and 1.125-1.210 g/mL, respectively.

L.LDL.P 0.05801 0.055 0.05801 L.LDL.PL -0.038 0.055 0.4916

The genetic score for Mendelian randomization is based on the CETP SNPs rs247616, rs12720922 and rs1968905. A P-value<0.00134 is regarded statistical significant. Full names and descriptions of metabolic measures are listed in Supplementary table 6.C.1. Results are based on analyses weighted towards the reference BMI distribution of the general Dutch population.

 a The beta coefficient is the natural-log transformed SD difference in metabolic measure per 1 μ g/mL increase in CETP concentration.

Table 6.C.3: Replication of the causal associations between CETP concentration and 111 circulating metabolic measures, which were assessed with a Mendelian randomization approach based on a CETP genetic score, using summary statistics from the MAGNETIC NMR GWAS dataset (n 20,000).

The genetic score for Mendelian randomization is based on the CETP SNPs rs247616, rs12720922 and rs1968905. A P-value<0.00227 is regarded statistical significant. Full names and descriptions of metabolic measures are listed in Supplementary table 6.C.1. N/A, not quantified.

a The beta coefficient is the natural-log transformed SD difference in metabolic measure per 1 µg/mL increase in CETP concentration.

Table 6.C.4: Observational associations between CETP concentration and 159 circulating metabolic measures, in the Netherlands Epidemiology of Obesity (NEO) study (n=5,672).

Metabolic measure	Beta coefficient ^a	SE	P-value
Very-low-density lipoproteins (VLDL)			
XXL.VLDL.C	0.140	0.027	2E-07
XXL.VLDL.CE	0.186	0.027	8E-12
XXL.VLDL.FC	0.084	0.026	0.0014
XXL.VLDL.L	0.104	0.026	0.0001
XXL.VLDL.P	0.102	0.026	0.0001
XXL.VLDL.PL	0.098	0.026	0.0002
XXL.VLDL.TG	0.095	0.026	0.0003
XL.VLDL.C	0.091	0.028	0.0010
XL.VLDL.CE	0.098	0.028	0.0004
XL.VLDL.FC	0.084	0.028	0.0024
XL.VLDL.L	0.051	0.027	0.0640
XL.VLDL.P	0.047	0.027	0.0873
XL.VLDL.PL	0.070	0.028	0.0112
XL.VLDL.TG	0.034	0.027	0.2096
L.VLDL.C	0.081	0.030	0.0074
L.VLDL.CE	0.113	0.030	0.0002
L.VLDL.FC	0.049	0.030	0.1032
L.VLDL.L	0.044	0.030	0.1457
L.VLDL.P	0.041	0.030	0.1769
L.VLDL.PL	0.057	0.030	0.0597
L.VLDL.TG	0.027	0.030	0.3828
M.VLDL.C	0.165	0.030	5E-08
M.VLDL.CE	0.237	0.030	$1E-15$
M.VLDL.FC	0.073	0.031	0.0167
M.VLDL.L	0.079	0.031	0.0103
M.VLDL.P	0.072	0.031	0.0203
M.VLDL.PL	0.094	0.031	0.0023
M.VLDL.TG	0.034	0.031	0.2804
S.VLDL.C	0.356	0.032	3E-28
S.VLDL.CE	0.415	0.032	2E-38
S.VLDL.FC	0.220	0.032	6E-12
S.VLDL.L	0.200	0.032	4E-10
S.VLDL.P	0.178	0.032	3E-08
S.VLDL.PL	0.211	0.032	8E-11
S.VLDL.TG	0.064	0.032	0.0421
XS.VLDL.C	0.452	0.030	1E-49
XS.VLDL.CE	0.432	0.030	5E-45
XS.VLDL.FC	0.476	0.030	2E-55

A P-value<0.00134 is regarded statistical significant. Full names and descriptions of metabolic measures are listed in Supplementary table 6.C.1. Results are based on analyses weighted towards the reference BMI distribution of the general Dutch population.

^a The beta coefficient is the natural-log transformed SD difference in metabolic measure per 1 μ g/mL increase in CETP concentration.