Investigating the role of triglycerides and triglyceride-containing lipoproteins in cardiovascular disease, using observational and genetic epidemiological methods

Roshni Joshi

Thesis submitted for the degree of Doctor of Philosophy at University College London

Institute of Cardiovascular Science
School of Life and Medical Sciences
University College London

I, Roshni Joshi confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in this thesis.

I dedicate this thesis to my mum, Manjula Joshi, who sadly lost her life to cancer in January 2021. Her words to me upon receiving the terminal diagnosis, "you've come this far, you must keep going". I remember those words every day, life is for living.

Abstract

Despite effective low-density lipoprotein cholesterol (LDL-C) lowering by statins, there remains a residual risk of CVD in individuals, which may in part be due to elevated triglyceride (TG) levels. Existing research evaluating the relationship between triglycerides and CVD has so far been mixed, and therapeutic agents to reduce triglycerides for CVD prevention are not routinely prescribed. Previous studies investigate total serum TG, which represents the summation of TG carried across all lipoproteins. It is possible that certain TG containing lipoproteins are more atherogenic than others, and thus using total serum TG measurement may be insufficiently precise to delineate any causal effect. ^{1}H-nuclear magnetic resonance (NMR) spectroscopy classifies lipoproteins into 14 different lipoprotein subfractions based on size and lipid composition, offering a more detailed interrogation to help confirm or refute the possible causal relationship between TG and CVD.

This thesis assesses the distribution of cholesterol and triglyceride content in 14 lipoprotein subfractions and establishes reference interval ranges based on the 2.5th and 97.5th percentiles. The largest interval range for TG content was observed in the medium VLDL subfraction (2.5th 97.5th percentile; 0.08 to 0.68 mmol/L), and for cholesterol content in the large LDL subfraction (0.47-1.45 mmol/L). TG concentrations in all sub-classes increased with increasing age and BMI. Increases in cholesterol concentrations were largely comparable between men and women by age, smoking status, and between fasting and non-fasting states. TG subfraction concentrations were significantly higher in ever smokers compared to never smokers, among subjects with CVD and type 2 diabetes as compared to disease-free subjects.

The TG content in the 14 lipoprotein subfractions is evaluated for association with CVD, and the extent to which the effect is independent of LDL-C and HDL-C is explored in observational analysis in Chapter 5. The results in this chapter demonstrate TG in 13

lipoprotein subfractions were positively associated with CHD (OR in the range 1.12 to 1.22). The positive effect estimates attenuated after adjustment for HDL-C and LDL-C. There was an absence of evidence demonstrating any association TG lipoprotein subfraction with stroke. Next, to elucidate potential causal relationships, observational and Mendelian randomisation (MR) approaches are used to investigate the total and direct effects of triglyceride and cholesterol content on CHD. There was a total causal association of TG content in five lipoprotein subfractions, and total association of cholesterol content in 10 lipoprotein subfractions with CHD. Multivariable MR analysis was used to explore the direct effects of TG content of the 14 lipoprotein subfractions conditioning on the cholesterol content, and vice versa for cholesterol associations. Here we found that there was a direct association of CHD for TG in four lipoprotein subfractions and cholesterol in 10 lipoprotein subfractions. Cholesterol content in triglyceride-rich lipoproteins (TRL) displayed the largest effects (MVMR OR in the range 2.73 to 14.31), an association that was not observed for TG in TRL. The observational and MR associations between TG content in lipoprotein subfractions and CHD presented here may be relevant in the context of ongoing drug development targeting TG-mediated pathways for disease reduction. An emerging approach to lower TG concentrations and lower risk of CHD is through inhibition of LPL function. Angiopoietin-like proteins 3 and 4 (ANGPLT3/4) are negative regulators of LPL and have recently emerged as novel drug targets to manage dyslipidaemia. The final section of this thesis discusses the contribution of the results to the current understanding of role of TG in CVD and translational applications to clinical care.

Impact statement

Triglycerides have long been thought to contribute to the residual cardiovascular disease risk observed in patients who achieve guideline recommended low-density lipoprotein cholesterol (LDL-C) targets. The detailed evaluation of the association of the triglyceride content in fourteen lipoprotein subfractions with cardiovascular disease presented in this thesis, contribute to a better understanding of the complex relationship between triglycerides, triglyceride-rich lipoproteins (TRL) and disease. Advances in this field increase the understanding of the causality of triglycerides and the atherogenicity of lipoproteins other than LDL, which are crucial for identifying novel lipid lowering targets, and for the interpretation of future clinical trials. The findings in this thesis support the recently published European Atherosclerosis Society Statement on Triglyceride-rich Lipoproteins and their Remnants.

An important aspect of this thesis are the reported distributions and determinants of cholesterol and triglycerides in the fourteen lipoprotein subfractions presented in chapter 4. Population based reference intervals are a widely used tool to interpret patient laboratory results, for example, the measurement of LDL-C for assessment of cardiovascular disease risk and to determine if lipid lowering is indicated. The determination of the reference range intervals in a generalisable population in this thesis forms the basis for scientific advances, and translation in the field of lipidomics and clinical practice. NMR quantification of lipoprotein lipids is becoming increasingly common in biobanks and has the potential to be made available in clinical care, where it is envisioned the reference intervals described in chapter 4 will serve as a valuable tool to aid decision making.

The observational results in chapter 5 advance the current understanding of the association of triglycerides in cardiovascular disease. Thanks to methods like Mendelian randomisation, it is possible to ascertain causality with more certainty than using traditional epidemiological methods. Chapter 6 identifies cholesterol in certain lipoprotein subfractions as the predominate lipid associated with disease. Identifying the components, that is, triglycerides, TRL or the cholesterol content of TRL, that give rise to risk, is essential to understand the pathological consequences of triglycerides in the context of residual cardiovascular risk. The results in chapter 6 provide important insights for future work both within and beyond academia and contribute to the growing body of evidence in this field. The latter is key given recent investigations of the proatherogenic effects of apoB-containing lipoproteins, non-HDL-cholesterol and remnant cholesterol.

Knowledge of TG metabolism, pathobiology of the different lipoproteins and the atherogenic potential of their lipid composition is still very much a work-in-progress. There are several promising contenders targeting triglyceride pathways showing promising cardiovascular benefit. Academic research and clinical trials of these candidates are ongoing at the time of writing this thesis. Anticipated results are envisioned to provide further elucidation of the proatherogenic effects of triglycerides, TRL and their remnants to ultimately translate to a reduction in cardiovascular disease risk.

Acknowledgements

My first acknowledgement is to say that I have had a wonderful experience throughout the PhD. I have been very lucky to have been a part of an enthusiastic, warm and welcoming research group and institute. I have made lifelong friends who have made the last few years a joyous adventure. They have also shown me immense support through very difficult times.

My heartfelt appreciation and gratitude to my PhD supervisors, Professor Aroon Hingorani, Professor Goya Wannamethee and Dr Floriaan Schmidt. They have shown patience and understanding in supervising me. Their empathetic and compassionate supervision approach has helped me navigate through the ups and downs of the last four years. It is through their encouragement and kindness that I have been able to complete this PhD.

My little sister Priya Joshi, who in many ways looks after me like a big sister. She has shared this PhD experience with me. Thank you for reading my work and texting me telling me to ‘get on with it’. Your strength and insight motivate me.

My Dad, Vinodray Joshi, who always gently enquires about the status of my thesis. I thank my parents for being open-minded and for the love and stability they have consistently provided.

My darling mum, Manjula Joshi who passed away in the final year of my PhD. She made many sacrifices, many of which I’ll never know, to facilitate my education. She encouraged independence and strength of character, and taught me that an education will give me a voice to help me understand the world. It seems her whole life focussed on this very goal. It is

because of her reassurance, inspiration and unwavering confidence in me that has led me to this point. I miss her in every moment, and I'll celebrate every accomplishment in her honour.

Funding

This work was supported by the British Heart Foundation 4-year studentship in Cardiovascular Research at University College London.

Abbreviations

ASCVD: atherosclerotic cardiovascular disease	LD: linkage disequilibrium
BMI: body mass index	LDL: low-density lipoprotein
BRHS: British regional heart study	MI: myocardial infarction
BWHHS: British women's heart and health	MR: mendelian randomisation
CHD: coronary heart disease	MVMR: multivariable mendelian randomisation
CAPS: Caerphilly prospective study	NMR: nuclear magnetic resonance
CARDIOGRAMplusC4D: coronary artery disease genome-wide replication and meta-analysis plus the coronary artery disease	OR: odds ratio
CI: confidence interval	RCT: randomisation controlled trial
CVD: cardiovascular disease	SABRE: Southall and brent revisited
DBP: diastolic blood pressure	SBP: systolic blood pressure
FDR: false discovery rate	SD: standard deviation
GWAS: genome-wide association study	SE: standard error
HDL: high-density lipoprotein	SNP: single nucleotide polymorphism
IDL: intermediate-density lipoprotein	UCLEB: University College London-Edinburgh-Bristol Consortium
IS: ischaemic stroke	VLDL: very low-density lipoprotein
IVW: inverse-variance weighted	WHII: Whitehall II

Related academic work

Publications

1. **Joshi, Roshni**, et al, In preparation "Evaluating the causal relevance of triglyceride and cholesterol content in 14 NMR measured lipoprotein subfractions with risk of coronary heart disease: An observational and genetic analysis", 2021
2. **Joshi, R.,** Wannamethee, G., Engmann, J., Gaunt, T., Lawlor, D. A., Price, J., Papacosta, O., Shah, T., Tillin, T., Whincup, P., Chaturvedi, N., Kivimaki, M., Kuh, D., Kumari, M., Hughes, A. D., Casas, J. P., Humphries, S. E., Hingorani, A. D., Schmidt, A. F., & UCLEB Consortium (2021). Establishing reference intervals for triglyceride-containing lipoprotein subfraction metabolites measured using nuclear magnetic resonance spectroscopy in a UK population. *Annals of clinical biochemistry*, *58*(1), 47–53. https://doi.org/10.1177/0004563220961753
3. **Joshi, R**., Wannamethee, S. G., Engmann, J., Gaunt, T., Lawlor, D. A., Price, J., Papacosta, O., Shah, T., Tillin, T., Chaturvedi, N., Kivimaki, M., Kuh, D., Kumari, M., Hughes, A. D., Casas, J. P., Humphries, S., Hingorani, A. D., & Schmidt, A. F. (2020). Triglyceride-containing lipoprotein subfractions and risk of coronary heart disease and stroke: A prospective analysis in 11,560 adults. *European Journal of Preventive Cardiology*, *27*(15), 1617–1626. https://doi.org/10.1177/2047487319899621

Publications separate to the PhD thesis

4. Garfield, V., **Joshi, Roshni**., Garcia-Hernandez, J., Tillin, T., & Chaturvedi, N. (2019). The relationship between sleep quality and all-cause, CVD and cancer mortality: the Southall and Brent REvisited study (SABRE). Sleep medicine, 60, 230–235. https://doi.org/10.1016/j.sleep.2019.03.012

Senior author publication

5. Topriceanu, Constantin-Cristian, Therese Tillin, Nishi Chaturvedi, **Roshni Joshi**, and Victoria Garfield. "The association between plasma metabolites and sleep quality in the Southall and Brent Revisited (SABRE) Study: A cross-sectional analysis." Journal of Sleep Research (2020)

Oral Presentations

I have delivered an oral presentation related to the contents of this PhD at the following conference

- BHF student conference, 2018
- European Atherosclerosis Society, 2020

Poster presentations

I have delivered poster presentations related to the contents of this PhD at the following conferences and meetings

- International Genetic Epidemiology Society, 2020
- American Heart Association Congress, Chicago 2018
- British Atherosclerosis Society, Cambridge 2018

Funded workshop attendance

- International Atherosclerosis Research School, European Atherosclerosis Society, Prague, 2019

1 Introduction

This introductory chapter will provide an overview of atherosclerosis, describe the risk factors for atherosclerotic cardiovascular disease, and the composition and role of lipoprotein lipids in disease pathogenesis.

1.1 Atherosclerosis and cardiovascular disease

Atherosclerosis is a disease of large and medium size arteries. The normal artery comprises three layers: the innermost intima (in close contact with the bloodstream), the tunica media, and the adventitia[1] Atherosclerosis refers to the accumulation of fatty, fibrous and inflammatory material in the innermost layer of arteries, the intima.. Under homeostatic conditions the endothelial layer lining the intima resists the adherence of blood leukocytes and platelets due to the actions of nitric oxide and prostaglandins.[2]. Endothelial damage or dysfunction due to exposure to risk factors such as smoking, high blood pressure and diabetes, results in an increase in endothelial permeability to low-density lipoproteins (LDL) and increased affinity for cholesterol (discussed further in the sections below), an early marker of atherosclerosis [3,4]. Endothelial dysfunction results in the expression of leukocyte adhesion molecules that promote adherence of blood monocytes to the endothelial layer where once attached, chemokines promote monocyte migration into the subendothelial space[5]. Once in the intima, monocytes differentiate into macrophages that accumulate and transform into lipid enriched foam cells – the hallmark of atherosclerotic lesions[6]. Foam cells signal to other cells through pro-inflammatory cytokines causing inflammation and release of growth factors at the site of endothelial injury, resulting in smooth muscle cell migration and proliferation in the intima[3]. Simultaneously, once localised at the arterial wall, T cells activated by macrophages release inflammatory cytokines and activate endothelial cells to attract more white blood cells This activation contributes to an increase in foam cell accumulation, lipid content, and foam cell death with an increase in cell debris in the developing atherosclerotic lesion (plaque)[4,7]. Slowly growing plaques, characterised by a small lipid core and a thick fibrous cap providing structural integrity, gradually

expand due to the accumulation of lipids in foam cells, smooth muscle cells and cell debris, tend to stabilise and are not prone to rupture[8]. In coronary arteries, stable plaques remain largely clinically silent or in the long term may lead to stable angina due to blood flow limitation. In contrast, unstable plaques are more prone to rupture. These plaques are characterised by large lipid cores, a thin fibrous cap, and an abundance of inflammatory cells[2]. Unstable plaques grow more rapidly due to more rapid lipid deposition and are able to invade the arterial lumen, obstruct blood flow and lead to tissue ischemia[9,10]. Atheromata that do not produce obstruction limiting blood flow can lead to acute occlusion of a vessel and precipitate an acute coronary event triggered by plaque rupture and thrombosis[3].

In its early stages, beginning in the second decade of life, atherosclerosis is asymptomatic, has a long latency period of many years, and often co-exists in more than one vascular bed[2]. The prevalence and extent of atherosclerosis increases with age, though the speed of progression and severity of atherosclerosis can depend on the affected site and degree of arterial occlusion[3,11]. Atherosclerotic disease of the coronary arteries may result in clinical manifestations such as myocardial infarction (MI) or angina, leading to the need for revascularisation procedures (together termed coronary heart disease; CHD). Atherosclerosis in other territories can lead to peripheral artery and renovascular disease, and stroke. Collectively CHD and clinical disease in other territories is termed cardiovascular disease (CVD) and many CVD events e.g. stroke or MI can be fatal[12]. Outcomes for patients presenting to health systems with an acute manifestation of atherosclerosis are positive, with the majority surviving largely due to the progress made in cardiovascular interventions, management, and clinical application of scientific discoveries yielding beneficial

outcomes for patients[3,12]. Despite these successes, atherosclerotic cardiovascular disease remains a major cause of mortality worldwide. Moreover, increases in the number of people surviving an initial event contribute to the global burden of CVD morbidity e.g., due to heart failure, arrythmia or physical incapacity from disease, and residual CVD risk among these individuals remains high. The long latency pre-clinical phase of atherosclerosis offers an opportunity for prediction and prevention of disease. The rest of this chapter discusses the epidemiology and risk factors for atherosclerotic CVD.

1.2 Epidemiology of Cardiovascular Disease

Cardiovascular disease is a leading cause of morbidity and pre-mature mortality worldwide[13]. In 2017 the World Health Organisation (WHO) estimated 17.9 million deaths due to CVD related illness, of which 85% were attributable to myocardial infarction (MI) and stroke[14]. The Global Burden of Disease report an age-standardised mortality rate of 278 per 100,000 per year, with higher income countries experiencing more positive outcomes compared to lower income countries[15,16]. Advances in public health interventions, drug treatments and adoption of new interventional technologies (e.g. coronary angioplasty and stenting) have resulted in a decline in mortality rates since the mid 1990s and a consequential rise of prevalent CVD[17]. More individuals who now survive an initial CVD event live longer, only to suffer the consequences of atherosclerosis later in life. Such individuals require long-term treatment and management (e.g. with antiplatelet drugs, statins and medication for heart failure) to control symptoms, reduce subsequent events and prevent premature mortality[6].

1.3 Modifiable risk factors for cardiovascular disease

Epidemiological studies have played an important role in elucidating the factors that predispose CVD, signposting opportunities for disease prevention. Effective prevention and management of CVD is based on knowledge of the important risk factors that have a cumulative effect throughout life. Evidence from prospective observational studies, such as the Framingham Heart Study (FHS)[18] and many others, and case-control studies such as INTERHEART[19,20] have identified modifiable metabolic, social and dietary risk factors for CVD. The findings from the FHS and many other studies have been synthesised using meta-analysis through efforts of the Prospective Studies Collaboration and the Emerging Risk Factors Collaborations, these are discussed later in this chapter. The FHS, considered one the most influential investigations of CVD, is a prospective long-term study that began in 1948 with 5209 adults and is now on its third generation of participants[18]. Monitoring of the FHS population initially led to the identification of the major CVD risk factors, high blood pressure, high blood cholesterol, smoking, obesity, and diabetes. The INTERHEART study examined 15,152 cases and 14,820 age- and sex-matched controls in 52 countries and report 9 modifiable risk factors account for 90% of acute myocardial infarction in men and women across all ages and ethnic groups[20]. The INTERSTROKE case–control study of 13,447 cases and 13,472 age- and sex-matched controls in 32 countries also demonstrates 91% of stroke burden is attributable to the same 9 modifiable risk factors, with the addition of cardiac causes (such as atrial fibrillation)[21]. The identified risk factors were, abnormal lipids defined as elevated low-density lipoprotein cholesterol (LDL-C), low high-density lipoprotein cholesterol (HDL-C) and elevated triglyceride (TG), and elevated blood pressure, tobacco smoking, raised blood glucose, abdominal obesity, physical

inactivity, unhealthy diet, no alcohol consumption, and psychosocial stress. The study further identified low income and poor education associations with increased CVD mortality as well as, increased risk of CVD risk factors tobacco smoking, obesity, and elevated blood pressure. Prospective cohort studies such as the FHS are better equipped to minimise reverse causation and recall biases compared to case-control studies. Differential recall between cases and controls may contribute to an overestimation of the 90% population attributable risk (PAR) reported in INTERHEART[22]. While INTERHEART and INTERSTROKE took measures to minimise such bias (e.g., recruiting appropriate 'at risk of outcome' controls from the study centres instead of from the general population), this should be a consideration when interpreting the results. Moreover, INTERHEART use self-reported hypertension data as opposed to continuous blood pressure values, which may underestimate the contribution of this risk factor. This was assessed in an examination of the risk factors identified in INTERHEART to population attributable (PAR) risk of 10 year follow up of MI in the prospective cohort British Regional Heart Study (BRHS)[22]. Whincup and colleagues observed a much higher contribution of diastolic blood pressure to PAR in BRHS (54%) than in INTERHEART (23%). However, when risk factors were taken together, the contributions of total cholesterol, diastolic blood pressure and smoking (including current, ex and passive) account for 91% of PAR in BRHS, very close to the 90% PAR observed in INTERHEART[20–22]. Of the considered modifiable risk factors, the Comparative Risk Assessment study found 80% of CHD deaths and 70% of stroke deaths were attributable to the joint effect of tobacco smoking, elevated blood pressure and raised serum cholesterol, these are discussed further below[13,16].

Substantial evidence from studies such as the seminal British Doctors study and others[23–25] have confirmed the association of smoking in CVD. In 2011, a meta-analysis of 75 studies concluded smoking increased the risk of CHD in men by 72% and in women by 92%. Further evidence has demonstrated that smoking cessation by age 50 reduces mortality risk by 50%, and cessation by age 30 avoids adverse smoking effects almost entirely[26].

Higher blood pressure is associated with increased vascular disease risk. Meta-analysis data from over 1 million participants has demonstrated a log-linear association of blood pressure with CVD mortality such that, a 20mmHg higher systolic blood pressure (SBP), and a 10mmHg higher diastolic blood pressure (DBP) is associated with a two-fold increased vascular disease[27]. Triangulation of evidence from observational studies with the findings from Mendelian randomisation (MR) studies and randomised trials of blood pressure lowering drugs, provide strong evidence that elevated blood pressure is causally related to increased risk of CHD and stroke[28,29].

1.4 The role of lipoproteins and lipids in atherosclerotic cardiovascular disease

This section discusses lipoprotein particle size, density, composition, and lipoprotein particle relationship with CVD. The subsequent section covers traditional methods of lipoprotein measurement, lipid content estimation, and more recent advancements for quantitative measurement of lipoprotein lipid particle size and content using high-throughput serum ^{1}H-Nuclear Magnetic Resonance (NMR) spectroscopy.

1.4.1 Lipoprotein Composition

The circulating blood lipids mainly comprise triglyceride (TG) and cholesterol. These are insoluble in water and must be transported within membrane bound lipoprotein particles (lipoproteins) which also contain certain proteins with enzymatic or targeting function. The function and role of TG and cholesterol in atherosclerosis are discussed further in sections below. Lipoproteins play a key role in dietary lipid transport and absorption by the small intestine, in the transport of lipids from liver to tissues, and the transport of lipids from peripheral tissues to the liver and intestine (reverse cholesterol transport)[30]. Lipoproteins are complex particles comprising a hydrophobic lipid-rich core and a hydrophilic phospholipid outer layer containing surface apolipoprotein molecules required for stabilisation and metabolism[31], see figure 1.1. The main lipoprotein groups are traditionally divided into seven classes, categorised based on the relative densities of the aggregates from ultracentrifugation as; chylomicrons (CMR), CMR remnants, very-low density lipoproteins (VLDL), intermediate density lipoproteins (IDL), low-density lipoproteins (LDL), high-density lipoproteins (HDL), and Lp (a)[32]. In the context of this thesis, when referring to the lipoprotein particle, the lipid suffix is omitted (e.g., VLDL, IDL, LDL or HDL particle). The size, composition of the lipid core, and protein components of the outer membrane varies depending on the lipoprotein, see table 1. When referring to the lipid content of a particular lipoprotein particle, lipid suffix is added, e.g., LDL = low density lipoprotein particles, and LDL-C = the cholesterol (C) content of low-density lipoprotein particles. Lipoprotein lipid content and metabolism is discussed below.

Figure 1.1 Lipoprotein structure

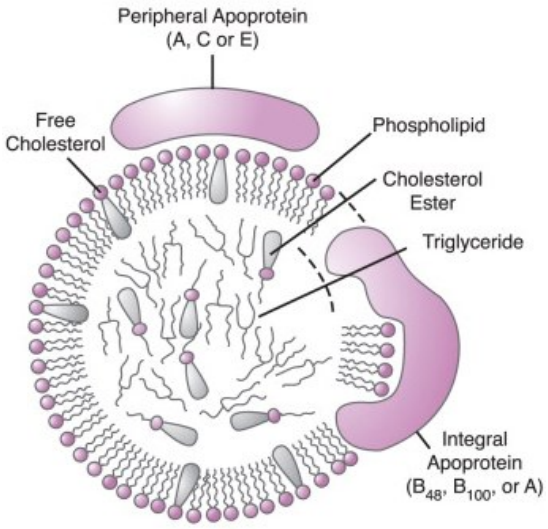


Figure adapted from; *Hegle R. A. (2013), in Emery and Rimoin's Principles and Practice of Medical Genetics*

Table 1.1 Lipoprotein classes

Lipoprotein	Density (g/ml)	Size (nm)	Major lipids	Major Apoproteins
Chylomicrons	<0.930	75-1200	Triglycerides	Apo B-48, Apo C, Apo E, Apo A-I, A-II, A-IV
Chylomicron remnants	0.930-1.006	30-80	Triglycerides Cholesterol	Apo B-48, Apo E
VLDL	0.930-1.006	30-80	Triglycerides	Apo B-100, Apo E, Apo C
IDL	1.006-1.019	25-35	Triglycerides Cholesterol	Apo B-100, Apo E, Apo C
LDL	1.019-1.063	18- 25	Cholesterol	Apo B-100
HDL	1.063-1.210	5- 12	Cholesterol Phospholipids	Apo A-I, Apo A-II, Apo C, Apo E
Lp(a)	1.055-1.085	~30	Cholesterol	Apo B-100, Apo (a)

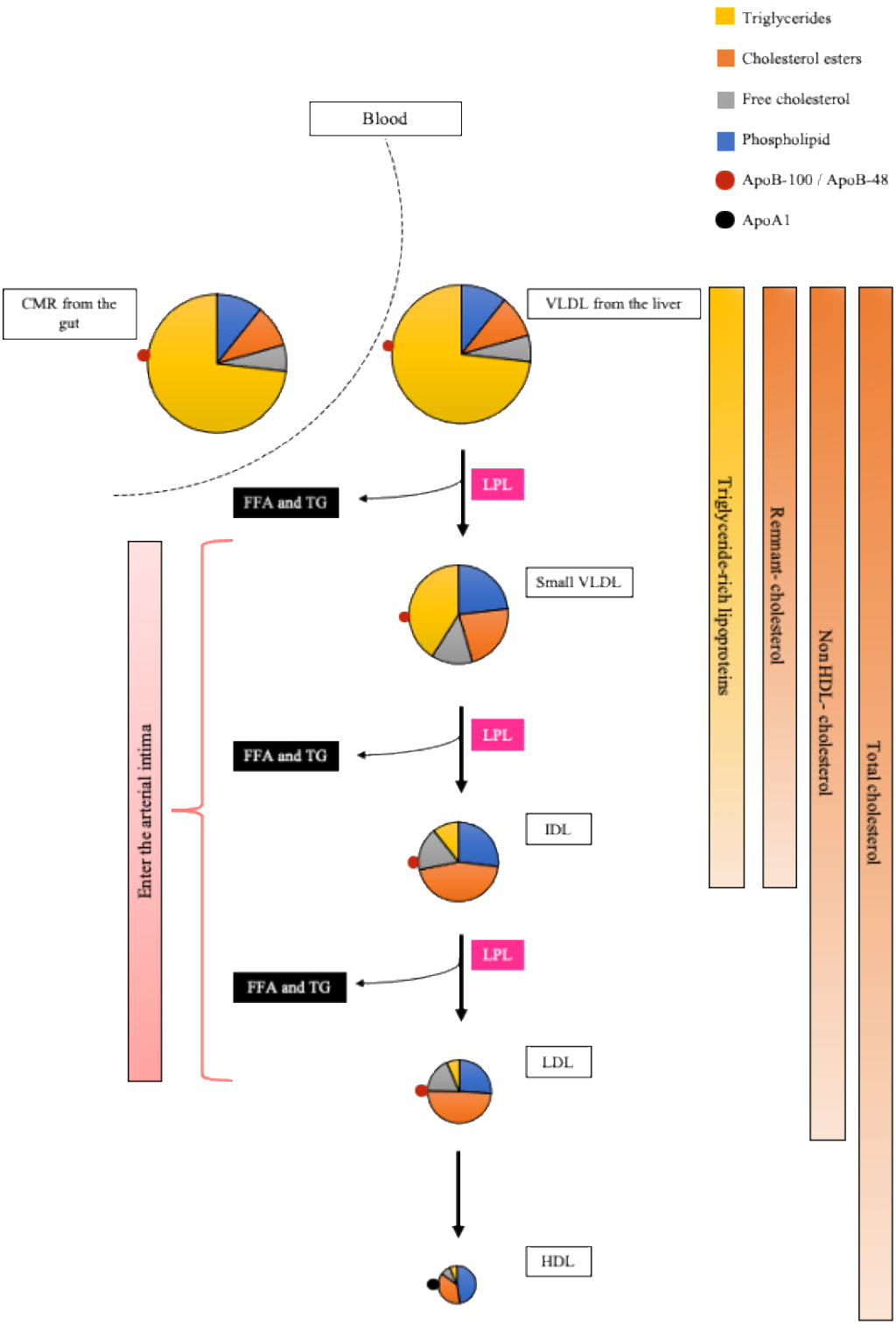
Chylomicrons (CMR) are large TG rich particles made by the intestine involved in the transport of dietary TG and cholesterol to the peripheral tissues and liver. Each CMR has one ApoB-48 as the core structural protein. Chylomicron size (75-1200 nm) is dependent on the amount of fat consumed in the diet[33]. A high fat meal increases the amount of TG transported and the formation of large CMR particles, whereas in the fasting state, CMR are relatively smaller and carry decreased quantities of TG. Triglycerides are liberated from CMR by lipoprotein lipase (LPL) in peripheral tissues[33]. The lipolysis products, fatty acids and glycerol, are stored in adipose and muscle tissues and the resulting TG-depleted remnants are taken up by the liver. Chylomicron remnants, enriched in cholesterol are considered to be pro-atherogenic[34]. Very-low density lipoproteins produced by the liver are TG-rich and contain one core structural ApoB-100 molecule[33]. VLDL particle size (30-80 nm) varies depending on the quantity of TG carried. Further removal of TG from VLDL by LPL in muscle and adipose tissue results in pro-atherogenic VLDL-remnants and IDL particles (25-35 nm) that are enriched in cholesterol and contain the ApoB-100 isotope and ApoE[33]. As most TG has been removed, the lipoprotein becomes denser and is referred to as a low-density lipoprotein (LDL) that contains one ApoB-100 molecule and carry most of circulatory cholesterol[33]. Smaller, more dense LDL particles have a lower affinity for the LDL receptor causing prolonged retention in the circulation and are considered to be more atherogenic than larger LDL[33,34]. The smaller LDL particles can more easily enter the arterial wall and bind to proteoglycans trapping them. This in turn increases susceptibility to oxidation and enhanced uptake by macrophages, to promote foam cell formation and induce inflammation, initiating the development and progression of atherosclerotic lesions[1,5]. The accumulation of cholesterol in the arterial wall is a slow process over

many decades and is accelerated by CVD risk factors, discussed in preceding sections. It is important to recognise that a VLDL particle, a remnant VLDL particle, an IDL and LDL particle are different names for the same circulating apoB lipoproteins at different stages in its lifecycle, depending on the lipid content it carries. HDL particles may contain multiple structural ApoA-I proteins and are enriched in cholesterol and phospholipids[33]. HDL play an important role in reverse cholesterol transport from peripheral tissues to the liver, which is considered one of the potential mechanisms by which HDL may be anti-atherogenic[35,36]. HDL particles have been shown to have antioxidant, anti-inflammatory, anti-thrombotic and anti-apoptotic properties, which has triggered extensive research in the ability of raising HDL-C to inhibit atherosclerosis. Lp (a) is an oxidised LDL particle attached to ApoB-100 via a disulphide bond.

1.4.2 Lipoprotein lipid terminology

In the current literature, CMR, VLDL, VLDL-remnants and IDL particles are termed triglyceride-rich lipoproteins (TRL). Hydrolysis of TG in TRL via LPL results in TRL enriched in cholesterol. Remnant cholesterol (RC) refers to the cholesterol content of VLDL and IDL and can be estimated from a standard lipid profile using the equation; RC (mmol/L) = total cholesterol minus LDL-C minus HDL-C. Non-HDL cholesterol refers to the cholesterol content in all atherogenic apoB containing lipoproteins (VLDL, IDL and LDL) and is estimated from a standard lipid profile as: non HDL-cholesterol (mmol/L) = total cholesterol minus HDL-C, see figure 1.2 for graphical depiction and succeeding sections for discussion on lipoprotein lipid measurement.

Figure 1.2 Graphical depiction of lipoprotein subfractions and lipid content



LPL = lipoprotein lipase; FFA = Free fatty acids; TG = triglycerides

1.4.3 Lipoprotein composition measurement and lipid content quantification

Early separation of lipoproteins was performed using electrophoresis and later superseded by preparative ultracentrifugation with quantitative clinical chemical measurement[37]. These techniques contributed to early animal and human insights on the relationship between blood lipids and CVD, which were based on lipoprotein lipid content initially beginning with the measurement of total cholesterol concentrations, across all lipoprotein particles[38]. Using analytical and preparative ultracentrifugation McFarlane isolated the 'X-protein', later renamed LDL[39], and Gofman observed different species within the α- and β-lipoproteins, changing the nomenclature of lipoproteins to VLDL, LDL, and HDL to reflect the different density regions[38]. Isolation of lipoproteins and early observations of the relationship between total cholesterol and atherosclerosis promoted the establishment of longitudinal cardiovascular epidemiologic programmes such as the FHS[37] and led to the observation of positive and negative associations with CHD of LDL-C and HDL-C respectively.

Early investigations of lipid associations with CVD began with total cholesterol concentrations measured as a summation across all lipoprotein subclasses. Advancement of lipoprotein separation methods aided the investigations of HDL and LDL-C, and finally total TG concentrations with disease. Clinical measurement of total cholesterol concentrations became readily available in the late 1970s with the advent of enzymatic-based reagents, which superseded previous methods such as analytical ultracentrifugation, Cohn factorisation and electrophoresis[38,40]. Enzymatic ultracentrifugation paved the way for automation, and the simplified methodology became the basis for TRL, and for beta-quantification of LDL and HDL in clinical

laboratories. Methods of LDL-C measurement comprise non-direct methods including ultracentrifugation and electrophoresis, and direct methods such as chemical precipitation, immuno-separation, and homogenous assay methods[37,41]. The most common method for estimating LDL-C in clinical laboratories is using the Friedewald equation; LDL-cholesterol (mmol/L); total cholesterol minus HDL-cholesterol minus TG concentrations/2.2[42]. More recent reports since the 1990s dispute the estimation of LDL-C using the Friedewald equation as it may not be sufficiently accurate at high TG concentrations or non-fasting assay samples[43]. Moreover, the Friedewald estimation method is nonspecific to LDL-C and includes cholesterol carried in IDL and some VLDL particles[44]. Similar to the Friedewald estimation, precipitation estimations of LDL-C from beta-quantification of lipoproteins also includes IDL-cholesterol in reported LDL-C values. Recent reports suggest remnant cholesterol (RC) is associated with increased CVD, however, there is debate around using RC as a lipid measure because it is estimated using a standard lipid profile, in which LDL-C is often inaccurately estimated dependant on the LDL-C assay employed as described above[44,45]. Triglycerides can be measured using direct and indirect methods in the clinical laborotory. Indirect estimations are calculated from the difference between serum concentrations of total fatty acids and concentration of cholesterol and phospholipid fatty acid esters[37]. Direct methods are relatively more preciese and include flurommetric, colorimetric and enzymatic estimation.

Lipoproteins can be further divided into subclasses to gain a more granular understanding of the association of lipoprotein size and density, lipid composition, and particle number with CVD. Recent evidence suggests individuals with

predominantly small LDL particles have a higher risk of CVD than those with large LDL, making accurate quantification of lipoprotein subclasses essential for CVD prevention and diagnosis[46]. Lipoprotein subclassification can be achieved by size and density specific of measurement of lipoproteins using high-performance liquid chromatography (HPLC)[47], mass spectroscopy[46], or nuclear magnetic resonance (NMR) spectroscopy[48]. Detailed quantification of lipoprotein subclass subfractions and their contents is made available by Nuclear Magnetic Resonance (NMR) spectroscopy. NMR spectroscopy offers the opportunity to overcome lipid estimation issues and interrogate the relationship between the lipid content of lipoproteins and CVD, further to what has been done before. NMR spectroscopy is discussed further below[44].

1.5 Nuclear magnetic resonance technology for the classification of lipid content of lipoproteins

Nuclear magnetic resonance was first used to study lipoprotein structure in the 1960s[49]. NMR for the separation and measurement of lipoprotein subclasses was first reported by Otvos[41] and Ala-Korpela[50]. Recent developments in quantitative profiling technologies and appealing results from application in understanding health and disease has made NMR metabolomics more common in epidemiology. The growth in NMR metabolomics reflects the advantages over MS and HPLC[51]. While MS may arguably be more sensitive, NMR spectroscopy is highly automatable, reproducible, easily quantifiable and requires little or no sample treatment or chemical derivatisation, making it the preferred technique for high throughput of plasma or serum samples in large-scale population studies[52]. The high-throughput proton (^{1}H) NMR metabolomics assay developed by Nightingale[53] provides

quantitative information on 220 measures per sample including 14 lipoprotein subclasses, 6 for VLDL, 1 for IDL, 3 for LDL and 4 for HDL, with molecular information on cholesterol, TG, and phospholipid concentration in each subclass, as well as fatty acids (e.g. ω-3 and ω-6 fatty acids) and other lipid species. Nightingale's broad biomarker analysis also measures numerous low molecular-weight metabolites (including amino acids, glycolysis-related measures and ketone bodies), as well as measures of glycoprotein acetylation (GlycA) and apolipoproteins A-I and B[53]. The lipoproteins are quantified based on the different chemical compositions and size, which experience different magnetic susceptibility[52]. This gives rise to distinctive NMR signals that are determined by the electron density and rotational diffusion of lipoprotein vehicles[54]. In particular, the methyl (–CH3) signals arising from large and less dense particles (i.e. VLDL, IDL and LDL) are different in shape and resonate at lower field strength (higher frequency) than the lipid signals emitted by smaller lipoproteins (i.e. HDL)[50]. Due to the overlapping signals of the lipid methylene and methyl envelopes, NMR quantification of lipoproteins requires a calibration step, such as the ultracentrifugation method of lipoprotein separation using Partial Least Squares (PLS) regression[55]. Very low-density lipoprotein is divided into six sizes, the largest is extremely large VLDL (XXL-VLDL), with a mean particle diameter of 75nm or more, and five remaining VLDL subclasses with decreasing mean particle diameters as; extra-large (XL-VLDL, average diameter 65nm), large (L-VLDL, 54nm), medium (M-VLDL, 44.5nm), small (S-VLDL, 36.8nm) and extra-small (XS-VLDL, 31.3nm). IDL (average diameter 28.6nm), LDL is divided as large (average diameter, 25.5nm), medium (23nm) and small LDL (18.7nm). HDL into four subfractions, very-large (XL-HDL, 14.3nm) large (L-HDL, 12.1nm), medium (M-HDL, 10.9nm), and small (S-HDL, 8.7nm[53]). The lipid content

of lipoproteins are in a constant state of flux, however in general the larger, less dense particles are triglyceride-rich and the more dense, smaller lipoproteins are cholesterol-rich. The composition, role and function of cholesterol and TG within lipoproteins are discussed in the sections below.

1.6 The role of cholesterol in atherosclerotic cardiovascular disease

1.6.1 Composition and metabolism

Cholesterol functions as a vital structural component for cell membranes. It is converted to steroid hormones oestrogen and testosterone, and to bile acids in the liver where bile salts emulsify dietary fat to make it absorbable[5]. As discussed in preceding sections, LDL, the major lipoprotein carrier of cholesterol, accumulates in the intima and stimulates the expression of adhesion molecules on the surface of endothelial cells, monocyte migration and differentiation into macrophages that accumulate cholesterol and eventually become foam cells[2]. Early animal model studies by Anitschkov demonstrated a causal role of cholesterol in the pathogenesis of atherosclerosis and later, Muller described families with high cholesterol and increased CVD[35,56]. Several decades later, through their discoveries of the LDL-receptor and oxidised-LDL, Brown and Goldstein, and Steinbrecher et al have revolutionised knowledge about the regulation and metabolism of cholesterol and the treatment of diseases caused by abnormally elevated LDL particles and LDL-cholesterol (LDL-C; the cholesterol carried in LDL particles) concentrations in the circulation[9,57].

1.6.2 The association of cholesterol with atherosclerotic cardiovascular disease

The European Atherosclerosis Society Consensus Panel[58] and the Emerging Risk Factors Collaboration[59] have appraised clinical and genetic evidence that indicates elevated LDL-C as causal in CVD. Over 200 prospective observational studies have been meta-analysed to quantify positive log-linear associations of LDL-C or non-HDL cholesterol (a surrogate for LDL-C) and heart disease. Randomised trials of more than 2 million participants with over 20 million-person years of follow-up and more than 150,000 cardiovascular events demonstrate consistent dose-dependent, log-linear association of exposure to higher circulating LDL-C and risk of CVD, with the risk appearing to increase with increasing exposure duration[58]. Moreover, evidence from statin trials have demonstrated for each 1.0 mmol/L reduction in LDL-C, coronary event risk is reduced by 25% (RR: 0.76, 95% CI 0.73 to 0.79) and ischemic stroke by 20% (RR: 0.80, 95% CI 0.74 to 0.86). Evidence from Mendelian randomisation (MR) studies (discussed further below) show single nucleotide polymorphisms (SNPs) associated with higher plasma LDL-C exhibit allele dose-dependent increase in risk for CVD whereas, SNPs associated with lower LDL-C levels are also associated with a lower risk of CVD. Triangulation of observational and randomised trial findings with more recent MR studies have demonstrated LDL-C lowering reduces CVD events. The compelling evidence has surmounted in LDL-C to be regarded as a causal risk factor in atherosclerosis development and progression, and LDL-C lowering forms the mainstay of primary and secondary prevention of CVD.

Given the success in lowering and achieving clinical LDL-C concentration targets to prevent primary and secondary CVD events, emerging research in the last

10 years has turned the focus to remnant cholesterol and non-HDL cholesterol. Clinical evidence suggests the residual cardiovascular risk observed in patients with well-controlled LDL-C might be explained in part by risk factors including the cholesterol content of remnant of TRL, also called remnant cholesterol (RC)[60,61]. Observational and genetic studies have suggested RC as a causal risk factor for CHD[62,63]. The mechanism by which RC causes atherosclerosis is different to that of LDL-C, as the latter requires modification prior to uptake by macrophages. Unlike LDL, remnant particles pass directly into the arterial all and are taken up by macrophages and smooth muscle cells[62]. Remnant particles are also larger than LDL and carry up to 40 times more cholesterol per particle, potentially making them more atherogenic[62]. Similar to RC, several studies have shown non-HDL-C levels, which recapitulates the cholesterol content of all atherogenic apoB containing lipoproteins, are a better predictor of cardiovascular risk compared to LDL-C concentrations alone[64,65]. Mendelian randomisation studies have also shown the clinical benefit of lowering LDL-C may be better explained by the absolute reduction in apoB-containing lipoproteins, which represent the total number of atherogenic lipoproteins irrespective of lipid content[66,67].

While the evidence of positive associations of LDL-C and atherosclerosis is vast and convincing, the role of cholesterol within HDL is more uncertain. HDL-C is inversely associated with risk of CVD and was considered a key component of predicting cardiovascular risk. The first compelling reports of a strong inverse association of HDL-C and CHD were described in the Framingham Heart Study[18]. Observational data from Framingham and many others formed the view of HDL-C as the 'good cholesterol'[8], the concept of reverse cholesterol transport[68], and led

investigators to hypothesise that raising HDL-C would reduce risk of CVD. The HDL-C hypothesis was further reinforced by pre-clinical data in animal studies, in which HDL-C infused rabbits exhibited inhibition of atherosclerosis, making HDL-C a target for novel therapeutic approaches[69]. However, despite consistent associations with atheroprotection, the casual relationship between HDL-C and atherosclerosis is uncertain. Several randomised clinical trials of HDL-C-raising drugs have failed to show a reduction in CVD events or were terminated early due to increased CHD events and total mortality in patients randomised to treatment[36,70–73]. The REVEAL IT trial of Anacetrapib[74], a CETP inhibitor, which is a protein that facilitates the transfer of cholesterol and TG between HDL and atherogenic particles in the blood to lower TG and LDL-C and raise HDL-C, report patients in the treatment arm had fewer first major coronary events at the end of the follow-up period compared to patients in the placebo arm [RR: 0.91 (95% CI 0.85 to 0.97)]. Despite the on-treatment reduction in major coronary events as compared to placebo, anacetrapib has not been pursued for commercial reasons. It is difficult to ascertain if failed trials are due to failure of a compound or the biomarker. Drugs specifically target a single protein in a biochemical pathway regulating the level of a biomarker, and so it may be difficult to distinguish if negative findings from drug trials or meta-analyses are reflective of a failure of a compound e.g. niacin (where the solution is to develop a more effective compound against the same protein target), or of a drug target biomarker i.e. HDL-C (where the solution is to develop a drug molecule that alters the same biomarker but through a different protein target), or failure of the biomarker (redirect efforts to a different biomarker). Moreover, if trials are successful, it is difficult to infer whether this is a target specific effect or if it is mediated through HDL-C alone or via additional effects on TG, or LDL-C, as unlike

most LDL-C lowering drugs, HDL-C elevating drugs have effects on other major lipid fractions. Results from MR studies of HDL-C using genome wide instruments remain equivocal as causal in risk of CVD[75,76], but this does not preclude the possibility that raising HDL-C through a target such as CETP might be beneficial for CVD outcomes.

1.7 The role of triglycerides in atherosclerotic cardiovascular disease

1.7.1 Composition and metabolism

Large amounts of fatty acids from meals are transported as TG to avoid toxicity[33]. Triglycerides are an insoluble key energy source made up of three free fatty acids ester-linked to a glycerol backbone. The exogenous source of TG originates from diet, while endogenous TG are synthesised in intestinal and liver cells[33,77]. Triglycerides play an essential part in intestinal lipid absorption via chylomicrons and apoB48, and endogenously in VLDL requiring apoB100[33] (see figure X). Circulating TRL exchange cholesterol esters with other particles through CETP and lecithin-choline acetyl transferase (LCAT)[64]. Lipolysis of CMR and VLDL releasing free fatty acids by endothelial bound LPL, yield CMR remnants and IDL that are taken up in the liver by hepatic LDL receptors[34,68]. In hepatocytes the TG is packaged with cholesterol and apoB-100 isoform into VLDL and released into the blood where after additional hydrolysis by lipases, free fatty acids and VLDL remnants (IDL) are released to yeild LDL[31]. Experimental studies have shown that these VLDL remnant particles deliver cholesterol into the intima which may sequester more atherogenic LDL-C[45]. Trigylceride concentrations in lipoproteins have traditionally been assessed under fasting conditions in the clinic due to the postprandial biological variablilty of TG. Recent changes to clinical guidelines

endorse non-fasting lipid measures to better reflect dietry habits and time spent in the prolonged post-prandial state throughout the day[78].

1.7.2 Association of triglycerides with cardiovascular disease

Triglycerides and HDL-C are inversely correlated with one another, with other CVD risk factors, and are both associated with CVD, TG positively and HDL-C inversely. As discussed previously, HDL-C exhibits inverse associations both with CVD and TG, whereas TG demonstrate positive associations[59]. Observational associations for both TG and HDL-C attenuate following statistical adjustment for one another, and for other lipid fractions such as LDL-C, making their role in atherosclerotic disease formation contentious. With the exception of a recently published trial[79,80], randomised trials to lower triglycerides (e.g. using fibrates to target PPAR-α) have largely been unsuccessful. Other RCTs of drugs to lower TG through other drug targets could help to resolve the uncertainty, but RCTs are expensive, long in duration and have a high failure rate. Thus, despite the recent emerging evidence from the REVEAL[74] and REDUCE-IT[80] trials, neither HDL-C elevating nor TG lowering therapies are routinely used in CVD prevention.

Conversely, insights from genetic studies suggest that elevated TG is a causal risk factor for CVD. Mutations in specific genes such as *LPL*, *APOC2*, *APOA5* or *LMFI* can increase TG concentrations[81]. A mutation in the *LPL* gene resulting in lipoprotein lipase deficiency (the protein responsible for plasma triglyceride degradation), causes high TG concentrations and is associated with an increased risk for CVD[82]. Relatively recent studies investigating mutations in the *APOA5*, *APOC3*

and *ANGPTL3* genes also support the hypothesis that elevated TG is related to CVD.[83]

The evidence presented here represents the association of total TG concentrations in the blood, which is the sum of TG in nascent VLDL, VLDL remnants in a fasting state, in CMR and CMR remnants in the postprandial state. It is not the association of TG concentration in a specific lipoprotein subfraction. Quantifying TG concentrations across 14 lipoprotein subfractions using NMR spectroscopy offers a novel way to further interrogate the association of TG within specific lipoprotein particles and CVD. The next section is a review of the epidemiological and genetic literature on association of TG with CVD to outline the rationale for the specific aims of this PhD.

To this end, the following aims of this thesis were therefore to:

i. Establish the distribution, determinants and reference range intervals for triglyceride-containing and cholesterol containing lipoprotein subfraction metabolites measured using NMR methods in the population

ii. Investigate (non-genetic) observational associations of the 14-triglyceride containing lipoprotein subfractions with CHD and stroke

iii. Compare and evaluate 14-triglyceride and 14-cholesterol lipoprotein subfraction observational effects with CHD to identify the critical atherogenic lipoprotein and lipid components

iv. Identify genetic instruments for triglyceride and cholesterol containing lipoprotein subfractions and use these to undertake multivariable Mendelian

randomisation analysis to explore which lipid trait predominates as causal in CHD.

1.8 References

1. Moore, K. J. & Tabas, I. Macrophages in the pathogenesis of atherosclerosis. *Cell* vol. 145 341–355 (2011).
2. Libby, P., Ridker, P. M. & Maseri, A. Inflammation and atherosclerosis. *Circulation* **105**, 1135–43 (2002).
3. Libby, P. *et al.* Atherosclerosis. *Nat. Rev. Dis. Prim.* **5**, 1–18 (2019).
4. Ross, R. Cell biology of atherosclerosis. *Annu. Rev. Physiol.* **57**, 791–804 (1995).
5. Weber, C. & Noels, H. Atherosclerosis: Current pathogenesis and therapeutic options. *Nature Medicine* vol. 17 1410–1422 (2011).
6. Libby, P. The changing landscape of atherosclerosis. *524 | Nat. |* **592**, (2021).
7. Bergheanu, S. C., Bodde, M. C. & Jukema, J. W. Pathophysiology and treatment of atherosclerosis. *Netherlands Hear. J.* **25**, 231–242 (2017).
8. Toth, P. P. The "good cholesterol" high-density lipoprotein. *Circulation* **111**, e89–e91 (2005).
9. Goldstein, J. L. & Brown, M. S. The LDL receptor. *Arteriosclerosis, Thrombosis, and Vascular Biology* vol. 29 431–438 (2009).
10. Finn, A. V., Nakano, M., Narula, J., Kolodgie, F. D. & Virmani, R. Concept of vulnerable/unstable plaque. *Arteriosclerosis, Thrombosis, and Vascular Biology* vol. 30 1282–1292 (2010).
11. Toth, P. P. Subclinical atherosclerosis: What it is, what it means and what we can do about it. *International Journal of Clinical Practice* vol. 62 1246–1254 (2008).
12. Herrington, W., Lacey, B., Sherliker, P., Armitage, J. & Lewington, S. Epidemiology of Atherosclerosis and the Potential to Reduce the Global Burden of Atherothrombotic Disease. *Circ. Res.* **118**, 535–546 (2016).
13. Joseph, P. *et al.* Reducing the Global Burden of Cardiovascular Disease, Part 1. *Circ. Res.* **121**, 677–694 (2017).
14. Shing, C. *et al. Self-reported sleep duration and quality and cardiovascular disease and mortality: a dose-response meta-analysis Running title: Sleep duration and quality and CVD and mortality*. http://eprints.keele.ac.uk/4932/1/20180521_ccg_Manuscript final.pdf.
15. Dicker, D. *et al.* Global, regional, and national age-sex-specific mortality and life expectancy, 1950-2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1684–1735 (2018).
16. Bansilal, S., Castellano, J. M. & Fuster, V. Global burden of CVD: Focus on secondary prevention of cardiovascular disease. *Int. J. Cardiol.* **201**, S1–S7 (2015).
17. Ballantyne, C., Arroll, B. & Shepherd, J. Lipids and CVD management: towards a global consensus. *Eur. Heart J.* **26**, 2224–2231 (2005).
18. Tsao, C. W. & Vasan, R. S. Cohort Profile: The Framingham Heart Study (FHS): overview of milestones in cardiovascular epidemiology. *Int. J.*

Epidemiol. **44**, 1800–1813 (2015).
19. Rosengren, A. *et al.* Association of psychosocial risk factors with risk of acute myocardial infarction in 11 119 cases and 13 648 controls from 52 countries (the INTERHEART study): Case-control study. *Lancet* **364**, 953–962 (2004).
20. Yusuf, P. S. *et al.* Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): Case-control study. *Lancet* **364**, 937–952 (2004).
21. O'Donnell, M. J. *et al.* Global and regional effects of potentially modifiable risk factors associated with acute stroke in 32 countries (INTERSTROKE): a case-control study. *Lancet* **388**, 761–775 (2016).
22. Whincup, P., Emberson, J., Morris, R. & Shaper pwhincup, A. G. *Correspondence*. http://www.ktl.fi/publications/ doi:10.1016/S0140-6736(05)17691-0.
23. Doll, R., Peto, R., Boreham, J. & Sutherland, I. Mortality from cancer in relation to smoking: 50 years observations on British doctors. *Br. J. Cancer* **92**, 426–9 (2005).
24. Pirie, K. *et al.* The 21st century hazards of smoking and benefits of stopping: a prospective study of one million women in the UK. *Lancet (London, England)* **381**, 133–41 (2013).
25. Lu, L., Mackay, D. F. & Pell, J. P. Meta-analysis of the association between cigarette smoking and peripheral arterial disease. *Heart* **100**, 414–423 (2014).
26. Huxley, R. R. & Woodward, M. Cigarette smoking as a risk factor for coronary heart disease in women compared with men: a systematic review and meta-analysis of prospective cohort studies. *Lancet* **378**, 1297–1305 (2011).
27. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet* **360**, 1903–1913 (2002).
28. Gill, D. *et al.* Genetically Predicted Midlife Blood Pressure and Coronary Artery Disease Risk: Mendelian Randomization Analysis. *J. Am. Heart Assoc.* **9**, e016773 (2020).
29. Ettehad, D. *et al.* Blood pressure lowering for prevention of cardiovascular disease and death: A systematic review and meta-analysis. *Lancet* **387**, 957–967 (2016).
30. The Role of Lipids and Lipoproteins in Atherosclerosis - Endotext - NCBI Bookshelf. https://www.ncbi.nlm.nih.gov/books/NBK343489/.
31. Hegele, R. A. Plasma lipoproteins: genetic influences and clinical implications. *Nat. Rev. Genet.* **10**, 109 (2009).
32. De Leon, H., Boue, S., Szostak, J., Peitsch, M. & Hoeng, J. *Systems Biology Research into Cardiovascular Disease: Contributions of Lipidomics-based Approaches to Biomarker Discovery*. *Current drug discovery technologies* vol. 12 (2015).
33. Feingold, K. R. & Grunfeld, C. *Introduction to Lipids and Lipoproteins*. *Endotext* (MDText.com, Inc., 2000).
34. Kindel, T., Lee, D. M. & Tso, P. The mechanism of the formation and

secretion of chylomicrons. *Atherosclerosis Supplements* vol. 11 11–16 (2010).
35. Linton, M. F. *et al. The Role of Lipids and Lipoproteins in Atherosclerosis*. *Endotext* (MDText.com, Inc., 2000).
36. Guyton, J. R. *et al.* Relationship of Lipoproteins to Cardiovascular Events. *J. Am. Coll. Cardiol.* **62**, 1580–1584 (2013).
37. McNamara, J. R., Warnick, G. R. & Cooper, G. R. A brief history of lipid and lipoprotein measurements and their contribution to clinical chemistry. *Clinica Chimica Acta* vol. 369 158–167 (2006).
38. Lindgren, F. T., Elliott, H. A. & Gofman, J. W. The ultracentrifugal characterization and isolation of human blood lipids and lipoproteins, with applications to the study of atherosclerosis. *J. Phys. Chem.* **55**, 80–93 (1951).
39. McFarlane, A. S. The ultracentrifugal protein sedimentation diagram of normal human, cow and horse serum. *Biochem. J.* **29**, 660 (1935).
40. McNamara, J. R. *et al.* Immunoseparation method for measuring low-density lipoprotein cholesterol directly from serum evaluated. *Clin. Chem.* **41**, 232–240 (1995).
41. Otvos, J. D., Jeyarajah, E. J. & Bennett, D. W. Quantification of plasma lipoproteins by proton nuclear magnetic resonance spectroscopy. *Clin. Chem.* **37**, 377–386 (1991).
42. Friedewald, W. T., Levy, R. I. & Fredrickson, D. S. Estimation of the Concentration of Low-Density Lipoprotein Cholesterol in Plasma, Without Use of the Preparative Ultracentrifuge. *Clin. Chem.* **18**, (1972).
43. Martin, S. S. *et al.* Friedewald-estimated versus directly measured low-density lipoprotein cholesterol and treatment implications. *J. Am. Coll. Cardiol.* **62**, 732–739 (2013).
44. Holmes, M. V. & Ala-Korpela, M. What is 'LDL cholesterol'? *Nature Reviews Cardiology* vol. 16 197–198 (2019).
45. Nordestgaard, B. G. & Varbo, A. Triglycerides and cardiovascular disease. *Lancet* **384**, 626–635 (2014).
46. Kulkarni, K. R., Garber, D. W., Jones, M. K. & Segrest, J. P. Identification and cholesterol quantification of low density lipoprotein subclasses in young adults by VAP-II methodology. *J. Lipid Res.* **36**, 2291–2302 (1995).
47. Jaroszewski, J. W. Hyphenated NMR, methods and applications. in *Encyclopedia of Spectroscopy and Spectrometry* 170–173 (Elsevier, 2016). doi:10.1016/B978-0-12-803224-4.00080-7.
48. Würtz, P. *et al.* High-throughput quantification of circulating metabolites improves prediction of subclinical atherosclerosis. *Eur. Heart J.* **33**, 2307–2316 (2012).
49. Steim, J. M., Edner, O. J. & Bargoot, F. G. Structure of human serum lipoproteins: nuclear magnetic resonance supports a micellar model. *Science (80-.).* **162**, 909–911 (1968).
50. Ala-Korpela, M. *et al.* A comparative study of 1H NMR lineshape fitting analyses and biochemical lipid analyses of the lipoprotein fractions VLDL, LDL and HDL, and total human blood plasm. *NMR Biomed.* **6**, 225–233

(1993).
51. Mora, S. *et al.* Lipoprotein Particle Profiles by Nuclear Magnetic Resonance Compared With Standard Lipids and Apolipoproteins in Predicting Incident Cardiovascular Disease in Women. *Circulation* **119**, 931–939 (2009).
52. Würtz, P. *et al.* Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Large-Scale Epidemiology: A Primer on -Omic Technologies. *Am. J. Epidemiol.* **186**, 1084–1096 (2017).
53. Lifelong health belongs to everyone. https://nightingalehealth.com/.
54. Mihaleva, V. V. *et al.* Automated quantum mechanical total line shape fitting model for quantitative NMR-based profiling of human serum metabolites. *Anal. Bioanal. Chem.* **406**, 3091–3102 (2014).
55. Fischer, K. *et al.* Biomarker Profiling by Nuclear Magnetic Resonance Spectroscopy for the Prediction of All-Cause Mortality: An Observational Study of 17,345 Persons. *PLoS Med.* **11**, e1001606 (2014).
56. Williams, K. J. & Tabas, I. The response-to-retention hypothesis of early atherogenesis. *Arteriosclerosis, Thrombosis, and Vascular Biology* vol. 15 551–562 (1995).
57. Steinbrecher, U. P., Parthasarathy, S., Leake, D. S., Witztum, J. L. & Steinberg, D. Modification of low density lipoprotein by endothelial cells involves lipid peroxidation and degradation of low density lipoprotein phospholipids. *Proc. Natl. Acad. Sci.* **81**, 3883–3887 (1984).
58. Ference, B. A. *et al.* Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. *Eur. Heart J.* **38**, 2459–2472 (2017).
59. The Emerging Risk Factors Collaboration*, T. E. R. F. Major Lipids, Apolipoproteins, and Risk of Vascular Disease. *JAMA* **302**, 1993 (2009).
60. Nordestgaard, B. G. A new start for triglycerides and remnant cholesterol-nonfasting. *Clinical Chemistry* vol. 63 1418–1419 (2017).
61. Varbo, A. & Nordestgaard, B. G. Remnant cholesterol and ischemic heart disease. *Current Opinion in Lipidology* vol. 25 266–273 (2014).
62. Nordestgaard, B. G. Triglyceride-Rich Lipoproteins and Atherosclerotic Cardiovascular Disease: New Insights From Epidemiology, Genetics, and Biology. *Circ. Res.* **118**, 547–63 (2016).
63. Jørgensen, A. B. *et al.* Genetically elevated non-fasting triglycerides and calculated remnant cholesterol as causal risk factors for myocardial infarction. *Eur. Heart J.* **34**, 1826–1833 (2013).
64. Hurt-Camejo, E. & Camejo, G. ApoB-100 Lipoprotein Complex Formation with Intima Proteoglycans as a Cause of Atherosclerosis and Its Possible Ex Vivo Evaluation as a Disease Biomarker. *Journal of Cardiovascular Development and Disease* vol. 5 (2018).
65. Ference, B. A. *et al.* Association of genetic variants related to CETP inhibitors and statins with lipoprotein levels and cardiovascular risk. *JAMA - J. Am. Med. Assoc.* **318**, 947–956 (2017).

66. Ala-Korpela, M. The culprit is the carrier, not the loads: cholesterol, triglycerides and apolipoprotein B in atherosclerosis and coronary heart disease. *Int. J. Epidemiol.* **48**, 1389–1392 (2019).
67. Richardson, T. *et al.* Apolipoprotein B underlies the causal relationship of circulating blood lipids with coronary heart disease. *Apolipoprotein B underlies causal Relatsh. Circ. blood lipids with Coron. Hear. Dis.* 19004895 (2019) doi:10.1101/19004895.
68. Glomset, J. A., Janssen, E. T., Kennedy, R. & Dobbins, J. Role of plasma lecithin: cholesterol acyltransferase in the metabolism of high density lipoproteins. *J. Lipid Res.* **7**, 639–648 (1966).
69. Badimon, J. J., Badimon, L. & Fuster, V. Regression of atherosclerotic lesions by high density lipoprotein plasma fraction in the cholesterol-fed rabbit. *J. Clin. Invest.* **85**, 1234–1241 (1990).
70. Lincoff, A. M. *et al.* Evacetrapib and cardiovascular outcomes in high-risk vascular disease. *N. Engl. J. Med.* **376**, 1933–1942 (2017).
71. Keech, A. *et al.* Effects of long-term fenofibrate therapy on cardiovascular events in 9795 people with type 2 diabetes mellitus (the FIELD study): randomised controlled trial. *Lancet* **366**, 1849–1861 (2005).
72. Schwartz, G. G. *et al.* Effects of dalcetrapib in patients with a recent acute coronary syndrome. *N. Engl. J. Med.* **367**, 2089–2099 (2012).
73. Rader, D. J. & Hovingh, G. K. HDL and cardiovascular disease. *The Lancet* vol. 384 618–625 (2014).
74. Group, T. H. C. Effects of Anacetrapib in Patients with Atherosclerotic Vascular Disease. *N. Engl. J. Med.* **377**, 1217–1227 (2017).
75. Khera, A. V. & Kathiresan, S. Genetics of coronary artery disease: discovery, biology and clinical translation. *Nat. Rev. Genet.* **18**, 331–344 (2017).
76. Gordillo-Marañón, M. *et al.* Validation of lipid-related therapeutic targets for coronary heart disease prevention using human genetics. *bioRxiv* 2020.11.11.377747 (2020) doi:10.1101/2020.11.11.377747.
77. Talayero, B. G. & Sacks, F. M. The role of triglycerides in atherosclerosis. *Curr. Cardiol. Rep.* **13**, 544–52 (2011).
78. Bansal, S. *et al.* Fasting Compared With Nonfasting Triglycerides and Risk of Cardiovascular Events in Women. *JAMA* **298**, 309 (2007).
79. Kastelein, J. J. P. & Stroes, E. S. G. FISHing for the Miracle of Eicosapentaenoic Acid. *N. Engl. J. Med.* **380**, 89–90 (2019).
80. Bhatt, D. L. *et al.* Cardiovascular Risk Reduction with Icosapent Ethyl for Hypertriglyceridemia. *N. Engl. J. Med.* NEJMoa1812792 (2018) doi:10.1056/NEJMoa1812792.
81. Singh, A. K. & Singh, R. Triglyceride and cardiovascular risk: A critical appraisal. *Indian J. Endocrinol. Metab.* **20**, 418–28 (2016).
82. Nordestgaard, B., Abildgaard, S., Circulation, H. W.- & 1997, undefined. Heterozygous lipoprotein lipase deficiency: frequency in the general population, effect on plasma lipid levels, and risk of ischemic heart disease. *Am Hear. Assoc*.

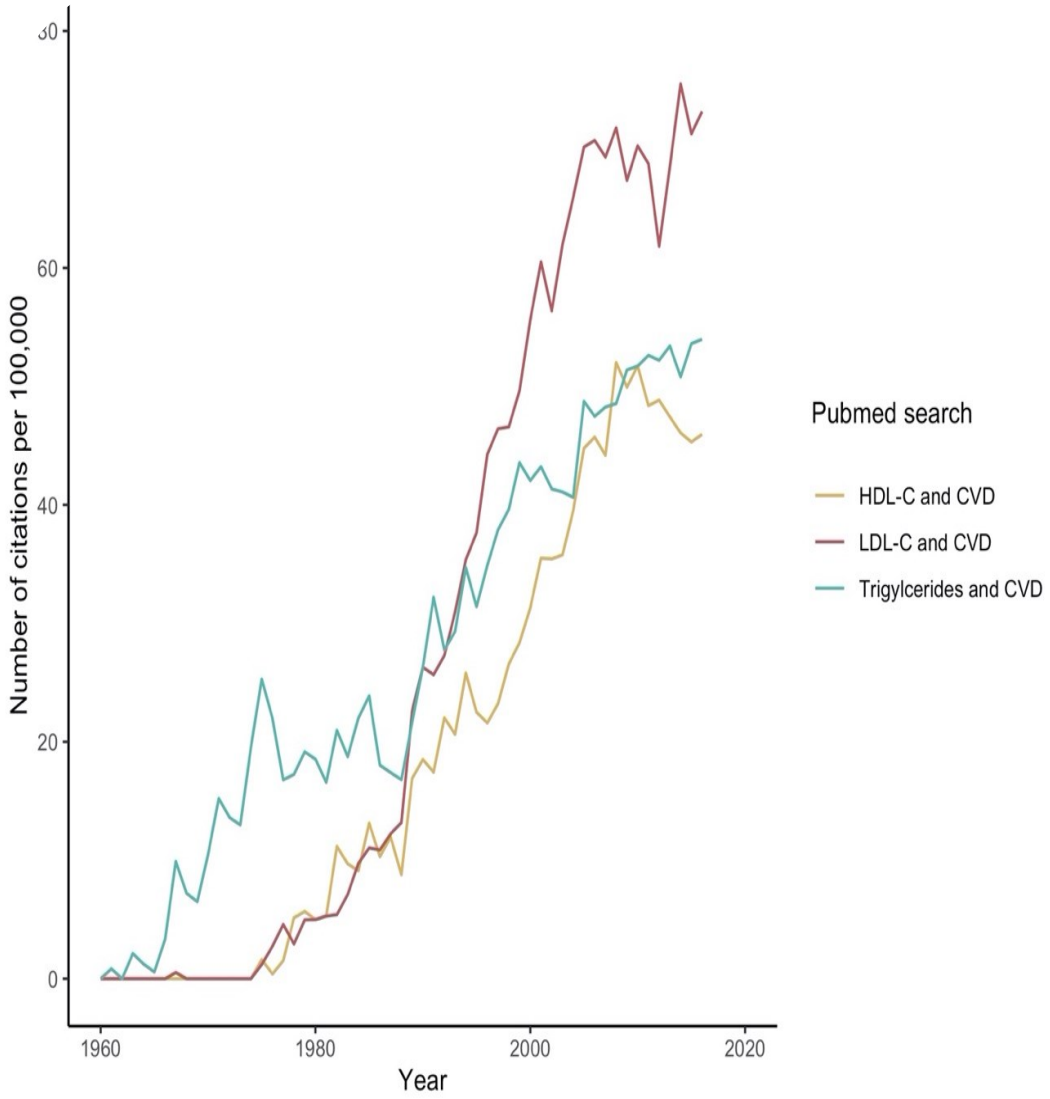
83. Rosenson, R., Davidson, M., … B. H.-J. of the & 2014, undefined. Genetics and causality of triglyceride-rich lipoproteins in atherosclerotic cardiovascular disease. *onlinejacc.org*.

2 Review of the current literature on associations of triglycerides with cardiovascular disease

The role triglyceride (TG) play in cardiovascular disease (CVD) risk is incompletely understood, such that clinical focus on elevated TG concentrations has fluctuated over the years. This is primarily due to changes in the evidence base. The absence of clinical trials that demonstrate a treatment benefit by lowering TG, combined with successful statin trials, has led to the majority of clinical focus being placed on lowering low-density lipoprotein cholesterol (LDL-C) for primary and secondary prevention of CVD. A residual CVD risk remains despite the current clinical attention on LDL-C, significant LDL-C lowering with statin therapies, and more recently, LDL-C lowering using PCSK-9 inhibitors with monoclonal antibodies for additional reduction to reach target LDL-C levels. Consequently, research interest in TG in relation to atherosclerosis and CVD has started to gain traction again in recent years (see figure 2.1).

A Pubmed search was conducted to year end 2017, when this PhD commenced. The following sections describes the significant epidemiological, genetic, and randomised control trial evidence relating to TG and CVD.

Figure 2.1 Number of PubMed citates for lipids and cardiovascular disease from 1960 to 20173



High-density lipoprotein cholesterol; HDL-C, low-density lipoprotein cholesterol; LDL-C, cardiovascular disease; CVD
Figure based on Pubmed search separate terms for high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, and triglyceride associations with cardiovascular disease from years 1960 to 2017.

2.1 Epidemiological data

There are strong positive epidemiological associations between TG and CVD[1–3]. However, it is unclear to what extent these associations are independent of CVD risk factors including HDL-C, the latter with which TG is inversely correlated. This has been demonstrated in a large scale meta-analysis of 68 long-term prospective studies conducted by The Emerging Risk Factors Collaboration (ERFC)[3]. The ERFC reports the hazard rate (HR) for CHD with TG was 1.37 after adjustment for nonlipid risk factors, but was reduced to 0.99 (95% CI, 0.94-1.05) after further adjustment for HDL-C and non–HDL-C. The attenuation of association between TG and CVD when accounting for HDL-C may imply the associations is confounded or, may represent an overadjustment for processes in the causal pathway. An added complexity is the right skewed distribution of TG in the population. To address this the ERFC authors log transformed TG measures to enable comparability to other lipid fractions, but forcing measures to a symmetrical distribution may also introduce variability, limit inferences or generalisability, and generate inaccurate estimates[4]. Recent findings have sparked further debate on the role of TG and CVD[5]. In a study with similar statistical power to the ERFC, the Copenhagen General Population Study[6] demonstrate increased risks for CHD and all-cause mortality for extremely high concentrations of TG, as well as a 5 fold increase for MI and 3.2 fold increase for stroke for plasma concentration of TG of 6.6 mmol/L versus 0.8 mmol/L[7].

Understanding the causal role of TG in CVD is pivotal to find effective approaches for disease prevention and treatment. Observational studies provide estimates of the likely role of TG concentrations in disease development and progression, but such studies are prone to bias. This has a direct impact for drug

development and randomised controlled trials. Novel drug development tracked to non-causal biomarkers can lead to expensive testing and subsequent failure of these drugs in phase III randomised controlled trials (RCTs), as seen for the drug trials discussed above[8]. Mendelian randomisation (MR) studies are used to ascertain causality of biomarkers, facilitated by the availability of large-scale genetic data. The principles, assumptions and limitations of MR studies are discussed next.

2.2 Genetics and Mendelian Randomisation

Mendelian randomisation analysis is facilitated by genome-wide association studies (GWAS). Genotyping platforms analysing millions of genetic variants termed single-nucleotide polymorphisms (SNPs), from large-scale global genetics consortia have identified genetic associations with phenotypic traits[9]. The fundamental principle of MR analysis and from which the name 'Mendelian randomisation' is derived, is in reference to Mendel's Second Law on the independent random assortment of alleles during meiosis, where DNA is transferred from parent to offspring at the time of gamete formation[10]. Inheritance of a particular variant or group of variants in an individual's DNA is inherited independent of other characteristics. Therefore, when grouping individuals based on genotype associating with TG concentrations, all other confounding characteristics should be similar, other than one group has genetically higher TG concentrations and the other group has genetically lower TG concentrations, analogous to randomisation in RCTs[9]. The valuable advantages of MR studies over observational studies are that genetic variants used as instrumental variables in MR studies are not susceptible to reverse causality, are not subjective to confounding due to Mendel's second law, and are measured with precision, reducing regression dilution bias due to measurement error.

The application of publicly available genetic data used in MR studies to evaluate relationship between risk factor-disease outcomes has made it a powerful tool in determining causal inference. The three key assumptions that form the definition of an instrumental variable for a valid MR study are[10,11]:

1. the genetic variant associates with the biomarker (the relevance assumption),
2. the genetic variant does not share common causes with the outcome (the independence assumption)
3. the genetic variants affect the outcome only through their effect on the risk factor (exclusion restriction)

There are two approaches to selecting instruments for MR analysis and each seek to answer different questions. Selecting a genetic instrument for the exposure, typically in the gene of interest, referred to as '*cis*-effect'. This approach is common when the exposure of interest is a specific drug target such as a protein, and is used in drug target validation studies to address whether modification of the encoded protein is will result in a reduction of disease outcome[12]. In the context of this thesis, cis-MR was not appropriate to ascertain the causal relevance of TG in CHD. Therefore, the second approach was used, which selects multiple genetic variants from across the genome, termed 'genome-wide MR'. Large sample size GWAS have been performed for complex traits such as lipids, identifying hundreds of independent variants reaching the established genome wide significance level. Independence of these traits is ensured in post analysis linkage disequilibrium (LD) clumping and pruning. In MR analysis, SNPs selected from across the genome robustly associated with the biomarker are used as an instrumental variable to test

whether the effects of the variants on the exposure result in proportional effects on the disease outcome.

There are three main potential limitations affecting the reliability of causal estimates obtained from genome wide MR analysis. These are weak instrument bias, statistical power and pleiotropy[13,14]. Weak instrument strength is determined by the magnitude and precision of associations of the genetic instrument with the risk factor. A higher value of the F-statistic (>10) indicates a strong instrument. Vertical pleiotropy does not invalidate the instrumental variable assumptions and does not result in bias. This is because vertically pleiotropic genetic variants affect the outcome through a pathway affected by the risk factor of interest[14,15]. A fundamental assumption of MR is the 'no horizontal pleiotropy' (exclusion restriction) assumption, which requires the genetic variants act on the disease outcome exclusively through the exposure of interest. Horizontal pleiotropy occurs when the genetic variant or variants affect the outcome outside the pathway of the exposure of interest[14,16]. A violation of this assumption can lead to biased causal estimates and potential false-positive causal relationships. *Cis* analyses may be less prone to horizontal pleiotropy as instruments are selected proxy to the causal gene of interest. The same is not true for genome wide MR analysis, whereby selecting variants from across the genome increases the potential for horizontal pleiotropy[17,18]. Moreover, emerging evidence suggests many traits are genetically correlated with each other and individual variants identified from larger sample sizes of GWAS are associated with multiple traits, both of which pose limitations on the validity of MR studies[19,20]. Recent method development has focused on attempting to identify the causal effect accounting for horizontally pleiotropic effects. These are discussed further.

In genome-wide MR analysis, a fixed effects inverse variance weighted (IVW) meta-analysis method is used, where the contribution of each variant to the overall estimate is the inverse of the variable of its effect on the outcome, with the intercept constrained through zero. MR-Egger regression relaxes the exclusion restriction assumption and allows for a non-zero intercept whereby the intercept term represents an estimate of the pleiotropic effect[14,21]. The IVW and MR-Egger methods both depend heavily on the InSIDE assumption (Instrument Strength Independent of Direct Effect), with the latter relaxing the assumption that the average pleiotropic effect is zero. InsIDE violation is likely when a large proportion of the horizontal pleiotropy occurs through a confounder of the exposure-outcome relationship[18]. The IVW and MR-Egger method model a single exposure variable effect on the outcome, using 'classical' univariate statistical techniques. Multivariable MR (MVMR) allows the for the adjustment of confounding variables affecting the exposure-disease pathway[21]. MVMR estimates the effect of the exposure on the outcome, conditioning the SNP-exposure effects on their corresponding effects on other known exposure traits[22,23]. An example of this is the genetic overlap between TG, LDL-C and HDL-C in estimating the influence of LDL-C on CHD. If the SNP-LDL-C effects are proportional to SNP-CHD effects, even after adjusting for the SNP-TG and SNP-HDL effects, this would support the conclusion that LDL-C has causal influence on CHD[21]. The MR-Egger method can be applied to MVMR to account for potential horizontal pleiotropy, using the same principles discussed above. The Rucker framework has been adapted for the MR context to select between the different models[14,24]. This begins by estimating the effects using IVW analysis and calculating the Cochran's Q statistic for heterogeneity, which indicates whether SNP-outcome associations are inconsistent and could cause biased effects. Next the heterogeneity

is re-estimated using MR-Egger analysis with a non-zero intercept (using Rucker's Q' statistic). If Q-Q' is large, this indicates the presence of horizontal pleiotropy, suggesting it is more appropriate to use the MR-Egger framework[14].

The role of MR studies to infer causality have been confirmed in proof of concept studies of genetic determinants of LDL-C concentrations associated with higher CHD evaluated against evidence from RCT of lipid lowering medication[25,26]. Similarly, studies instrumenting multiple TG associated SNPs from across the genome have reported similar findings for genetically elevated TG and an increased CVD [27,28]. Holmes *et al* used weighted allele scores based on multiple SNPs and found allele scores for TG were associated with CHD events in both unrestricted (67 SNPs, OR: 1.62; 95% CI: 1.24 to 2.11), and restricted (27 SNPs, OR: 1.61; 95% CI: 1.00 to 2.59) analysis[28]. In a cis-MR approach selecting instrument(s) associated with proteins that affect TG, a study by The Emerging Risk Factors Collaboration[29] compared MR estimates instrumenting a SNP in the *APOA5* gene and risk of CHD, with estimates obtained from prospective studies. For each inherited allele, individuals had a dose-dependent 0.25mmol/L higher mean TG concentration and an 18% increased risk for CHD (Odds Ratio: 1. 18; 95% CI 1.11 to 1. 26). The findings from the MR study were concordant with the hazard ratio of an equivalent TG increase obtained from prospective studies (HR: 1.10; 95% CI 1.08 to 1.12). Correspondingly, TG lowering mutations in the *APOC3* gene have demonstrated a reduction in CHD risk by 39- 41%[30,31]. Genetic inactivation of *APOC3, ANGPLT3* and *ANGPTL4* genes that encode for inhibitors of LPL, the enzyme involved in TG metabolism, are associated with lower TG levels and CVD risk. Based on these

findings, new therapies to reduce TG concentrations via lipase antagonism are currently in development, discussed below.

2.3 Triglyceride lowering treatment trials

The main known TG-lowering therapies are fibrates, niacin, CETP-inhibitors, or omega-3 fatty acids. These are not routinely prescribed for TG lowering for primary or secondary prevention of CVD due to the lack of compelling findings from clinical trials.

Fibrates decrease TG concentrations by approximately 36%[32]. Despite TG lowering, cardiovascular outcome studies of fibrate therapy have produced variable results and failed to show reduction in CVD risk when administered with statins in combination therapy[33,34]. As mentioned above, observational evidence suggest risk of CHD is largest at TG concentrations of 6.6 mmol/L[6], and therefore results from the fibrate trials cannot show if a reduction of TG concentration provides CVD benefit as most fibrate trials exclude participants with TG concentrations greater than 4.5 mmol/L[35].

Similarly, niacin reduces TG concentrations by up to 20% and CVD events by 37%. In post hoc trial analysis in a subset of patients with high TG greater than 2.3mmol/L and HDL-C less than 0.8mmol/L, but did not show added treatment benefit when adding niacin to statin therapy[32,36]. However, it is noted that authors report post-hoc subgroup analysis and such findings should be interpreted with caution as significant associations may due to chance[37]. Due to the lack of treatment benefit coupled with the high incidence of adverse effects in patients, niacin use is

limited. In addition to HDL-C raising, the REVEAL (The Randomized EValuation of the Effects of Anacetrapib Through Lipid-modification) trial[38] also report a reduction in TG in the treatment arm of the trial and fewer first major coronary events at the end of the follow-up period compared to patients in the placebo arm [RR: 0.91 (95% CI 0.85 to 0.97)]. This was the first trial to report such findings.

Omega-3 fatty acid agents reduce elevated TG concentrations by up to 33% and are mostly used in individuals with extreme hypertriglyceridemia to modulate inflammatory and immunological responses in conditions such as acute pancreatitis[39]. Although efficacious in response to acute pancreatitis, the same TG-lowering benefit has not been observed for cardiovascular outcomes, with the exception of one recent trial[40–43]. The largest Cochrane review to date of trials of omega-3 fatty acids for the primary and secondary prevention of CVD concluded that with the possible exception of alpha-linoleic acid, a type of omega-3 fatty acid found in plants, fatty acid agents 'have little or no effect on mortality or cardiovascular health'[44]. Against this background, the recent publication of the multi-site, multi-country Reduction of Cardiovascular Events with Icosapent Ethyl–Intervention Trial (REDUCE-IT)[45] came as a surprise. In the REDUCE-IT trial, 8179 high-risk CVD patients receiving statin therapy, were randomised to receive 4mg daily dose of icosapent ethyl (EPA) or placebo containing mineral oil. At baseline, LDL-C were well controlled, and TG were slightly elevated (median 1.94 and 2.44 mmol/L, respectively). The addition of EPA to statin therapy resulted in a 21.6% reduction in TG at the end of 5 year follow up and 25% lower risk in the main composite end point of CVD events among those in the treatment arm compared to placebo (hazard ratio, 0.75; 95% CI, 0.68 to 0.83; P<0.001). The results from the

REDUCE-IT trial contradict results from other recent randomised trials directed at reducing risk of cardiovascular events beyond LDL-C lowering. Ongoing REDUCE-IT critique deliberates the higher than expected cardiovascular benefit of EPA on the basis of changes in TG levels. Kastelein and colleagues[46] comment the observed median reduction of 0.36 mml/L in non-HDL-C from baseline would translate to a lower risk of CVD events of 6-8%, not the 25% reduced risk observed in the REDUCE-IT. Kastelein *et al* further argue results were similar irrespective of whether normal TG concentration was attained, which suggests the findings may argue against the theory that TG lowering reduces cardiovascular events. REDUCE-IT authors attribute significant findings in part to the formulation (highly purified EPA ethyl) and the daily 4mg dose used that was different to those in previous outcome trials. Bhatt et al[45] further comment that REDUCE-IT results were similar to those of the Japan EPA Lipid Intervention Study[47] (JELIS). Authors for JELIS report 19% lower cardiovascular events with stain therapy plus 1.8g of EPA daily than with statin therapy alone. It is possible the positive findings in REDUCE-IT may be explained by pleiotropic non-lipid mechanisms by EPA, much like effects exerted by statins, including anti-inflammatory, stabilisation of coronary plaques or improvement in endothelial dysfunction, which may result in the significant reduction in clinical endpoints[48]. Given the uncertainty of the precise mechanism of action of EPA, the debate over whether EPA or high doses of EPA are beneficial for CVD events is ongoing. The anticipated results of the STRENGH (Statin Residual Risk Reduction With Epanova in High Cardiovascular Risk Patients with Hypertriglyceridemia) trial of Epanova, is predicted to provide further clarity [49].

Angiopoietin-related proteins are important regulators of lipoprotein metabolism. ANGPTL3 is an endogenous inhibitor of lipoprotein lipase (LPL), the main enzyme involved in hydrolysis of TRL[50]. In light of favourable consequences of ANGPTL3 deficiency, an ANGPTL3 antibody (evinacumab) and anti-sense oligonucleotide (ASO) have entered clinical trials with encouraging results. Following administration of the ASO, *ANGPTL3* mRNA expression and plasma ANGPTL3 protein levels were significantly decreased in the mouse models[50,51]. The ASO decreased TG concentrations and slowed progression of atherosclerosis in LDL-receptor knockout mice[50]. In 44 human participants randomised to receive subcutaneous ASO injections or placebo, lower levels of ANGPTL3 protein, TG, LDL-C, non-HDL-C, and total cholesterol were observed in the ASO group compared with placebo group after 6 weeks of treatment[52]. In a Phase I study, evinacumab was tested in 83 healthy human volunteers with mildly raised TG or LDL-C concentrations[53]. The participants were randomised to receive evinacumab or placebo, and TG concentrations were reduced by 76%, LDL-C by 23.2% and HDL-C by 18.4% in the treatment vs placebo arm. These findings support *ANPTL3* pathway as important in lipid regulation and CVD.

2.4 Chapter summary

In the post-statin era there remains a residual CVD risk, renewing the research and clinical interest in TG. While the observational evidence between the association of elevated TG and cardiovascular events is supported, it is not unconfounded, with considerable attenuation of effects when adjusting for CVD risk factors. Moreover, the observational effects do not provide estimates of causality. In contrast, genetic evidence from MR studies suggests a causal role of TG. Though MR results from

such studies are compelling, instrumented genetic variants associated with elevated TG concentrations also likely associate with other lipoproteins, namely HDL-C and LDL-C, violating the 'no horizontal' pleiotropy' assumption of the MR principle. Finally, with the exception of the REDUCE-IT trial, experimental TG lowering has not yet shown to improve CVD outcomes.

The described inconsistent associations across the different types of studies make it plausible that an effective agent has not yet been found or, that the hypothesis is flawed, resulting in causal uncertainty regarding TG and its role in atherosclerotic disease and CVD progression. Given the complexity of the metabolic pathway, as discussed in Chapter 1, and the incomplete understanding of the role of TG in CVD, it remains possible that certain TG containing lipoproteins are pro-atherogenic, while others are not, such that total TG measurement may be insufficiently precise to disentangle the causal relationships of individual lipoprotein fractions with CVD. Using a combined genetic and metabolomics approach and applying the principles and tools for observational epidemiological data analysis and MR, this PhD aims to identify the causal relevance of TG containing lipoprotein subfractions in cardiovascular disease.

The application of NMR metabolomics to measure and quantify lipid lipoproteins offers the opportunity to interrogate and evaluate the causal associations of TG measured in 14 lipoprotein subfractions in the context of the current evidence, some of which has been described in this chapter. The following chapters focus on addressing the aims of this thesis outlined above, and the final chapter summarises the understanding of the role of TG in CVD and directions for future development.

2.5 References

1. Bansal, S. *et al.* Fasting Compared With Nonfasting Triglycerides and Risk of Cardiovascular Events in Women. *JAMA* **298**, 309 (2007).
2. Langsted, A., Freiberg, J. J. & Nordestgaard, B. G. Fasting and nonfasting lipid levels influence of normal food intake on lipids, lipoproteins, apolipoproteins, and cardiovascular risk prediction. *Circulation* **118**, 2047–2056 (2008).
3. The Emerging Risk Factors Collaboration*, T. E. R. F. Major Lipids, Apolipoproteins, and Risk of Vascular Disease. *JAMA* **302**, 1993 (2009).
4. Feng, C. *et al.* Log-transformation and its implications for data analysis. *Shanghai Arch. psychiatry* **26**, 105–9 (2014).
5. Wiesner, P. & Watson, K. E. Triglycerides: A reappraisal. *Trends Cardiovasc. Med.* **27**, 428–432 (2017).
6. Freiberg, J., Tybjærg-Hansen, A., Jama, J. J.- & 2008, undefined. Nonfasting triglycerides and risk of ischemic stroke in the general population. *jamanetwork.com*.
7. Nordestgaard, B. G. Triglyceride-Rich Lipoproteins and Atherosclerotic Cardiovascular Disease: New Insights From Epidemiology, Genetics, and Biology. *Circ. Res.* **118**, 547–63 (2016).
8. Hingorani, A. D. *et al.* Improving the odds of drug development success through human genomics: modelling study. *Sci. Rep.* **9**, 18911 (2019).
9. Bennett, D. A. & Holmes, M. V. Mendelian randomisation in cardiovascular research: an introduction for clinicians. *Heart* **103**, 1400–1407 (2017).
10. Hingorani, A. & Humphries, S. Nature's randomised trials. *Lancet (London, England)* **366**, 1906–8 (2005).
11. Davies, N. M., Holmes, M. V & Smith, G. D. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* **362**, 601 (2018).
12. Andrikoula, M. & McDowell, I. F. W. The contribution of ApoB and ApoA1 measurements to cardiovascular risk assessment. *Diabetes, Obes. Metab.* **10**, 271–278 (2008).
13. Carter, A. R. *et al.* Mendelian randomisation for mediation analysis: current methods and challenges for implementation. doi:10.1101/835819.
14. Bowden, J. *et al.* A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Stat. Med.* **36**, 1783–1802 (2017).
15. Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M. & Smoller, J. W. Pleiotropy in complex traits: challenges and strategies. *Nat. Rev. Genet.* **14**, 483–495 (2013).
16. Schmidt, A. F. *et al.* Genetic drug target validation using Mendelian randomisation. *Nat. Commun.* **11**, 3255 (2020).
17. Schmidt, A. F. & Dudbridge, F. Mendelian randomization with Egger

pleiotropy correction and weakly informative Bayesian priors. *Int. J. Epidemiol.* **47**, 1217–1228 (2018).
18. Burgess, S. & Thompson, S. G. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur. J. Epidemiol.* **32**, 377–389 (2017).
19. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236 (2015).
20. Webb, T. R. *et al.* Systematic evaluation of pleiotropy identifies 6 further loci associated with coronary artery disease. *J. Am. Coll. Cardiol.* **69**, 823–836 (2017).
21. Burgess, S. & Thompson, S. G. Multivariable Mendelian Randomization: The Use of Pleiotropic Genetic Variants to Estimate Causal Effects. *Am. J. Epidemiol.* **181**, 251–260 (2015).
22. Rees, J. M. B., Wood, A. M. & Burgess, S. Extending the MR-Egger method for multivariable Mendelian randomization to correct for both measured and unmeasured pleiotropy. *Stat. Med.* **36**, 4705–4718 (2017).
23. Allara, E. *et al.* Genetic Determinants of Lipids and Cardiovascular Disease Outcomes: A Wide-Angled Mendelian Randomization Investigation. *Circ. Genomic Precis. Med.* **12**, 543–551 (2019).
24. Hemani, G., Bowden, J. & Davey Smith, G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Human Molecular Genetics* vol. 27 R195–R208 (2018).
25. Linsel-Nitschke, P. *et al.* Lifelong reduction of LDL-cholesterol related to a common variant in the LDL-receptor gene decreases the risk of coronary artery disease—a Mendelian randomisation study. *PLoS One* **3**, (2008).
26. Swerdlow, D. I., Hingorani, A. D. & Humphries, S. E. Genetic Risk Factors and Mendelian Randomization in Cardiovascular Disease. *Current Cardiology Reports* vol. 17 1–11 (2015).
27. Thomsen, M., Varbo, A., Tybjærg-Hansen, A. & Nordestgaard, B. G. Low nonfasting triglycerides and reduced all-cause mortality: a mendelian randomization study. *Clin. Chem.* **60**, 737–46 (2014).
28. Holmes, M. V. *et al.* Mendelian randomization of blood lipids for coronary heart disease. *Eur. Heart J.* **36**, 539–550 (2015).
29. Triglyceride-mediated pathways and coronary disease: collaborative analysis of 101 studies. *Lancet* **375**, 1634–1639 (2010).
30. Jørgensen, A. B., Frikke-Schmidt, R., Nordestgaard, B. G. & Tybjærg-Hansen, A. Loss-of-Function Mutations in *APOC3* and Risk of Ischemic Vascular Disease. *N. Engl. J. Med.* **371**, 32–41 (2014).
31. Loss-of-Function Mutations in *APOC3,* Triglycerides, and Coronary Disease. *N. Engl. J. Med.* **371**, 22–31 (2014).
32. Toth, P. P. Triglyceride-rich lipoproteins as a causal factor for cardiovascular disease. *Vasc. Health Risk Manag.* **12**, 171–83 (2016).
33. Keech, A. *et al.* Effects of long-term fenofibrate therapy on cardiovascular events in 9795 people with type 2 diabetes mellitus (the FIELD study):

randomised controlled trial. *Lancet* **366**, 1849–1861 (2005).
34. Effects of Combination Lipid Therapy in Type 2 Diabetes Mellitus. *N. Engl. J. Med.* **362**, 1563–1574 (2010).
35. Nordestgaard, B. G. & Varbo, A. Triglycerides and cardiovascular disease. *Lancet* **384**, 626–635 (2014).
36. Guyton, J. R. *et al.* Relationship of Lipoproteins to Cardiovascular Events. *J. Am. Coll. Cardiol.* **62**, 1580–1584 (2013).
37. Peto, R. Current misconception 3: that subgroup-specific trial mortality results often provide a good basis for individualising patient care. *Br. J. Cancer* **104**, 1057 (2011).
38. Group, T. H. C. Effects of Anacetrapib in Patients with Atherosclerotic Vascular Disease. *N. Engl. J. Med.* **377**, 1217–1227 (2017).
39. Lei, Q. C. *et al.* The role of omega-3 fatty acids in acute pancreatitis: a meta-analysis of randomized controlled trials. *Nutrients* **7**, 2261–73 (2015).
40. Kotwal, S., Jun, M., Sullivan, D., Perkovic, V. & Neal, B. Omega 3 Fatty Acids and Cardiovascular Outcomes. *Circ. Cardiovasc. Qual. Outcomes* **5**, 808–818 (2012).
41. Jun, M. *et al.* Effects of fibrates on cardiovascular outcomes: a systematic review and meta-analysis. *Lancet* **375**, 1875–1884 (2010).
42. Kromhout, D., Giltay, E. J., Geleijnse, J. M. & Alpha Omega Trial Group. n–3 Fatty Acids and Cardiovascular Events after Myocardial Infarction. *N. Engl. J. Med.* **363**, 2015–2026 (2010).
43. Rauch, B. *et al.* OMEGA, a Randomized, Placebo-Controlled Trial to Test the Effect of Highly Purified Omega-3 Fatty Acids on Top of Modern Guideline-Adjusted Therapy After Myocardial Infarction. *Circulation* **122**, 2152–2159 (2010).
44. Abdelhamid, A. S. *et al.* Omega-3 fatty acids for the primary and secondary prevention of cardiovascular disease. *Cochrane Database Syst. Rev.* (2018) doi:10.1002/14651858.CD003177.pub4.
45. Bhatt, D. L. *et al.* Cardiovascular Risk Reduction with Icosapent Ethyl for Hypertriglyceridemia. *N. Engl. J. Med.* NEJMoa1812792 (2018) doi:10.1056/NEJMoa1812792.
46. Kastelein, J. J. P. & Stroes, E. S. G. FISHing for the Miracle of Eicosapentaenoic Acid. *N. Engl. J. Med.* **380**, 89–90 (2019).
47. Yokoyama, M. *et al.* Effects of eicosapentaenoic acid on major coronary events in hypercholesterolaemic patients (JELIS): a randomised open-label, blinded endpoint analysis. *Lancet* **369**, 1090–1098 (2007).
48. Boden, W. E. *et al.* Profound reductions in first and total cardiovascular events with icosapent ethyl in the REDUCE-IT trial: why these results usher in a new era in dyslipidaemia therapeutics. *Eur. Heart J.* (2019) doi:10.1093/eurheartj/ehz778.
49. Outcomes Study to Assess STatin Residual Risk Reduction With EpaNova in HiGh CV Risk PatienTs With Hypertriglyceridemia - Full Text View - ClinicalTrials.gov. https://clinicaltrials.gov/ct2/show/NCT02104817.

50. Wang, X. & Musunuru, K. Angiopoietin-Like 3: From Discovery to Therapeutic Gene Editing. *JACC: Basic to Translational Science* vol. 4 755–762 (2019).
51. Stitziel, N. O. *et al.* ANGPTL3 Deficiency and Protection Against Coronary Artery Disease. *J. Am. Coll. Cardiol.* **69**, 2054–2063 (2017).
52. Graham, M. J. *et al.* Cardiovascular and metabolic effects of ANGPTL3 antisense oligonucleotides. *N Engl J Med* **377**, 222–232 (2017).
53. Dewey, F. E. *et al.* Genetic and pharmacologic inactivation of ANGPTL3 and cardiovascular disease. *N. Engl. J. Med.* **377**, 211–221 (2017).
54. Manninen, V. *et al.* Lipid alterations and decline in the incidence of coronary heart disease in the Helsinki Heart Study. *jamanetwork.com*.
55. Rubins, H. B. *et al.* Gemfibrozil for the Secondary Prevention of Coronary Heart Disease in Men with Low Levels of High-Density Lipoprotein Cholesterol. *N. Engl. J. Med.* **341**, 410–418 (1999).
56. Klempfner, R. *et al.* Elevated Triglyceride Level Is Independently Associated With Increased All-Cause Mortality in Patients With Established Coronary Heart Disease: Twenty-Two-Year Follow-Up of the Bezafibrate Infarction Prevention Study and Registry. *Circ. Cardiovasc. Qual. Outcomes* **9**, 100–8 (2016).
57. Scott, R., O'brien, R., Fulcher, G., … C. P.-D. & 2009, undefined. Effects of fenofibrate treatment on cardiovascular disease risk in 9,795 individuals with type 2 diabetes and various components of the metabolic syndrome. *Am Diabetes Assoc*.

3 Methods

This chapter provides and overview of the datasets, exposure and outcome phenotype measures used throughout this thesis. Detailed methods are discussed further in the succeeding results chapters.

3.1 Datasets

University College London-Edinburgh-Bristol (UCLEB) Consortium

This thesis uses data sourced from prospective observational studies from the UCLEB consortium[1]. These are the Whitehall-II Study (WHII), the British Regional Heart Study (BRHS), the Southall and Brent Revisited Study (SABRE), the MRC National Survey of Health and Development (NSHD), the Caerphilly Prospective Study (CAPS), and the British Women's Heart and Health Study (BWHHS). The strengths of UCLEB consortium include it being a stable long-term resource for large scale integrated metabolomics and genomic analyses. The integration of multiple layers of -omics data within the framework of cohort studies, large sample size and standardised lipid measurements contribute to a more comprehensive approach to address the aims of this thesis.

The studies are discussed below.

All studies with the exception of SABRE are almost all exclusively of European ancestry. The age of recruitment ranges from birth (NSHD) to 60-79 years (BWHHS), with most cohorts recruiting in mid-life. The current age of participants spans the 5th to 9th decades of life when the majority of cardiovascular disease manifest, making the consortium a valuable source for cases of incident disease. Each of the studies are prospective cohort design and have NMR metabolomics data quantified using the Nightingale platform (details of quantification methods discussed in Chapter 1) and DNA repository with published genetic analyses. The strength of cohort-based analyses is that genetic loci can be identified for very quantitative trait recorded in sufficiently large numbers. Each of the contributing studies has a defined inclusion criterion, procedures for the collection and recording

of demographic details, biological samples and clinical measures. The studies have a wide range of clinical and biological measures with overlap across studies to facilitate pooled analyses. Studies have used common measurement methods, with many blood markers measured in the same laboratory. All studies follow participants for disease and have an ongoing clinical assessments and biological sampling. Table 3.1 below shows the variables obtained from each contributing study. The individual cohorts are discussed further below. All ethical approval was collected locally by individual studies.

Table 3.1 Description of UCLEB variables used in this thesis

Variable	Unit of measure	Variable	Unit of measure
Triglyceride and cholesterol in 14 lipoprotein subfraction metabolites			
Extremely large VLDL	mmol/L	Large LDL	mmol/L
Very large VLDL	mmol/L	Medium LDL	mmol/L
Large VLDL	mmol/L	Small LDL	mmol/L
Medium VLDL	mmol/L	Very large HDL	mmol/L
Small VLDL	mmol/L	Large HDL	mmol/L
Very small VLDL	mmol/L	Medium HDL	mmol/L
IDL	mmol/L	Small HDL	mmol/L
NMR measured lipids			
Total cholesterol	mmol/L	Apolipoprotein AI	g/L
Total triglyceride	mmol/L	Apolipoprotein B	g/L
LDL-cholesterol	mmol/L		
HDL-cholesterol	mmol/L		
Clinical chemistry measured lipids			

Total cholesterol	mmol/L	LDL-cholesterol	mmol/L
Total triglycerides	mmol/L	HDL-cholesterol	mmol/L
Anthropometric measures			
Age	years	Height	Meters
Sex	Male/female	Weight	kg
Participant characteristics and lifestyle measures			
Smoking	Yes/no	SBP	mmHg
Alcohol	Yes/no	DBP	mmHg
Type 2 diabetes	Yes/no	HbAIc	mmol/L
Disease outcome			
CHD[b]	Yes/no		
Stroke[b]	Yes/no		

SBP: systolic blood pressure; DBP: diastolic blood pressure; LDL-C: low-density lipoprotein cholesterol; HDL-C: high-density lipoprotein cholesterol; TG: triglycerides; CHD: coronary heart disease.
[a] Calculated using Friedewald et al.(1972). [LDL-chol] = [Total chol] - [HDL-chol] - ([TG]/2.2) where all concentrations are given in mmol/L (note that if calculated using all concentrations in mg/dL then the equation is [LDL-chol] = [Total chol] - [HDL-chol] - ([TG]/5))
[b] CHD is defined as first event of myocardial infarction or revascularisation and stroke as first event of ischaemic or haemorrhagic stroke.

3.1.1 British Regional Heart Study (BRHS)

The BRHS recruited 7735 men in 1978 to 1980 aged 40-59 years from general practices across the UK[2]. At re-examination in 1998-2000, when the men were 60-79 years, a wide range of phenotypic measures such as lipids, blood pressure, inflammatory markers, and anthropometric and behavioural variables such as BMI, cigarette smoking and alcohol consumption were collected in 4252 participants and DNA was extracted for 3945. NMR metabolomics quantification of blood collected

in the fasting and non-fasting state at re-examination were analysed in Finland using the Nightingale platform[3]. Incident outcome variables for coronary heart disease and stroke were collected through self-report questionnaires and validated through medical records at 10 year follow up from re-examination in 2010. A positive answer to self-report question "Have you ever been told by a doctor that you have had a heart attack (coronary thrombosis or myocardial infarction)?" was validated using a history of typical features including chest pain, supported by ECG evidence and/or abnormal enzyme levels, (WHO criteria classify, two of the three). All new major CHD events reported by the practices are followed-up with an enquiry form to the GP or hospital consultant to obtain confirmatory evidence that case criteria have been met. Similarly, a positive answer to "Have you ever been told by a doctor that you have had a stroke?" and validated using record-review of an acute disturbance of cerebral function of vascular origin, lasting >24 hours. Case definition includes subarachnoid haemorrhage, cerebral haemorrhage or thrombosis.

3.1.2 Whitehall II Study (WHII)

The Whitehall II study recruited over 10,000 participants between 1985 and 1988 from 20 London based Civil service departments, of which 66% are men[4]. The study has 9 phases of follow-up (5 with clinical assessment and biological sampling) over 20 years. Data used in this thesis are obtained from phase 5 measured in 1997-1999 when NMR quantification was performed in 4762 participants using fasting and non-fasting blood samples using the Nightingale platform. Anthropometric, lifestyle and participant characteristic data are also collected from phase 5. Disease outcome data of incident CHD and stroke were collected at seven-year follow-up from phase 5 using self- report questionnaires as positive answer to "have you ever

been told by a doctor that you have had a heart attack, stroke or transient ischaemic attack"? Genetic samples were collected in 2004 from over 6,000 participants.

3.1.3 Medical Research Council National Survey of Health and Development (NSHD)

The NSHD is an ongoing birth cohort study consisting of all births in England, Scotland and Wales in one week in March 1946[5]. The original cohort comprised 2,547 women and 2,815 men who have been followed up over 20 times since their birth. NMR quantification was conducted using the Nightingale platform at follow up in 2006-2010 in 1790 participants aged 60-65 years using blood samples measured in the fasting and non-fasting state. Data permanenting to physical, lifestyle and anthropometric measures and outcome phenotypes for self-reported or doctor diagnosed prevalent CHD and stroke were also collected in 2006-2010.

3.1.4 Caerphilly prospective study (CaPS)

The CaPs study is based on men aged 45 to 59 years who lived in Caerphilly, South Wales and were recruited to the study between 1979 and 1983[6]. NMR metabolomics quantification was performed on fasting blood measures in 1500 participants at follow-up between 1989 and 1993, anthropometric and lifestyle variables were also collected at this time. DNA was extracted from blood samples collected in 1992-1994. Follow-up for disease outcomes is by self-report questionnaire for CHD, "Have you ever had a heart attack or coronary thrombosis?" and for stroke "Have you ever had a stroke?". Positive answers to self-report questions are linked to hospital episode discharge summaries for validation checks.

3.1.5 Southall and Brent Revisited study (SABRE)

The SABRE study is a tri-ethnic study including British men and women of European, South Asian and African Caribbean descent living in West and North London[7]. Participants were recruited in 1988-1991 and were re-examined at 20-year follow up in 2008-2011. NMR metabolite measures were quantified using both fasting and non-fasting blood samples measured at baseline (N = 3593) and at 20year follow-up (N = 1426). All study variables including anthropometric and lifestyle variables are obtained from baseline year. Incident disease outcome at 20 year follow up in 2008-2011 were ascertained using self-report questionnaires as positive answers to "Have you ever been told by a doctor that you have had a heart attack (coronary thrombosis or myocardial infarction)?" and validated using primary care record review adjudicated by 2 senior physicians, based on symptoms, cardiac enzymes, ECG findings and hospital discharge diagnosis. Similarly, stroke was defined as self-report as positive answer to question "Have you ever been told by a doctor that you have had a stroke or transient ischaemic attack with symptoms lasting at least 24 hours?" and validated similar to CHD events, with definite or probable diagnosis of stroke made according to predetermined criteria based on symptoms, duration of symptoms and MR or CT imagining. For the purpose of this thesis the SABRE cohorts are subdivided by ethnicity, defined as SABRE 1 (European), SABRE 2 (South Asian) and SABRE 3 (African-Caribbean).

3.1.6 Caerphilly prospective study (CaPS)

The BWHHS was established in 1999 as a parallel to the BRHS, using the same sampling frame and clinical protocols for follow-up[8]. From 1999-2001, 4286 women aged 60-69 were selected from 23 general practices across the UK. DNA and NMR

metabolomic quantification was performed on fasting blood samples from baseline year 1999-2001. Data on anthropometric and lifestyle variables were collected at the same time point. At 10-year follow-up in 2010, disease outcome data were ascertained using self-report questionnaires for CHD as a positive answer to question "Have you ever been told by a doctor that you have had a heart attack (coronary thrombosis or myocardial infarction)?". Self-report answers were validated by any one of the following; ECG evidence, raised enzyme levels, raised cardiac enzyme levels, raised troponin levels, hospital letter confirming diagnosis, recoded from unstable angina following review by study investigators. Similarly, stroke cases were ascertained as positive answer to questions "Have you ever been told by a doctor that you have had stroke?", and validated as any one of; ischaemic or haemorrhagic stroke at scan and symptoms > 24 hours, hospital letter confirming diagnosis and symptoms > 24 hours, Final diagnosis is stroke without scan or letter and symptoms >24 hours, recorded following record review by study investigators.

3.2 Genetic association data

3.2.1 Exposure data

Genetic association estimates for TG and cholesterol in 14 lipoprotein subfractions were obtained from an existing NMR metabolite GWAS meta-analysis conducted in my host laboratory in 25,000 samples from 7 studies contributing to the UCLEB consortium. Blood metabolites were quantified using the Nightingale[9] NMR platform as described above in cohort studies. Genotyping was completed using the Illumina Cardio-Metabochip array[10] and imputed based on 1000G for SNPs with MAF<0.001. Metabolites were tested against genotypes adjusting for age and gender and 10 principal components. There was no adjustment for population stratification

as UCLEB studies were previously shown to be homoegenous[1]. To correct for multiple testing, genome- and metabolome-wide statistical significance was set to $p<2.3e-9$.

3.2.2 Outcome data

Ascertainment of SNP-CHD genetic association via publicly available summary data from the CARDIoGRAMplusC4D in 63,746 cases and 130,681 controls, genotyped using either Metabochip or GWAS data imputed using HapMap[11]. Participants were of either European (95%) or South Asian ancestry. Cases were defined as documented history of acute coronary syndrome, coronary artery bypass graft, percutaneous coronary revascularisation, coronary artery stenosis greater than 50% in at least one coronary vessel, or angina pectoris. Corrections were made for age, sex and population stratification.

3.2.3 Statistical analysis methods

Specific methods of analysis are described in detail in the succeeding results chapters in the respective methods sections. As an overview, chapter 4 evaluates the influence of age, smoking, body mass index (BMI) on the triglyceride (TG) and cholesterol distributions using generalised linear model (GAM) curves. This approach was selected as it is considered more flexible and adapts the fitted curve to the data to identify hidden patterns as compared to other models such as a polynomial regression, which restricts the form of the curve. The influence of fasting status on TG distribution was assessed using the Kolmogorov–Smirnov (KS) test due to the capability of the KS test to detect variances in the population over other tests such as the t-test. In chapter 5, logistic regression models were used to assess

the relationship between TG in 14 lipoprotein subfractions and coronary heart disease and stroke. Had time-to CHD or stroke event data been available in UCLEB, Cox regression analyses may have been used. This is a potential limitation of UCLEB data. In the absence of such data, it is appropriate to use logistic regression given the binary outcome of CHD and stroke. Mendelian randomisation (MR) is used in the genetic analyses in chapter 6 to evaluate the predominant causal lipoprotein lipid in CHD. The assumptions underlying MR are discussed in chapter 2. Briefly, genetic associations with TG and cholesterol in 14 lipoprotein subfractions were obtained from a meta-analyses of Kettunen[12] and a de novo genome-wide association study (GWAS) of UCLEB measurements. The selected instruments were harmonised to genetic associations with CHD in the outcome database CARDIoGRAMplusC4D[13] . The selected instruments were used to evaluate the effect of TG and cholesterol content of each lipoprotein subfractions on CHD using the inverse variance weighted estimator and multivariable MR, correcting for horizontal pleiotropy. Please see methods section chapter 6 for further details.

3.3 References

1. Shah, T. *et al.* Population Genomics of Cardiometabolic Traits: Design of the University College London-London School of Hygiene and Tropical Medicine-Edinburgh-Bristol (UCLEB) Consortium. *PLoS One* **8**, e71345 (2013).
2. Shaper, A. G. *et al.* British Regional Heart Study: cardiovascular risk factors in middle-aged men in 24 towns. *BMJ* **283**, 179–186 (1981).
3. Soininen, P., Kangas, A. J., Würtz, P., Suna, T. & Ala-Korpela, M. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. *Circ. Cardiovasc. Genet.* **8**, 192–206 (2015).
4. Marmot, M. & Brunner, E. Cohort Profile: The Whitehall II study. *Int. J. Epidemiol.* **34**, 251–256 (2005).
5. Wadsworth, M., Kuh, D., Richards, M. & Hardy, R. Cohort Profile: The 1946 National Birth Cohort (MRC National Survey of Health and Development). *Int. J. Epidemiol.* **35**, 49–54 (2006).
6. Fone, D. L. *et al.* Cohort Profile: The Caerphilly Health and Social Needs Electronic Cohort Study (E-CATALyST). *Int. J. Epidemiol.* **42**, 1620–1628 (2013).
7. Tillin, T. *et al.* Southall And Brent REvisited: Cohort profile of SABRE, a UK population-based comparison of cardiovascular disease and diabetes in people of European, Indian Asian and African Caribbean origins. *Int. J. Epidemiol.* **41**, 33–42 (2012).
8. Lawlor, D. A., Bedford, C., Taylor, M. & Ebrahim, S. Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study. *J. Epidemiol. Community Health* **57**, 134–140 (2003).
9. Würtz, P. *et al.* High-throughput quantification of circulating metabolites improves prediction of subclinical atherosclerosis. *Eur. Heart J.* **33**, 2307–2316 (2012).
10. Voight, B. F. *et al.* Correction: The Metabochip, a Custom Genotyping Array for Genetic Studies of Metabolic, Cardiovascular, and Anthropometric Traits. *PLoS Genet.* **9**, (2013).
11. Deloukas, P. *et al.* Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat. Genet.* **45**, 25–33 (2013).
12. Kettunen, J. *et al.* Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat. Genet.* **44**, 269–276 (2012).
13. Nikpay, M. *et al.* A comprehensive 1000 Genomes–based genome-wide association meta-analysis of coronary artery disease. *Nat. Genet.* **47**, 1121 (2015).

4 Chapter 4 Establishing reference intervals for triglyceride and cholesterol concentrations in 14 lipoprotein subfraction metabolites measured using Nuclear Magnetic Resonance Spectroscopy in a UK population

Related publication

Joshi, R., Wannamethee, G., Engmann, J., Gaunt, T., Lawlor, D. A., Price, J., Papacosta, O., Shah, T., Tillin, T., Whincup, P., Chaturvedi, N., Kivimaki, M., Kuh, D., Kumari, M., Hughes, A. D., Casas, J. P., Humphries, S. E., Hingorani, A. D., Schmidt, A. F., & UCLEB Consortium (2021). Establishing reference intervals for triglyceride-containing lipoprotein subfraction metabolites measured using nuclear magnetic resonance spectroscopy in a UK population. *Annals of clinical biochemistry*, *58*(1), 47–53. https://doi.org/10.1177/0004563220961753

Data sources

UCLEB Consortium studies

- British Regional Heart Study (BRHS)
- Whitehall II study (WHII)
- Southall Brent REvisted study (SABRE)
- Caerphilly Prospective Study (CAPS)

Abstract

Background

Nuclear magnetic resonance (NMR) spectroscopy allows lipoproteins to be subclassified into 14 different classes based on particle size and lipid content. We recently showed that triglyceride in these subfractions have differential associations with cardiovascular disease (CVD) events. We report the distributions and define reference interval ranges for TG and cholesterol in 14 lipoprotein subfraction metabolites.

Methods

Lipoprotein subfractions using the Nightingale NMR platform were measured in 9,073 participants from 4 cohort studies contributing to the UCL-Edinburgh-Bristol (UCLEB) consortium. The distribution of each metabolite was assessed. Reference interval ranges were calculated for a disease-free population, by sex, age, BMI, smoking status and in participants with CVD or type 2 diabetes.

Results

The largest reference interval range for TG was observed in the medium VLDL subfraction (2.5^{th} 97.5^{th} percentile; 0.08 to 0.68 mmol/L), and for cholesterol in the large LDL subfraction (0.47-1.45 mmol/L). TG subfraction concentrations in VLDL, IDL, LDL and HDL sub-classes increased with increasing age and increasing BMI. Increases in cholesterol concentrations were largely comparable between men and women by age and smoking status with the exception of ever smokers in the HDL subclass. TG subfraction concentrations were significantly higher in ever smokers compared to never smokers, among those with clinical chemistry measured total TG

greater than 1.7 mmol/L, and in those with CVD , and type 2 diabetes as compared to disease-free subjects.

Conclusion

This is the first study to establish reference interval ranges for TG and cholesterol concentrations in 14 lipoprotein subfractions, in samples from the general population measured using the NMR platform. The utility of NMR lipid measures may lead to greater insights for the role of TG and cholesterol in CVD, emphasising the importance of appropriate reference interval ranges for future clinical decision making.

4.1 Introduction

Risk factors for atherosclerotic disease include elevated total cholesterol, LDL-cholesterol (LDL-C) and triglycerides (TG), and are used in disease risk assessment in clinical care[1]. Elevated LDL-C and TG are common among people with metabolic syndrome, obesity and type 2 diabetes (T2DM)[2,3], and are associated with an increased risk of cardiovascular disease (CVD)[4]. Reference intervals are the most common tool used for the interpretation of numerical pathology reports for comparisons to patient laboratory results to support clinical decision making. For example, population-based reference intervals are used as a tool to define thresholds for clinical decisions in the measurement of LDL-C for CVD risk assessment. LDL-C measurements together with other measurements such as body mass index (BMI) or systolic blood pressure (SBP) can help to determine if lipid-lowering is indicated[5].

The high-throughput proton (^{1}H) serum nuclear magnetic resonance (NMR) metabolomics platform developed by Nightingale provides quantitative information on lipoprotein particle size and lipid content representing multiple metabolic pathways[6–8]. NMR measures of lipoprotein subfractions are increasingly used in epidemiological and genetic studies, and may provide better insights into biological processes compared to clinical chemistry measures of total serum cholesterol or TG, which represent wide-ranging biological heterogeneity that is incorporated into a single measure [9]. Total cholesterol is a sum of all the cholesterol molecules in circulation and thus has no distinction for the lipoprotein particle it is carried with. The lipoprotein metabolism process is complex and involves various particle subclasses that have different and even opposite biological roles, for example LDL and high-density lipoprotein (HDL) particles being a well-known example in

reference to CVD risk. NMR quantification of lipoproteins from a serum sample capture a comprehensive molecular signature, which then allows the modelling of 14 lipoprotein subfractions characterised by particle size. Each particle gives a characteristic signal that is mechanically distinctive, the area of which is proportional to the concentration of the lipoprotein particle being identified. For each NMR measured lipoprotein subfraction, phospholipids, triglycerides, and esterified and free cholesterol are quantified, discussed in detail in preceding chapters.

The Nightingale NMR metabolomics approach has been used in two recent prospective cohort studies that found evidence to suggest a differential association of total serum TG and TG concentrations in subfractions with coronary heart disease (CHD) [10,11]. For example, total serum TG association with CHD was OR 1.19 (95% CI 1.10 to 1.28), TG in the VLDL subfractions was associated with CHD in the range of OR 1.12 to 1.22, whereas TG concentrations in the LDL subfractions conveyed a relatively lower risk (OR in the range 1.13 to 1.17)[11]. Similarly, cholesterol concentrations in VLDL, IDL and LDL subfractions had robust positive associations with CHD (OR in the range 1.14 to 1.28), whereas cholesterol concentrations in HDL subfractions were inversely associated (OR in the range 0.81 to 0.90).

NMR quantified lipoproteins better predict subclinical atherosclerosis when evaluated against conventional clinical chemistry lipid testing. In a study of 1595 individuals aged 24–39 years from the population-based Cardiovascular Risk in Young Finns Study, better prediction of 6-year incident high intima-media thickness (a marker for subclinical disease) was achieved when clinical chemistry (CC)

measured total and HDL-cholesterol were preplaced by NMR quantified lipoproteins[12].

The different associations of TG and cholesterol in lipoprotein subfractions with CHD and the increasing availability and implementation of NMR metabolomics in biobanks, epidemiological and, genetic studies, highlight the need to extend the standard lipid reference intervals to include TG and cholesterol lipoprotein subfractions[12].

This study aims to define reference intervals for TG and cholesterol concentrations in 14 lipoprotein subfractions using data from multiple UK based cohorts from the UCLEB consortium.

4.2 Methods

4.2.1 Population study sample

Data were sourced from the UCL-Edinburgh-Bristol (UCLEB) consortium, including NMR metabolite measures in 9,073 participants from 4 cohort studies: The British Regional Heart Study (BRHS), including men aged 60-79 at assessment in 1998-2000, the Whitehall II study (WHII), including UK government workers aged 45-69 years at assessment in 1997 to 1999, the Southall And Brent Revisited Study (SABRE), a tri-ethnic study including British men and women from European (SABRE1), South Asian (SABRE2) and African Caribbean (SABRE3) descent, and the Caerphilly Prospective Study (CAPS), including men registered in general practice aged 55-69 at assessment in 1989-1993. The design and data collection for the UCLEB Consortium of longitudinal population studies has been described in Chapter 3[13]. Age (years), sex (male/female), smoking (ever/never), body mass index (BMI), CHD, stroke and T2DM variables were collected at the time of NMR blood sample measurement.

4.2.2 Metabolite quantification

Using Nightingale NMR metabolomics platform[8], high-throughput metabolite quantification of TG and cholesterol (esterified and free cholesterol are summed in this study to represent the quantity of cholesterol present in each subfraction), in 14 lipoprotein subfractions (mmol/L) were ascertained in fasting and non-fasting serum samples in all contributing studies. To ensure long-term sample integrity, blood samples were stored and transported at -80 °C across all contributing UCLEB studies until NMR quantification in 2014. NMR metabolomics platform has been

extensively used in epidemiology and genetics studies[14–16], and its application reviewed and described in Chapter 1[6,17].

4.2.3 Statistical analysis

Individuals were removed based on any event of CHD, stroke or T2DM to include a healthy, 'disease free' population. The study-specific distribution of TG and cholesterol in each lipoprotein subfraction was first assessed. Data were then pooled using individual participant data form all four cohorts into one dataset. TG and cholesterol reference intervals for each subfraction were based on the 2.5th, and 97.5th percentiles stratified by age and sex. Age group bands were calculated as <55 years, 55-65 years and >65 years. The influence of age, smoking and BMI on the TG and cholesterol distributions were assessed statistically and graphically using "generalised linear model" (GAM) curves, and box plots. Given the influence of diet on total TG levels, the TG subfraction reference intervals were further assessed for influence of fasting status using the Kolmogorov–Smirnov (KS) test. TG and cholesterol reference intervals were additionally calculated in the following groups; 1) participants with CVD (defined as occurrence of either CHD or stroke, 2) participants with T2DM, 3) participants with clinical chemistry total TG greater, or less than 1.7 mmol/L or total cholesterol greater, or less than 5.2 mmol/L, for TG and cholesterol subfractions respectively and 4) TG and measured in the fasting and non-fasting state.

4.3 Results

4.3.1 Reference interval ranges

A total of 9,073 individuals were included in the main healthy, free of CVD and type 2 diabetes study sample, of which 5,574 (62.8%) were male (median age 61.7 years, IQR 52.0, 67.6), had median BMI of 26.0 (IQR 23.9, 28.4) kg/m^2 and 3,027 (54.2%) were current or ex-smokers (i.e., ever smokers). Women had a median age of 53.9 (IQR 49.9, 59.9), a BMI of 25.6 (23.6, 27.8) kg/m^2 and 431 (13.0%) were ever smokers. In general, concentrations of TG were highest in the VLDL subfractions, specifically medium and small VLDL, whereas cholesterol concentrations were mostly abundant in the LDL subfractions. Description of study population and median concentration of TG and cholesterol in 14 lipoprotein subfractions are shown in table 4.1.

Table 4.1 Description of study sample

	Men (n = 5574)	Women (n = 3299)
Age, years	61.7 (52.0, 67.6)	53.9 (49.9, 59.9)
BMI, kg/m2	26.0 (23.9, 28.4)	25.6 (23.6, 27.8)
Smoking, ever	3027/5574 (54.2)	431/3299 (13.0)
Triglyceride concentration (mmol/L)		
VLDL		
Extremely large	0.02 (0.01-0.03)	0.02 (0.01-0.02)
Very large	0.03 (0.01-0.05)	0.02 (0.01-0.04)
Large	0.10 (0.06-0.18)	0.09 (0.05-0.16)
Medium	0.23 (0.16-0.35)	0.24 (0.16-0.34)
Small	0.22 (0.18-0.29)	0.23 (0.17-0.30)
Very small	0.11 (0.09-0.13)	0.12 (0.09-0.14)
IDL	0.12 (0.10-0.14)	0.12 (0.10-0.15)
LDL		
Large	0.10 (0.08-0.12)	0.10 (0.09-0.12)
Medium	0.05 (0.04-0.06)	0.05 (0.04-0.05)
Small	0.03 (0.02-0.04)	0.03 (0.02-0.04)
HDL		
Very large	0.01 (0.01-0.02)	0.01 (0.01-0.02)
Large	0.03 (0.02-0.04)	0.02 (0.02-0.03)
Medium	0.05 (0.04-0.06)	0.05 (0.04-0.06)
Small	0.05 (0.04-0.06)	0.04 (0.04-0.05)
Cholesterol concentration (mmol/L)		
VLDL		
Extremely large	0.01 (<0.01-0.01)	<0.01 (<0.01-0.01)
Very large	0.01 (0.01-0.02)	0.01 (0.01-0.02)
Large	0.05 (0.03-0.05)	0.05 (0.03-0.07)
Medium	0.15 (0.11-0.21)	0.16 (0.12-0.21)
Small	0.24 (0.19-0.30)	0.27 (0.22-0.32)
Very small	0.28 (0.24-0.33)	0.32 (0.28-0.37)
IDL	0.73 (0.61-0.87)	0.83 (0.72-0.96)
LDL		
Large	0.89 (0.72-1.07)	0.97 (0.83-1.13)
Medium	0.50 (0.39-0.61)	0.54 (0.45-0.63)
Small	0.30 (0.24-0.37)	0.32 (0.27-0.38)
HDL		
Very large	0.20 (0.15-0.26)	0.20 (0.16-0.25)

Large	0.22 (0.14-0.35)	0.32 (0.24-0.41)
Medium	0.33 (0.26-0.42)	0.45 (0.39-0.52)
Small	0.44 (0.36-0.49)	0.42 (0.38-0.46)

Values are median (IQR) or %.

VLDL = Very-low density lipoprotein; IDL = Intermediate-density lipoprotein; LDL = Low-density lipoprotein; HDL = High-density lipoprotein

The sum of TG concentrations in the 14 lipoprotein subfractions was compared to CC measured total serum TG and found an increase of 0.34 mmol/L of NMR measured total TG for every 1 mmol/L increase in CC measured total serum TG, see appendix figures 4.1 and 4..2. The overall study population distribution for TG in the subfractions were comparable across contributing studies, showing agreement between ethnicities in the SABRE cohort (figure 4.1) and overlap of TG measured in the fasting and non-fasting state, see appendix figure 4.3. Of the 14 subfractions, TG in 12 subfractions had a skewed right tailed distribution, and two (medium and small HDL) had a more symmetrical distribution. When comparing the sum of cholesterol concentrations in 14 subfractions against CC measured total serum cholesterol, NMR measured cholesterol increased by 0.69 mmol/L per 1 mmol/L in CC cholesterol. Similar to TG concentrations, cholesterol distribution was comparable across the contributing UCLEB studies, see figure 2. Cholesterol in the VLDL subclass had right tailed distributions, cholesterol in the remaining LDL and HDL subclasses had more symmetrical distributions with the possible exception of cholesterol in small HDL.

Figure 4..1 The distribution of TG in 14 lipoprotein subfractions from contributing UCLEB studies

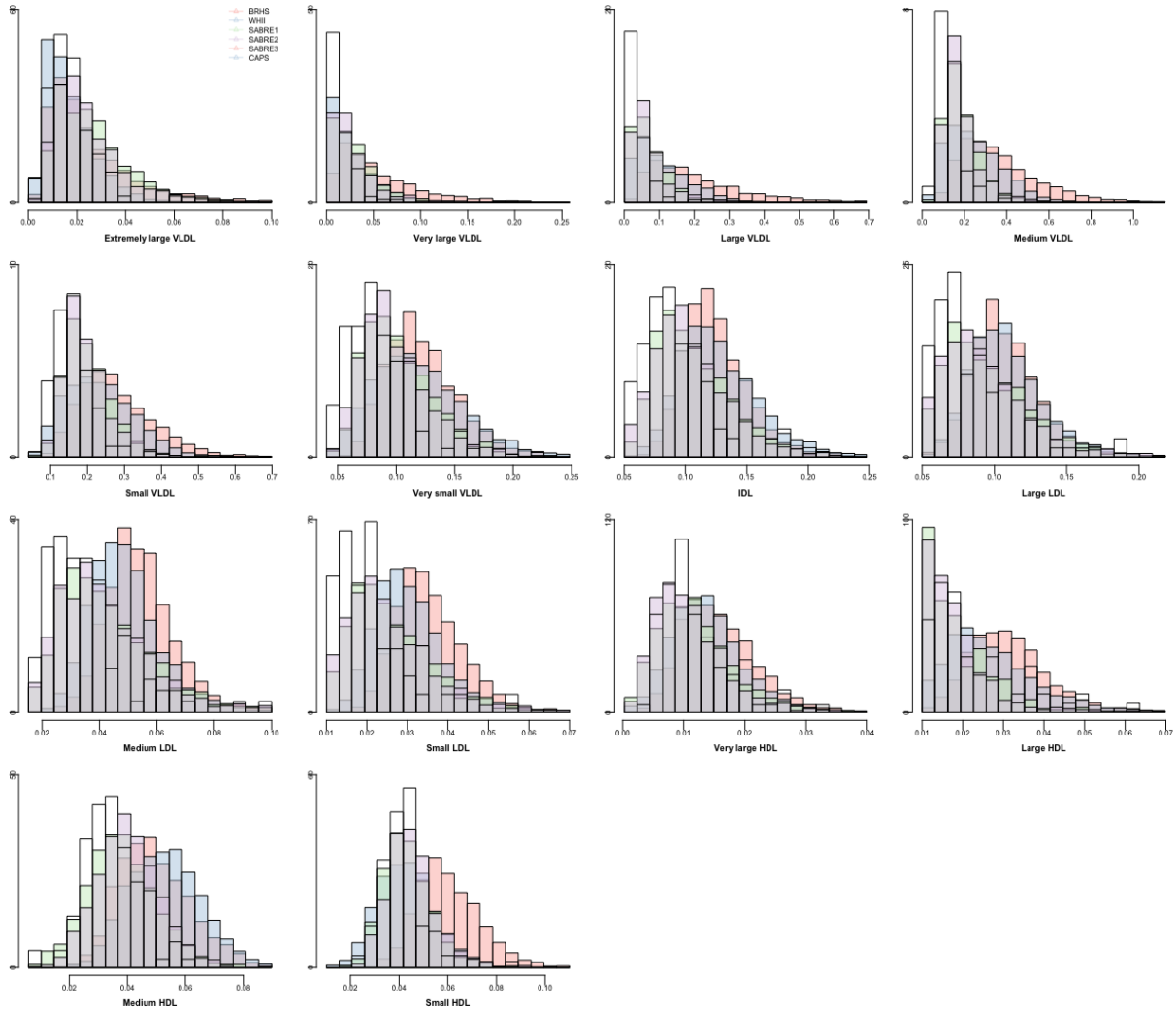
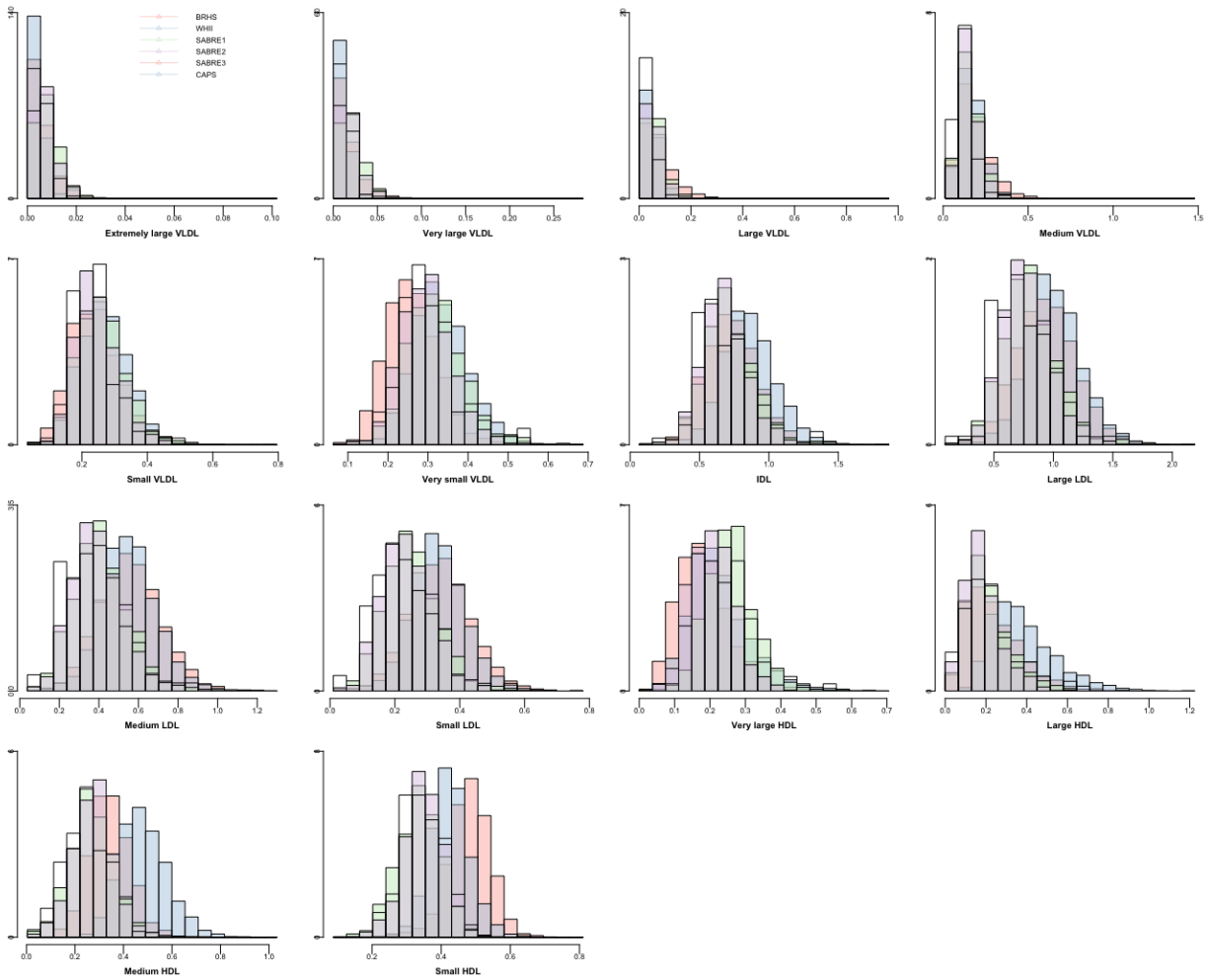


Figure 4.2 The distribution of cholesterol in 14 lipoprotein subfractions from contributing UCLEB studies



The reference intervals (2.5th – 97.5th percentile) for TG and cholesterol in the 14 subfractions are shown in table 4.2 and graphically in figure 4.3. Wide reference intervals were observed for TG in the VLDL subclass, for example the reference interval for TG in medium VLDL and small VLDL was, 0.08-0.67 mmol/L and 0.10-0.46 mmol/L, respectively. A smaller reference interval range was observed for TG in IDL, LDL and HDL subclass subfractions, for example, the reference interval range for TG in large HDL was 0.01-0.05 mmol/L. For cholesterol, a narrow reference interval range is observed in the VLDL subfractions whereas wide

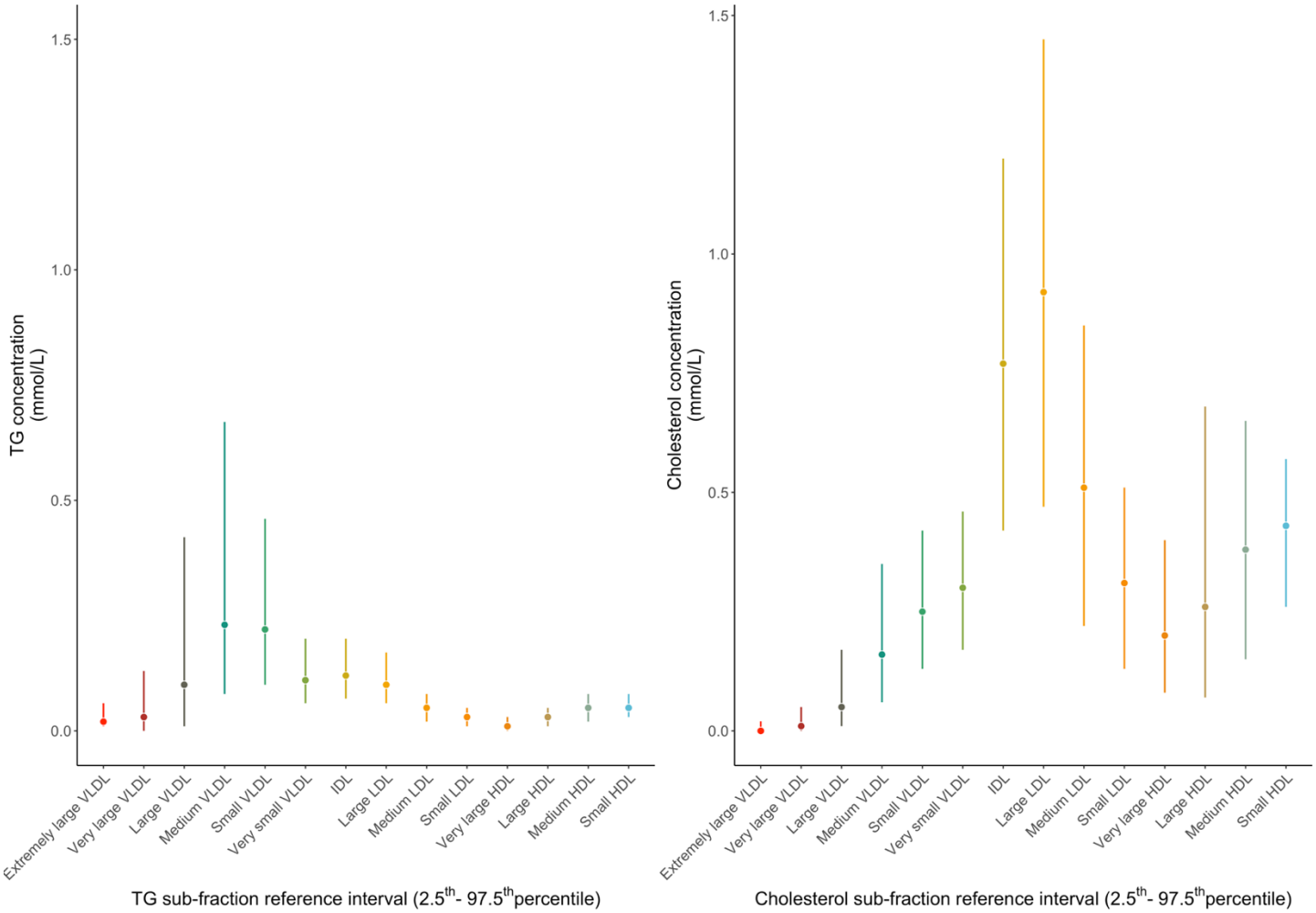
reference intervals are observed in the IDL (0.42-1.20 mmol/L), LDL and HDL subclass subfractions, specifically in the large LDL and large HDL subfractions for which the interval range was 0.47-1.45 and 0.07-0.68 mmol/L, respectively.

Table 4.2 Reference interval range of 14 TG and cholesterol subfractions (n=9,073)

	Triglyceride reference interval range (mmol/L)		Cholesterol reference interval range (mmol/L)	
Lipoprotein subfraction	2.50%	97.50%	2.50%	97.50%
VLDL				
Extremely large	0.01	0.06	<0.01	0.02
Very large	<0.01	0.13	<0.01	0.05
Large	0.01	0.42	0.01	0.17
Medium	0.08	0.67	0.06	0.35
Small	0.10	0.46	0.13	0.42
Very small	0.06	0.20	0.17	0.46
IDL	0.07	0.20	0.42	1.20
LDL				
Large	0.06	0.17	0.47	1.45
Medium	0.02	0.08	0.22	0.85
Small	0.01	0.05	0.13	0.51
HDL				
Very large	<0.01	0.03	0.08	0.40
Large	0.01	0.05	0.07	0.68
Medium	0.02	0.08	0.15	0.65
Small	0.03	0.08	0.26	0.57

VLDL = Very-low density lipoprotein; IDL = Intermediate-density lipoprotein; LDL = Low-density lipoprotein; HDL = High-density lipoprotein.

Figure 4.3 Reference interval range for 14 triglyceride-containing and cholesterol-containing lipoprotein subfractions (median, 2.5th, 97.5th) percentile



4.3.2 Age and sex stratified reference interval ranges

Age and sex stratified reference interval ranges for TG and cholesterol in the 14 subfractions are presented in appendix table 4.1. The reference interval ranges (2.5^{th}-97.5^{th} percentile) were largely comparable between men and women across VLDL subfractions. For example, among men aged <55 years, TG in the large VLDL reference interval was in the range 0.01 – 0.24 mmol/L, and for women of the same age band the reference interval range was 0.02-0.30 mmol/L. When considering cholesterol reference interval ranges, men overall had a wider interval range. For example, this is observed in cholesterol in large LDL, men aged 55-65 years had a range of 0.47-1.48 mmol/L, compared to women in the same age band had an interval range of 0.58-1.47 mmol/L.

4.3.3 Subgroup reference interval ranges

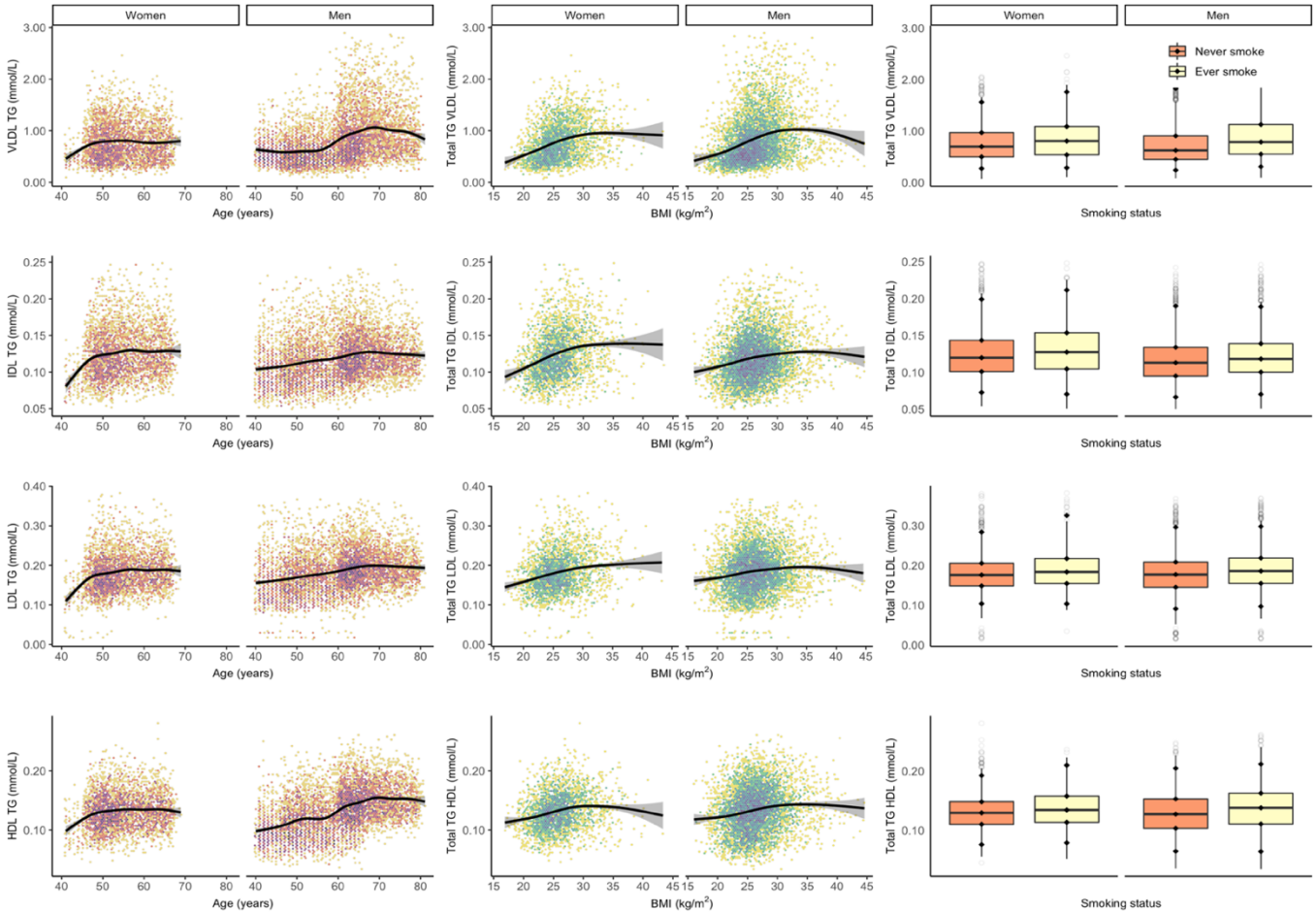
Figure 4 shows GAM curves and density distribution for the sum of TG in VLDL, IDL, LDL and HDL subclass subfractions for age and BMI, and box plot for TG distribution by smoking status. Among men, TG concentration increased with age, with the most prominent age differences observed in the HDL subclass (figure 4.4, left panel). By comparison, TG concentration differences were not as noticeable for women for whom concentrations were comparable for the VLDL, IDL, LDL and HDL subclasses by age. For both men and women, TG subfraction concentration increased with increasing BMI for the VLDL, IDL, LDL and HDL subclasses. Ever smokers had higher mean TG subfraction concentrations across all subclasses as compared to never smokers. Figure 4.5 shows the GAM curves and density distribution for cholesterol concentrations across VLDL, IDL, LDL and HDL subclasses. Increases in cholesterol concentrations were largely comparable between

men and women by age, with the possible exception of cholesterol among men in the HDL subclass for which there was an increase in concentration with increasing age until 50 years and then a plateau in concentration levels. Cholesterol concentrations increased steadily in the VLDL subclass with increasing BMI in both men and women, whereas concentrations remained level in the IDL and LDL subclasses. Conversely, in the HDL subclass, cholesterol concentration reduced with higher BMI, with the most prominent reduction seen in women and a U-shaped curve among men. Mean cholesterol concentration by smoking status was similar between men and women, with the exception of ever smokers in the HDL subclass, in which the mean concentration was lower than that observed among never smokers.

The reference interval ranges for TG in the 14 subfractions were comparable in the population with CVD (N = 2719 and those with T2DM (N = 1325) to the reference interval range reported in the 'disease-free' group in table 4.2. In general, the largest variation in reference interval ranges across the subfractions between the disease sub-groups was observed for TG in the VLDL subclass, appendix table 4.2. For example, the 2.5^{th} to 97.5^{th} reference interval range for TG in medium VLDL in the different groups were 0.08-0.67, CVD: 0.09, 0.79 mmol/L and in T2DM: 0.08, 0.84 mmol/L. Conversely, cholesterol concentrations and reference interval ranges were overall higher in the across the 14 sun-fractions in the 'disease free' group as compared to the CVD and T2DM groups. For example, the interval range for cholesterol in the large LDL subfraction was 0.47- 1.45, 0.44-1.39 and 0.36-1.38 in the CVD and T2SDM groups respectively.

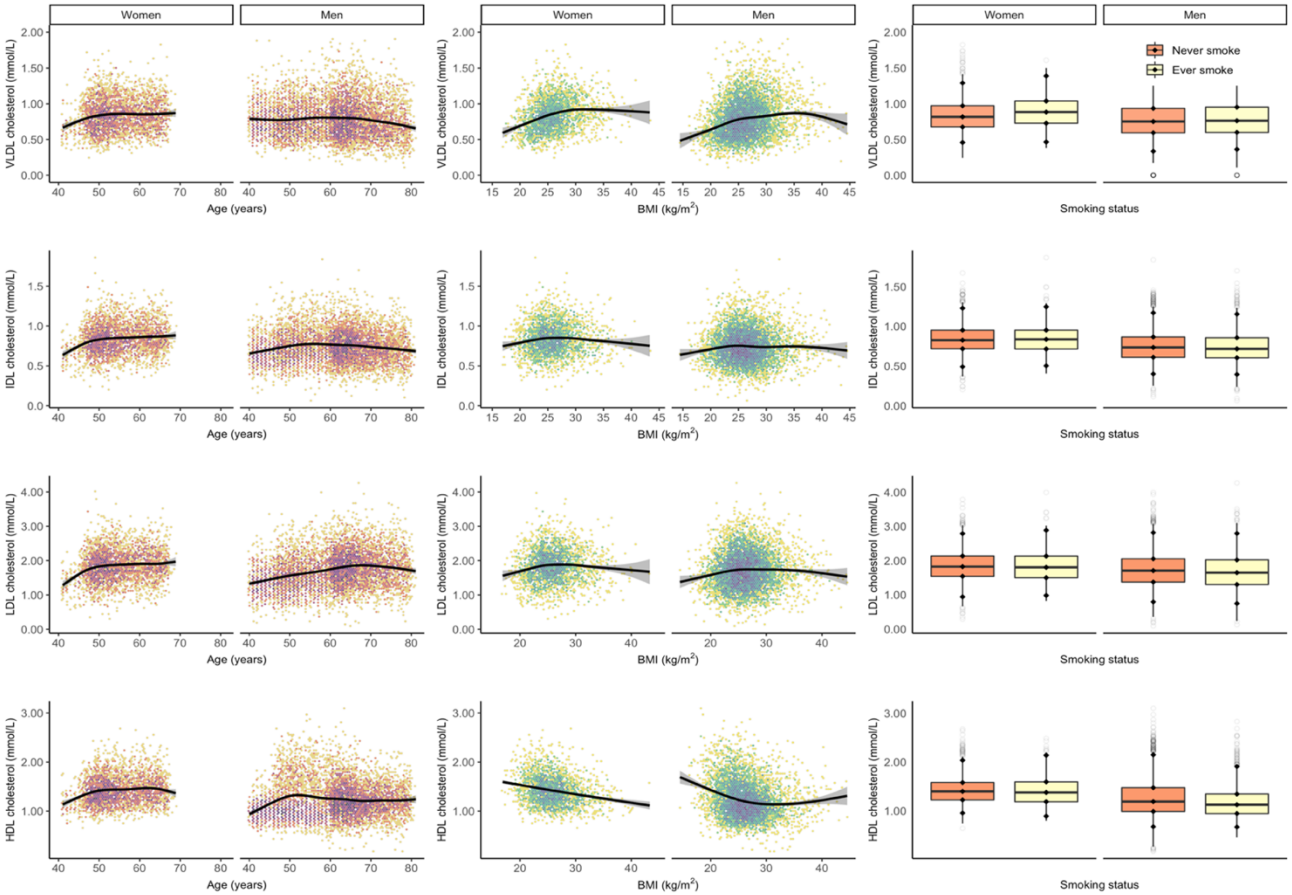
The reference interval ranges were assessed in subgroups stratified by clinically recognised lipid targets as lower than, 1.7mmol/L and lower than 5.2 mmol/L for total TG and total cholesterol respectively. TG and cholesterol concentrations in the 14 subfractions were higher in the group with CC measured total TG greater than 1.7mmol/L, and total cholesterol greater than 5.2 mmol/L. TG subfraction concentrations and reference interval range in the 14 subfractions were comparable in TG measured in the fasting and non-fasting state.

Figure 4.4 Distribution of TG concentration in VLDL, IDL, LDL and HDL subclass by age (left panel), body mass index (centre panel) and smoking status (right panel)



N.b. Slope indicates a GAM estimate with 95% confidence interval. Tile colours represent the number of observations, with purple coloured tiles indicating a higher density. Smoking distribution (right panel) is based on data from BRHS, SABRE and WHII studies. All were significant at P value threshold for <0.001

Figure 4.5 Distribution of cholesterol concentration in VLDL, IDL, LDL and HDL subclass by age (left panel), body mass index (centre panel) and smoking status (right panel



N.b. Slope indicates a GAM estimate with 95% confidence interval. Tile colours represent the number of observations, with purple coloured tiles indicating a higher density. Smoking distribution (right panel) is based on data from BRHS, SABRE and WHII studies. All were significant at P value threshold for <0.001

4.4 Discussion

This study provides reference interval ranges (2.5^{th} to 97.5^{th} percentiles) for TG and cholesterol in 14 lipoprotein subfraction metabolites as measured by NMR spectroscopy based on a sample of UK adults. There was agreement in the distribution of triglyceride subfraction and cholesterol subfraction concentrations between ethnicities. Triglyceride concentrations for men and women increase with increasing age and BMI, are higher among ever smokers and in those with CVD and T2DM as compared to disease-free subjects, and in individuals with total TG concentrations greater than 1.7 mmol/L. Cholesterol concentrations gradually increase with increasing age and BMI in VLDL, IDL LDL subclass, with the most prominent increase observed in the VLDL subclass.

Lipid reference interval ranges are derived using clinical chemistry measurement of blood samples from a reference population and are necessary to support clinical decision making and apply analytical data in healthcare delivery. For example, clinical chemistry estimates of LDL-C are measured in individuals and evaluated against an interval range to inform lifestyle or therapeutic intervention for CHD prevention. While there is incontrovertible evidence that LDL-C has a causal role, the role of TG in CHD risk is less clear. Meta-analysis from prospective observational studies have demonstrated higher concentrations of clinical chemistry measured total serum TG are associated with higher risk of CHD, but effect estimates attenuate to the null after adjustment for HDL-C (17). On the other hand, Mendelian randomisation studies support a potential causal association for TG (18). The association of the major blood lipid fractions (LDL-C, TG and HDL-C) with CHD is seen across the whole of the concentration range, with no threshold value

and the same is expected to be true of TG in the lipoprotein subfractions. The reference ranges reported here should not therefore be taken to imply that individuals whose measurements lie within these ranges are free of CHD risk. Rather, reported here are the observed values in general UK populations.

NMR methodology offers the potential for more granular quantification of TG and cholesterol in different lipoproteins that would otherwise be unavailable using conventional approaches, enabling a more detailed investigation of TG and cholesterol lipoprotein subfractions in relation CHD risk and prognosis. Evidence from studies using this approach suggest CHD risk may be divergent depending on the type of lipoprotein subfraction. Two recent studies report observations of TG in VLDL subfractions may be more atherogenic and associated with a higher risk of CHD compared to TG in the IDL, LDL and HDL subclass subfractions[10,11].

4.4.1 Research in context

This study evaluates the concentration distribution and range of TG and cholesterol in 14 subfractions and includes data from multiple UK population cohorts and from men and women from a range of age groups and ethnicities including European, South Asian and African-Caribbean ancestry. Total serum TG and cholesterol measured using clinical chemistry and the sum of NMR measured TG and cholesterol across the 14 subfractions. Discrepancies between clinical chemistry methods and NMR measured total TG have been reported previously[10,18]. In one such study, Balling, 2019 suggests differences in analytical calibration from measurement of TG between the two methods may lead to measurement differences, with NMR quantification deemed as the more accurate method[18,19].

4.4.2 Strengths and limitations

TG concentrations in lipoproteins are in a constant state of flux and are highly variable and, in addition to age, sex and ethnicity, can depend on factors such as food intake, fasting/non-fasting state, CVD and metabolic disorders such as type 2 diabetes[20]. Due to the relatively large sample sizes available, we observed a significant difference between the fasting and non-fasting distribution TG in the subfractions. This significant difference did not prove relevant for determining the reference interval ranges, which were comparable to 2 decimal points. Moreover, it is postulated that due to varying food in-take patterns, the non-fasting state predominates the fasting state in 24-hour cycle as fasting for more than 8 hours normally only occurs before breakfast. Nordestgaard and colleagues report the maximal mean changes measured in random non-fasting versus fasting blood samples are +0.3 mmol/L TG, -0.2mmol/L total cholesterol, -0.2 mmol/L LDL-C and -0.2mmol/L non-HDL cholesterol measures do not translate to clinically significant differences, especially when evaluating CVD risk[21]. A shift away from the longstanding tradition of using fasting to non-fasting lipid profiles is endorsed in multiple guidelines. This shift has been seen in countries including, Denmark, the United Kingdom, Europe, Canada and Brazil following the consensus view that non-fasting lipid profiles represent a simplified process for both clinicians and patients, without negative implications for prognostic or diagnostic options, for example in the case of CVD prevention[22,23]. Due to small numbers of current smokers in the available data across the contributing cohorts we stratified by ever and never smoking status, instead of the more informative "never", "ex" and "current" smokers. Participants with current CVD or T2DM were excluded in the main analyses, however it is possible TG or cholesterol levels in the study population were

altered by other diseases or by lipid lowering medication, which we were not able to account for in this study. It is likely TG and cholesterol monitoring is likely to occur in individuals at risk of, or with current CVD. Therefore, we provide additional reference intervals in participants with CVD and T2D. Cholesterol levels increase with increasing age, as shown in the main figure above and in the reference range interval for cholesterol in the 14 subfractions among the 'healthy, disease-free' population. Lower cholesterol concentrations across the 14 subfractions were observed among the group with CVD, likely due to lipid-lowering medications by statins. The same was not seen for TG concentrations, most likely because TG lowering is not normally prescribed except in extreme circumstances when serum TG levels are greater than 11mmol/L, which can lead to acute pancreatitis. With the possible exception of the omega-3 fatty acid trial REDUCE-IT[24], there is an absence of convincing CVD benefit from trials of TG lowering using niacin or fibrates.

This study establishes reference interval ranges for TG and cholesterol in 14 lipoprotein subfractions for men and women by age, BMI, smoking status, CVD, T2DM and stratified by clinical chemistry measured total TG and cholesterol, and TG fasting status for population-based cohorts from the UK population. Further studies would be needed to assess if the reference intervals presented here could be extended to a non-UK population and if the risks associated with the reference intervals identify a threshold within these ranges to inform CVD risk in a clinical setting. By doing so, the reference interval ranges may help to set realistic targets and guide research interests, contributing to the development of effective targeted TG lowering therapies, aimed at for example TG in VLDL subfractions which may be the most atherogenic[10]. NMR lipoprotein particle number and size have been

assessed in relation to CHD, however this study specifically presents reference range intervals for TG and cholesterol within the 14 subfractions[25]. Further investigations would be needed to compare subfraction lipid composition, particle number concentration and size.

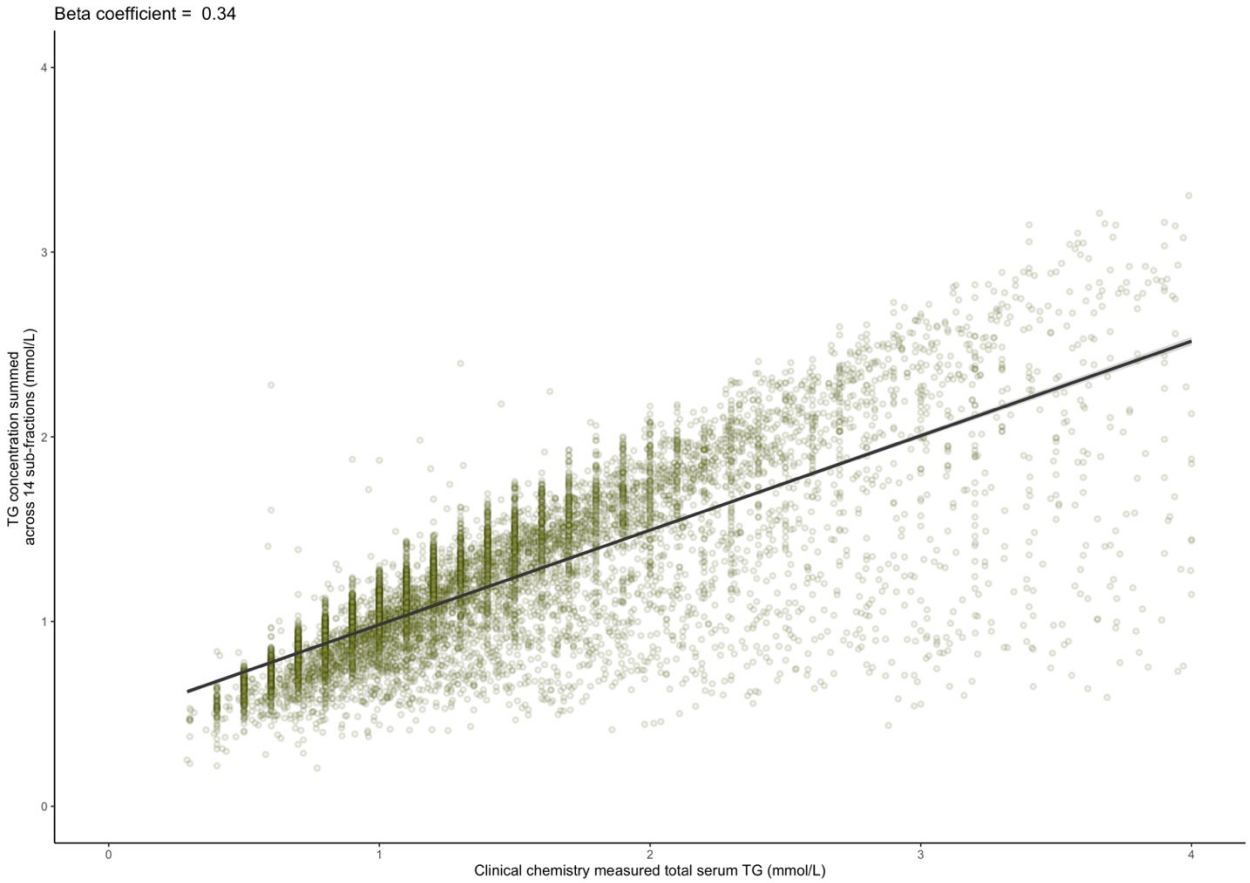
Metabolomics is becoming integrated with genomics to contribute to a better understanding of disease aetiologies and disease risk[26]. It is likely that quantitative metabolomics will be incorporated into large biobanks, which would extend the relevance of sample collection and encourage the life-long assessment of metabolic health[6]. TG and cholesterol subfraction reference interval ranges may help complement current routine clinical chemistry measures of lipids and become an integral tool in targeted patient management and improved disease risk prediction and prevention.

4.5 Conclusion

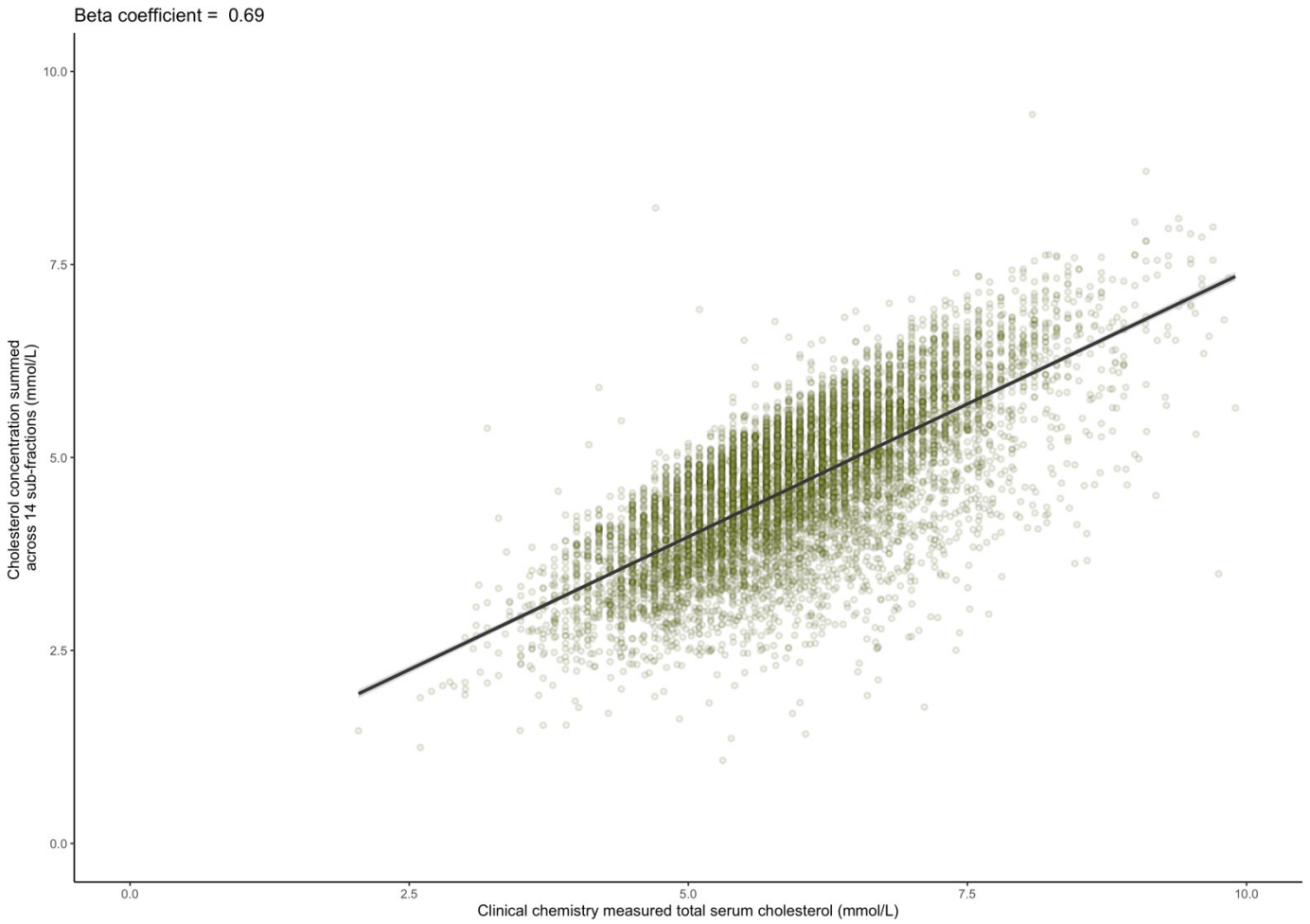
This study is the first to establish reference interval ranges for TG and cholesterol in 14 lipoprotein subfraction metabolites, measured using the Nightingale NMR platform for men and women in a UK population. NMR measures of lipoproteins may provide insights into biological processes compared to clinical chemistry measures of TG and cholesterol and lead to greater insights for the role of TG in CVD, emphasising the importance of appropriate reference interval ranges for future clinical decision making.

4.6 Chapter 4 appendix

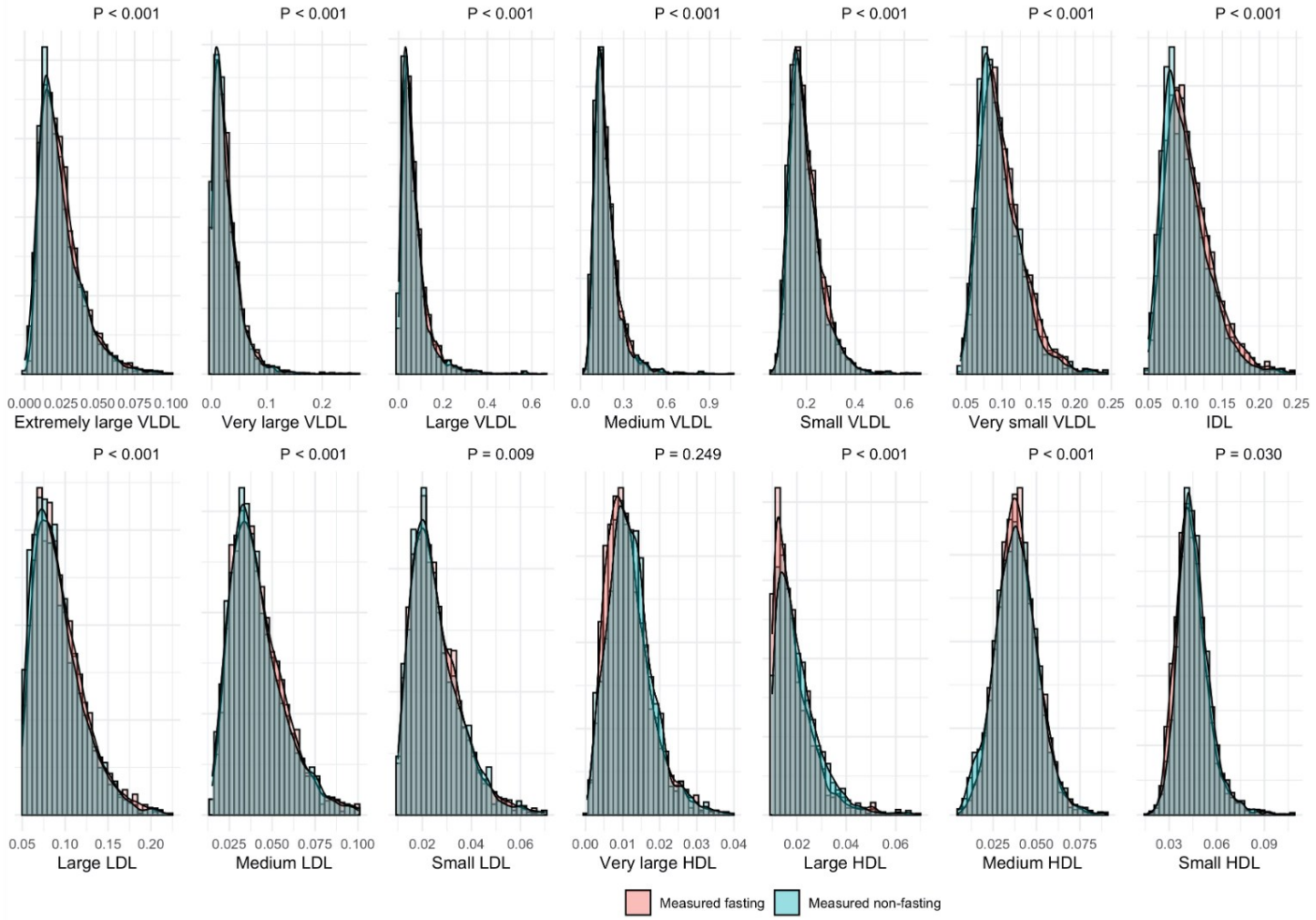
Appendix figure 4.1 Scatter plot to show the sum of TG across 14 subfraction vs clinical chemistry measured total TG



Appendix figure 4.2 Scatter plot to show the sum of cholesterol across 14 subfraction vs clinical chemistry measured total TG



Appendix figure 4.3 Comparison of TG in 14 subfractions measured in the fasting vs non-fasting using data from the SABRE cohort



Appendix table 4.1 Age and sex stratified reference interval ranges (2.5th, median 97.th percentile)

	Men								
	<55 (n = 1825)			55 to 65 (n = 1910)			>65 (n = 1921)		
Triglyceride concentration (mmol/L)	2.50%	median	97.50%	2.50%	median	97.50%	2.50%	median	97.50%
Extremely large VLDL	<0.01	0.02	0.06	0.01	0.02	0.07	0.01	0.02	0.07
Very large VLDL	<0.01	0.02	0.09	<0.01	0.03	0.15	0.01	0.04	0.16
Large VLDL	0.01	0.06	0.24	0.01	0.10	0.48	0.04	0.16	0.52
Medium VLDL	0.07	0.16	0.43	0.07	0.24	0.73	0.12	0.32	0.82
Small VLDL	0.09	0.18	0.37	0.10	0.23	0.48	0.13	0.26	0.51
Very small VLDL	0.05	0.01	0.18	0.06	0.11	0.19	0.07	0.12	0.19
IDL	0.06	0.11	0.19	0.07	0.12	0.19	0.08	0.12	0.19
Large LDL	0.05	0.09	0.17	0.06	0.10	0.17	0.07	0.11	0.17
Medium LDL	0.02	0.04	0.08	0.03	0.05	0.08	0.03	0.05	0.08
Small LDL	0.01	0.03	0.05	0.01	0.03	0.05	0.02	0.03	0.05
Very large HDL	<0.01	0.01	0.03	<0.01	0.01	0.03	0.01	0.02	0.03
Large HDL	0.01	0.02	0.05	0.01	0.03	0.05	0.02	0.03	0.06
Medium HDL	0.02	0.04	0.07	0.02	0.05	0.08	0.03	0.05	0.08
Small HDL	0.02	0.04	0.07	0.03	0.05	0.08	0.04	0.06	0.09
	Women								
	<55 (n = 1839)			>55 to <65 (n =1243)			>65 (n =204)		
	2.50%	median	97.50%	2.50%	median	97.50%	2.50%	median	97.50%
Extremely large VLDL	0.01	0.02	0.05	0.01	0.02	0.04	0.01	0.02	0.04
Very large VLDL	<0.01	0.02	0.09	<0.01	0.02	0.08	<0.01	0.02	0.08
Large VLDL	0.02	0.09	0.30	0.02	0.09	0.29	0.03	0.10	0.29
Medium VLDL	0.08	0.23	0.54	0.08	0.24	0.54	0.10	0.25	0.54

Small VLDL	0.10	0.23	0.44	0.10	0.23	0.43	0.12	0.24	0.42
Very small VLDL	0.06	0.11	0.20	0.06	0.12	0.20	0.07	0.12	0.19
IDL	0.07	0.12	0.20	0.08	0.12	0.20	0.08	0.12	0.20
Large LDL	0.06	0.10	0.17	0.07	0.11	0.16	0.07	0.11	0.17
Medium LDL	0.02	0.04	0.08	0.03	0.05	0.08	0.03	0.05	0.07
Small LDL	0.01	0.03	0.05	0.02	0.03	0.05	0.02	0.03	0.05
Very large HDL	<0.01	0.01	0.03	<0.01	0.01	0.03	0.01	0.01	0.02
Large HDL	0.01	0.02	0.05	0.01	0.02	0.05	0.01	0.02	0.05
Medium HDL	0.03	0.05	0.08	0.03	0.05	0.08	0.03	0.05	0.07
Small HDL	0.03	0.04	0.07	0.02	0.04	0.07	0.03	0.04	0.06
	Men								
	<55 (n = 1825)			55 to 65 (n = 1910)			>65 (n = 1921)		
Cholesterol concentration (mmol/L)	2.50%	median	97.50%	2.50%	median	97.50%	2.50%	median	97.50%
Extremely large VLDL	<0.01	0.01	0.02	<0.01	<0.01	0.02	<0.01	<0.01	0.02
Very large VLDL	<0.01	0.01	0.05	<0.01	0.01	0.05	<0.01	0.01	0.05
Large VLDL	0.01	0.05	0.14	0.01	0.05	0.20	0.01	0.06	0.21
Medium VLDL	0.06	0.14	0.29	0.06	0.16	0.37	0.05	0.16	0.40
Small VLDL	0.13	0.24	0.40	0.13	0.25	0.42	0.11	0.24	0.41
Very small VLDL	0.19	0.30	0.45	0.15	0.29	0.47	0.15	0.26	0.42
IDL	0.40	0.71	1.16	0.41	0.75	1.21	0.39	0.72	1.14
Large LDL	0.41	0.81	1.37	0.47	0.92	1.48	0.47	0.93	1.47
Medium LDL	0.17	0.42	0.76	0.22	0.52	0.86	0.26	0.55	0.88
Small LDL	0.11	0.26	0.46	0.14	0.32	0.53	0.16	0.34	0.53
Very large HDL	0.11	0.23	0.43	0.08	0.21	0.41	0.06	0.18	0.39
Large HDL	0.05	0.21	0.77	0.06	0.23	0.72	0.06	0.22	0.59
Medium HDL	0.11	0.31	0.67	0.14	0.35	0.66	0.19	0.33	0.52

Small HDL	0.23	0.37	0.51	0.24	0.44	0.58	0.32	0.48	0.58
	Women								
	<55 (n = 1839)			>55 to <65 (n =1243)			>65 (n =204)		
	2.50%	median	97.50%	2.50%	median	97.50%	2.50%	median	97.50%
Extremely large VLDL	<0.01	<0.01	0.01	<0.01	<0.01	0.01	<0.01	<0.01	0.01
Very large VLDL	<0.01	0.01	0.03	<0.01	0.01	0.03	<0.01	0.01	0.03
Large VLDL	0.01	0.05	0.12	0.01	0.05	0.12	0.02	0.05	0.12
Medium VLDL	0.07	0.16	0.32	0.07	0.17	0.31	0.08	0.17	0.30
Small VLDL	0.14	0.26	0.42	0.16	0.27	0.41	0.16	0.28	0.41
Very small VLDL	0.20	0.32	0.47	0.22	0.33	0.47	0.22	0.34	0.46
IDL	0.47	0.81	1.23	0.52	0.85	1.21	0.54	0.85	1.26
Large LDL	0.51	0.95	1.45	0.58	1.00	1.47	0.62	0.99	1.55
Medium LDL	0.24	0.52	0.83	0.29	0.56	0.83	0.32	0.54	0.87
Small LDL	0.14	0.31	0.49	0.17	0.34	0.50	0.20	0.33	0.53
Very large HDL	0.08	0.20	0.37	0.09	0.21	0.40	0.10	0.21	0.40
Large HDL	0.14	0.31	0.63	0.15	0.33	0.69	0.16	0.31	0.60
Medium HDL	0.25	0.45	0.66	0.27	0.46	0.68	0.27	0.44	0.61
Small HDL	0.30	0.41	0.54	0.31	0.42	0.55	0.33	0.42	0.53

Appendix table 4.2 Reference interval range of TG and cholesterol in 14 subfractions stratified by CVD, T2DM, clinical chemistry measured total TG and cholesterol and, TG measured in the fasting and non-fasting state.							
CVD (n = 2719)							
	TG concentration				Cholesterol concentration		
	2.50%	median	97.50%		2.50%	median	97.50%
Extremely large VLDL	0.01	0.02	0.07		<0.01	0.01	0.02
Very large VLDL	<0.01	0.04	0.17		<0.01	0.02	0.06
Large VLDL	0.02	0.13	0.53		0.02	0.07	0.23
Medium VLDL	0.09	0.29	0.79		0.06	0.18	0.40
Small VLDL	0.12	0.25	0.51		0.12	0.25	0.43
Very small VLDL	0.07	0.12	0.20		0.14	0.27	0.45
IDL	0.07	0.12	0.20		0.37	0.70	1.15
Large LDL	0.06	0.10	0.17		0.44	0.86	1.39
Medium LDL	0.03	0.05	0.08		0.20	0.48	0.81
Small LDL	0.02	0.03	0.05		0.11	0.29	0.49
Very large HDL	<0.01	0.01	0.03		0.06	0.18	0.36
Large HDL	0.01	0.03	0.05		0.05	0.19	0.53
Medium HDL	0.03	0.05	0.08		0.11	0.31	0.57
Small HDL	0.03	0.05	0.09		0.24	0.42	0.55
T2DM (n = 1325)							
Extremely large VLDL	0.01	0.02	0.08		<0.01	0.01	0.02
Very large VLDL	<0.01	0.04	0.18		<0.01	0.02	0.07
Large VLDL	0.02	0.13	0.55		0.01	0.07	0.25
Medium VLDL	0.08	0.27	0.84		0.06	0.18	0.44
Small VLDL	0.12	0.25	0.52		0.13	0.25	0.43
Very small VLDL	0.06	0.12	0.21		0.14	0.28	0.44
IDL	0.07	0.12	0.21		0.35	0.70	1.12
Large LDL	0.06	0.11	0.18		0.36	0.83	1.38
Medium LDL	0.02	0.05	0.09		0.15	0.46	0.80
Small LDL	0.01	0.03	0.06		0.08	0.28	0.48
Very large HDL	<0.01	0.01	0.03		0.06	0.18	0.36
Large HDL	0.01	0.02	0.05		0.05	0.2	0.58
Medium HDL	0.03	0.05	0.08		0.10	0.32	0.58
Small HDL	0.03	0.05	0.09		0.22	0.41	0.55
	CC measured total TG < 1.7 mmol/L (n = 6076)				CC measured total cholesterol < 5.2 mmol/L (n = 1977)		
Extremely large VLDL	<0.01	0.01	0.03		<0.01	<0.01	0.01

Very large VLDL	<0.01	0.02	0.05		<0.01	0.01	0.03
Large VLDL	0.01	0.07	0.19		0.01	0.04	0.12
Medium VLDL	0.07	0.19	0.37		0.05	0.11	0.26
Small VLDL	0.09	0.19	0.32		0.10	0.19	0.29
Very small VLDL	0.06	0.10	0.16		0.14	0.24	0.33
IDL	0.07	0.11	0.17		0.33	0.59	0.81
Large LDL	0.06	0.10	0.15		0.36	0.69	0.94
Medium LDL	0.02	0.04	0.07		0.16	0.37	0.54
Small LDL	0.01	0.03	0.04		0.10	0.23	0.33
Very large HDL	<0.01	0.01	0.02		0.06	0.18	0.36
Large HDL	0.01	0.03	0.05		0.07	0.26	0.65
Medium HDL	0.02	0.04	0.07		0.16	0.36	0.64
Small HDL	0.02	0.04	0.06		0.25	0.40	0.52
	CC measured total TG > 1.7 mmol/L (n = 2860)				CC measured total cholesterol > 5.2 mmol/L (n = 6719)		
Extremely large VLDL	0.01	0.03	0.08		<0.01	0.01	0.02
Very large VLDL	0.01	0.06	0.17		<0.01	0.01	0.05
Large VLDL	0.03	0.22	0.53		0.01	0.06	0.18
Medium VLDL	0.12	0.41	0.82		0.07	0.17	0.36
Small VLDL	0.15	0.32	0.53		0.15	0.27	0.42
Very small VLDL	0.08	0.14	0.22		0.20	0.32	0.47
IDL	0.08	0.14	0.21		0.51	0.82	1.23
Large LDL	0.07	0.12	0.18		0.57	0.99	1.49
Medium LDL	0.03	0.06	0.09		0.27	0.56	0.86
Small LDL	0.02	0.04	0.06		0.16	0.34	0.52
Very large HDL	<0.01	0.02	0.03		0.09	0.21	0.41
Large HDL	0.01	0.03	0.05		0.06	0.26	0.69
Medium HDL	0.03	0.06	0.08		0.15	0.38	0.65
Small HDL	0.04	0.06	0.09		0.26	0.43	0.58
TG measured in the fasting state (N= 2273)							
Extremely large VLDL	0.01	0.02	0.06				
Very large VLDL	<0.01	0.02	0.09				
Large VLDL	0.01	0.06	0.24				
Medium VLDL	0.08	0.16	0.42				
Small VLDL	0.11	0.19	0.36				
Very small VLDL	0.06	0.09	0.17				
IDL	0.06	0.1.0	0.18				
Large LDL	0.05	0.09	0.17				

Medium LDL	0.02	0.04	0.08
Small LDL	0.01	0.02	0.05
Very large HDL	<0.01	0.01	0.03
Large HDL	0.01	0.02	0.04
Medium HDL	0.02	0.04	0.06
Small HDL	0.03	0.04	0.07
Extremely large VLDL	0.01	0.02	0.06
Very large VLDL	<0.01	0.02	0.10
Large VLDL	0.01	0.05	0.25
Medium VLDL	0.07	0.16	0.44
Small VLDL	0.10	0.18	0.36
Very small VLDL	0.06	0.09	0.17
IDL	0.06	0.09	0.17
Large LDL	0.05	0.09	0.16
Medium LDL	0.02	0.04	0.08
Small LDL	0.01	0.02	0.05
Very large HDL	<0.01	0.01	0.03
Large HDL	0.01	0.02	0.04
Medium HDL	0.01	0.04	0.06
Small HDL	0.03	0.04	0.07

4.7 References

1. Nordestgaard, B. G. & Varbo, A. Triglycerides and cardiovascular disease. *Lancet* **384**, 626–635 (2014).
2. Toth, P. P. Triglyceride-rich lipoproteins as a causal factor for cardiovascular disease. *Vasc. Health Risk Manag.* **12**, 171–83 (2016).
3. Hegele, R. A. *et al.* The polygenic nature of hypertriglyceridaemia: Implications for definition, diagnosis, and management. *The Lancet Diabetes and Endocrinology* vol. 2 (2014).
4. Miller, M. *et al.* Triglycerides and Cardiovascular Disease. *Circulation* **123**, 2292–2333 (2011).
5. Pencina, M. J. *et al.* Predicting the 30-year risk of cardiovascular disease: the framingham heart study. *Circulation* **119**, 3078–84 (2009).
6. Soininen, P., Kangas, A. J., Würtz, P., Suna, T. & Ala-Korpela, M. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. *Circ. Cardiovasc. Genet.* **8**, 192–206 (2015).
7. Ala-Korpela, M., Kangas, A. J. & Soininen, P. Quantitative high-throughput metabolomics: a new era in epidemiology and genetics. *Genome Med.* **4**, 36 (2012).
8. Lifelong health belongs to everyone. https://nightingalehealth.com/.
9. Würtz, P. *et al.* Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Large-Scale Epidemiology: A Primer on -Omic Technologies. *Am. J. Epidemiol.* **186**, 1084–1096 (2017).
10. Holmes, M. V. *et al.* Lipids, Lipoproteins, and Metabolites and Risk of Myocardial Infarction and Stroke. *J. Am. Coll. Cardiol.* (2018) doi:10.1016/j.jacc.2017.12.006.
11. Roshni Joshi*a, S Goya Wannamethee*b, Jorgen Engmanna, Caroline Dalea, Tom Gauntc, Barbara Jefferisb, Deborah A Lawlorc-e, Jackie Pricef, Olia Papacostab, Tina Shaha, Therese Tilling, Nishi Chaturvedig, Mika Kivimakig, Diana Kuhh, Meena Kumarii, Alun D Hu, j on behalf of the U. C. Triglyceride-containing lipoprotein subfractions and risk of coronary heart disease and stroke: a prospective analysis in 11,560 adults.
12. Würtz, P. *et al.* High-throughput quantification of circulating metabolites improves prediction of subclinical atherosclerosis. *Eur. Heart J.* **33**, 2307–2316 (2012).
13. Shah, T. *et al.* Population Genomics of Cardiometabolic Traits: Design of the University College London-London School of Hygiene and Tropical Medicine-Edinburgh-Bristol (UCLEB) Consortium. *PLoS One* **8**, e71345 (2013).

14. Wurtz, P. *et al.* Circulating Metabolite Predictors of Glycemia in Middle-Aged Men and Women. *Diabetes Care* **35**, 1749–1756 (2012).
15. Würtz, P. *et al.* Characterization of systemic metabolic phenotypes associated with subclinical atherosclerosis. *Mol. BioSyst.* **7**, 385–393 (2011).
16. Würtz, P. *et al.* Metabolite Profiling and Cardiovascular Event Risk. *Circulation* **131**, 774–785 (2015).
17. Soininen, P. *et al.* High-throughput serum NMR metabonomics for cost-effective holistic studies on systemic metabolism. *Analyst* **134**, 1781 (2009).
18. Balling, M. *et al.* A third of nonfasting plasma cholesterol is in remnant lipoproteins: Lipoprotein subclass profiling in 9293 individuals. *Atherosclerosis* **286**, 97–104 (2019).
19. Holmes, M. V. & Ala-Korpela, M. What is 'LDL cholesterol'? *Nature Reviews Cardiology* vol. 16 197–198 (2019).
20. Brunzell, J. D. Hypertriglyceridemia. *N. Engl. J. Med.* **357**, 1009–1017 (2007).
21. Nordestgaard, B. G. A Test in Context: Lipid Profile, Fasting Versus Nonfasting. *Journal of the American College of Cardiology* vol. 70 1637–1646 (2017).
22. Langsted, A. & Nordestgaard, B. G. Nonfasting versus fasting lipid profile for cardiovascular risk prediction. *Pathology* vol. 51 131–141 (2019).
23. Langsted, A., Freiberg, J. J. & Nordestgaard, B. G. Fasting and nonfasting lipid levels influence of normal food intake on lipids, lipoproteins, apolipoproteins, and cardiovascular risk prediction. *Circulation* **118**, 2047–2056 (2008).
24. Bhatt, D. L. *et al.* Cardiovascular Risk Reduction with Icosapent Ethyl for Hypertriglyceridemia. *N. Engl. J. Med.* NEJMoa1812792 (2018) doi:10.1056/NEJMoa1812792.
25. El Harchaoui, K. *et al.* Value of Low-Density Lipoprotein Particle Number and Size as Predictors of Coronary Artery Disease in Apparently Healthy Men and Women. The EPIC-Norfolk Prospective Population Study. *J. Am. Coll. Cardiol.* **49**, 547–553 (2007).
26. Shah, S. H. & Newgard, C. B. Integrated Metabolomics and Genomics. *Circ. Cardiovasc. Genet.* **8**, 410–419 (2015).

5 Chapter 5 Triglyceride-containing lipoprotein subfractions and risk of coronary heart disease and stroke: a prospective analysis in 11,560 adults

Related publication

Joshi, R., Wannamethee, S. G., Engmann, J., Gaunt, T., Lawlor, D. A., Price, J., Papacosta, O., Shah, T., Tillin, T., Chaturvedi, N., Kivimaki, M., Kuh, D., Kumari, M., Hughes, A. D., Casas, J. P., Humphries, S., Hingorani, A. D., & Schmidt, A. F. (2020). Triglyceride-containing lipoprotein subfractions and risk of coronary heart disease and stroke: A prospective analysis in 11,560 adults. *European journal of preventive cardiology*, https://doi.org/10.1177/2047487319899621

Data sources

UCLEB Consortium studies

- British Regional Heart Study (BRHS)
- Whitehall II study (WHII)
- Southall Brent REvisted study (SABRE)

Abstract

Aims

Elevated low-density lipoprotein cholesterol (LDL-C) is a risk factor for cardiovascular disease however, there is uncertainty about the role of total triglycerides (TG) and the individual triglyceride-containing lipoprotein subfractions. We measured fourteen TG-containing lipoprotein subfractions using nuclear magnetic resonance (NMR) and examined associations with coronary heart disease (CHD) and stroke.

Methods

TG containing subfraction measures were available in 11, 560 participants from the three UK cohorts free of CHD and stroke at baseline. Multivariable logistic regression was used to estimate the association of each subfraction with CHD and stroke expressed as the odds ratio (OR) per standard deviation (SD) increment in the corresponding measure.

Results

The 14 TG-containing subfractions were positively correlated with one another and with total TG, and inversely correlated with HDL-C. Thirteen subfractions were positively associated with CHD (OR in the range 1.12 to 1.22), with the effect estimates for CHD being comparable in subgroup analysis of participants with and without type 2 diabetes, and were attenuated after adjustment for HDL-C and LDL-C. There was no evidence for a clear association of any TG lipoprotein subfraction with stroke.

Conclusions

TG subfractions are associated with increased risk of CHD but not stroke, with attenuation of effects on adjustment for HDL-C and LDL-C.

5.1 Introduction

Elevated low-density lipoprotein cholesterol (LDL-C) is thought to play a central role in atherogenesis[1] and is associated with increased risk of coronary heart disease (CHD) in observational studies, an association which is robust to adjustment for other risk factors[2]. Randomised controlled trials of LDL-C lowering drugs also now provide compelling evidence of its causal role in CHD[3]. On the other hand, the role of triglycerides (TG) in CHD is less clear. Observational data from a large meta-analysis of prospective studies also suggest a higher circulating concentration of TG, and a lower concentration of high-density lipoprotein cholesterol (HDL-C) is associated with coronary heart disease (CHD) but the association of each is attenuated to the null after adjustment for the other leading to uncertainty on the nature of these associations with CHD[2]. However, recently Mendelian randomisation studies have suggested a potential causal association between TG and CHD[4], which has gained some support following publication of the findings of the REDUCE-IT trial[5].

Total circulating TG concentration is made up of contributions from a number of different TG-containing lipoprotein subfractions which, as yet, are not routinely measured in clinical practice. TG are most abundant in chylomicrons transporting fatty acids from the intestine after a meal, and very large-density lipoproteins (VLDL), transporting TG from the liver[6]. In general, the concentration of TG decrease as the lipid content of these lipoproteins are hydrolysed and thus, different lipoprotein subfractions may display different associations with CHD risk[7].

High throughput technology enables the quantification of TG-containing lipoprotein sub fractions, among other lipoprotein subclasses and other metabolites

using serum Nuclear Magnetic Resonance (NMR) spectroscopy[8]. A study based on this platform in a prospective cohort in China found evidence that the association of TG with CHD may depend on the type of TG-containing lipoprotein subfraction[9]. A further study used the same platform in the Finnish population[10]. However, no study to our knowledge has yet investigated the association of TG-containing lipoprotein subfractions with risk of CHD or stroke in other populations. This chapter describes an observational analysis of 14 TG containing subfraction measurements in 11,560 participants to investigate potential subfraction specific associations with CHD and stroke in prospective longitudinal cohort studies from the UCL – Edinburgh - Bristol (UCLEB) consortium[11].

5.2 Methods

5.2.1 Study samples

The design and data collection for the UCL-Edinburgh-Bristol (UCLEB) consortium of longitudinal population studies has been described previously and in detail in Chapter 3[11]. NMR metabolite measurements for the current analysis were available in 11,560 participants enrolled in the British Regional Heart Study[12] (BRHS), including men aged 60 – 79 at metabolite assessment in 1998 – 2000 and 7 years of follow-up; the Whitehall II study[11] (WHII), including UK government workers aged 45 to 69 years at metabolite assessment in 1997 to 1999 and 7 years of follow-up; and The Southall And Brent REvisited Study[13] (SABRE), a tri-ethnic study including British men and women of European (SABRE 1), South Asian (SABRE 2) and African Caribbean descent (SABRE 3), with 20 year follow-up.

5.2.2 Metabolite Quantification

The Nightingale high-throughput NMR metabolomics platform was used to quantify concentrations of total and fourteen TG-containing lipoprotein subfraction metabolomics measures (referred to as TG subfractions in this chapter from here onwards) from plasma samples in either fasting and non-fasting states in all contributing studies. Apolipoprotein A1 (apoA1) and B (apoB) were measured using the same platform. Detailed experimental protocols and application of the metabolomics platform method have been described in previous chapters and reviewed[8,14,15].

5.2.3 Participant characteristics

The following participant information was recorded at time of metabolite measurement: age (years), sex, lifestyle factors; smoking (categorised here as ever/never) and alcohol (ever/never); BMI ($kg/m^{2)}$; and systolic and diastolic blood pressure (mm Hg) and type 2 diabetes mellitus (yes/no). Standard clinical chemistry was used to measure LDL-C and HDL-C (mmol/L) and serum triglycerides (mmol/L). This included, in all three studies LDL-C levels being estimated using Friedwald's equation from total cholesterol and TG (LDL-C = total cholesterol - (Triglyceride / 5) – HDL [16]).

5.2.4 Outcomes

Incident CHD was defined as the first occurrence of fatal or nonfatal, myocardial infarction (MI), or coronary revascularisation. Incident stroke was defined as the first occurrence of fatal or non-fatal ischaemic or haemorrhagic stroke. Methods of disease ascertainment for contributing UCLEB cohorts are described in Chapter 3[11]

5.2.5 Statistical analysis

The study-specific distribution of each TG subfraction were assessed first (figure 5.1). On finding agreement across studies, subfraction measurements were mean centred and standardised to an SD of 1. Spearman's correlation coefficient (r_s) was used to explore associations between the 14 TG subfractions, NMR measured total serum TG, and various participant characteristics. Study-specific logistic regression was used to estimate the odds ratio (OR) and 95% confidence interval

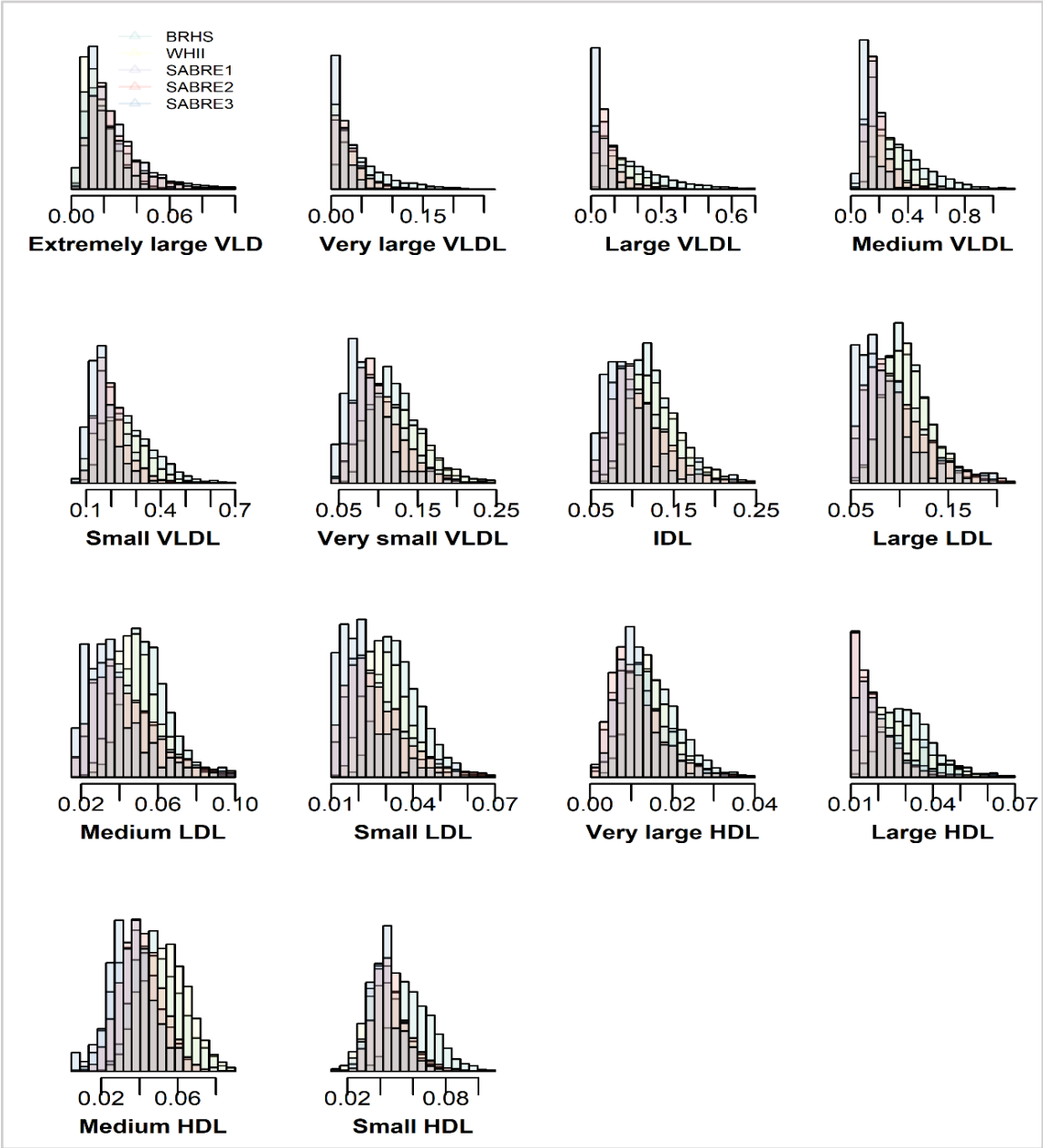
(CI) with CHD and stroke. Estimates were synthesised across cohorts, using the fixed-effect inverse variance weighted estimator. Different multivariable logistic regression models were evaluated to assess how much of the association with CHD and stroke remained after accounting for known CVD risk factors, specifically; age and sex (model 1), model 1 with additional correction for BMI, smoking, systolic blood pressure (SBP) and type 2 diabetes (model 2).

5.2.6 Sensitivity analysis

Subsequent conditioning on LDL-C and HDL-C measurements (model 3) or apoA1 and apoB (model 4) enabled comparisons with previous meta-analysis[2] and to assess the independent association of the 14 subfractions with CHD. Given the potential influence of food intake on the concentration of circulating TG subfractions, the association of fasted and non-fasted TG subfractions with CHD and stroke was compared using data from the SABRE study. A type 2 diabetes stratified analysis to determine if an interaction between diabetes and TG was present in this analysis as has been reported in previous work[17].

Analysis was conducted using R studio version 1.1.423[18] using the following packages; visualise the correlation matrix (Corrplot)[19], conduct meta-analyses (Metafor)[20].

Figure 5.1 Distributions of 14 triglyceride subfractions



N.B. Histograms show study specific distribution (mmol/L) of each NMR triglyceride subfraction.

5.3 Results

The association of the 14 TG subfractions and total NMR measured TG with CHD and stroke was assessed in a sample of 11,560 participants, of which 1,031 experienced CHD and 582 stroke. The mean age was 58 (SD: 9.2) years, 7,634 (66%) were men, mean BMI was 26.4 (SD: 3.8), and average SBP was 131 (SD: 22.7) mmHg, see Table 5.1

Table 5.1 Description of study populations

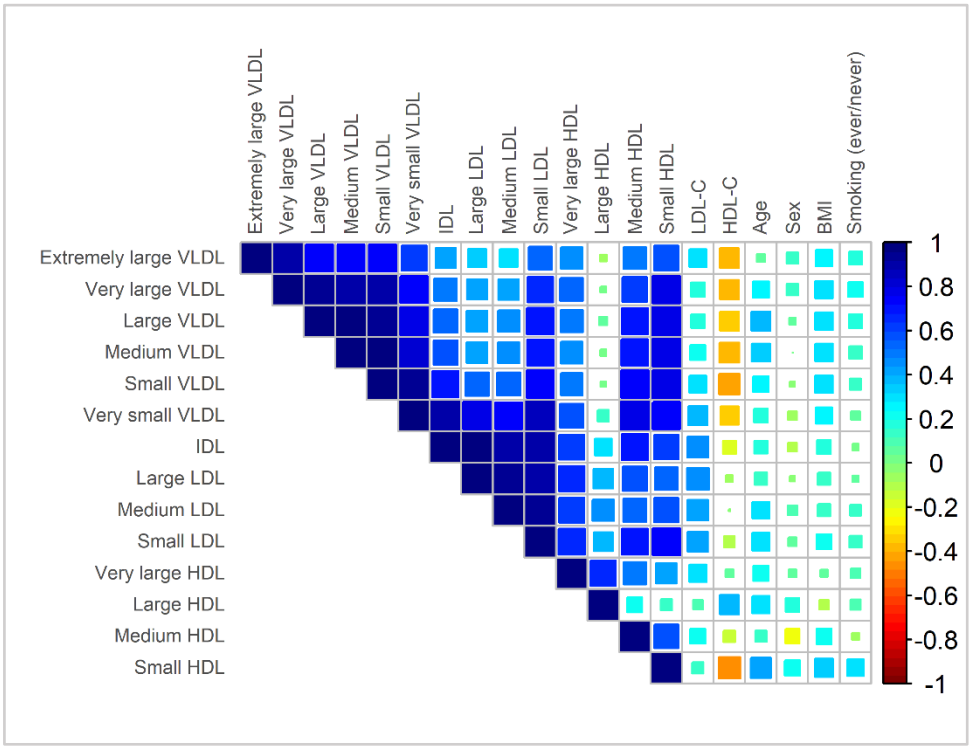
	SABRE2 N= 1526	SABRE1 N= 1561	SABRE3 N= 192	WHII N= 4728	BRHS N= 3553	Total participant sample N = 11,560	Missing (%)
Age, years	51.1 (7.0)	53.2 (7.3)	53.3 (5.7)	55.5 (6.0)	68.7 (5.5)	58.6 (9.2)	0.00
Sex, male (%)	1278 (83.8)	1356 (86.9)	180 (93.8)	1267 (26.8)	3553 (100)	7634 (66.0)	0.00
BMI, kg/m2	26.1 (3.6)	26.1 (3.9)	26.8 (3.7)	26.1 (3.9)	26.9 (3.6)	26.4 (3.8)	0.06
Smoking, ever	349 (22.9)	1097 (70.3)	79 (41.4)	617 (14.6)	2521 (71.1)	4663 (42.2)	0.04
Alcohol, ever	899 (59.3)	1467 (94.0)	176 (92.2)	4244 (91)	3145 (90.4)	9931 (87.0)	0.01
Type 2 diabetes, (%)	318 (20.1)	86 (5.5)	37 (19.3)	251 (5.3)	379 (11.1)	1071 (9.4)	0.01
SBP, mmHg	125.2 (18.2)	123.0 (17.1)	128.4 (18.7)	122.9 (16.1)	149.2 (24.2)	131.3 (22.7)	0.00
DBP, mmHg	80.1 (10.6)	76.8 (10.6)	81.5 (12.4)	77.4 (10.3)	85.3 (11.2)	80.2 (11.3)	0.00
HbA1c, mmol/L	6.2 (1.3)	5.6 (0.6)	6.0 (0.8)	-	5.0 (0.9)	5.4 (1.0)	0.46
Clinical chemistry measured lipids							
Total cholesterol, mmol/L*	5.9 (5.2, 6.6)	6.1 (5.4, 6.8)	5.8 (5.1, 6.6)	5.8 (5.2, 6.6)	6.0 (5.4, 6.7)	5.9 (5.3, 6.6)	0.00
LDL-C, mmol/L*	3.8 (3.2, 4.4)	4.0 (3.4, 4.7	3.8 (3.2, 4.5)	3.8 (3.2, 4.4)	3.9 (3.3, 4.5)	3.8 (3.3, 4.5)	0.00
HDL-C, mmol/L*	1.2 (1.0, 1.4)	1.3 (1.1, 1.5)	1.4 (1.1, 1.6)	1.4 (1.2, 1.7)	1.3 (1.1, 1.5)	1.3 (1.1, 1.6)	0.06
TG, mmol/L*	1.7 (1.2, 2.5)	1.4 (1.0, 2.1)	1.1 (0.8, 1.5)	1.1 (0.8, 1.6)	1.6 (1.2, 2.2)	1.4 (1.0, 2.0)	0.05

NMR measured lipids							
Total cholesterol, mmol/L*	3.8 (3.3, 4.4)	4.1 (3.5, 4.6)	3.8 (3.2, 4.4)	5.1 (4.5, 5.6)	4.6 (4.0, 5.2)	4.6 (3.9, 5.3	0.00
TG, mmol/L*	0.9 (0.7, 1.2)	0.9 (0.7, 1.2)	0.8 (0.6, 0.9)	1.2 (0.9, 1.5)	1.4 (1.1, 1.9)	1.2 (0.9, 1.5)	0.00
Apolipoprotein A1, mmol/L*	1.2 (1.1, 1.3)	1.2 (1.1, 1.3)	1.2 (1.1, 1.3)	1.6 (1.5, 1.7)	1.4 (1.3, 1.5)	1.4 (1.3, 1.6)	0.00
Apolipoprotein B, mmol/L*	0.8 (0.7, 0.9)	0.8 (0.7, 1.0)	0.8 (0.7, 0.9)	1.0 (0.9, 1.1)	1.0 (0.8, 1.1)	0.9 (0.8, 1.1)	0.00
CHD, events/N (%)	317/919 (17.4)	189/1,035 (15.4)	23/109 (17.4)	122/4,200 (2.8)	380/3,173 (10.7)	1,031/10,467 (9.9)	0.09
Stroke, events/N (%)	136/1,191 (10.2)	119/1,212 (8.9)	10/118 (13.2)	68/4,646 (1.4)	241/3,312 (6.8)	582/10,479 (5.3)	0.04

Values are mean ± SD or %.
* Median (interquartile range).
BMI = Body mass index; SBP = systolic blood pressure; DBP = diastolic blood pressure; LDL-C = low-density lipoprotein cholesterol; HDL-C = high-density lipoprotein cholesterol; TG = Triglycerides; CRP= C-reactive protein; Il6 = interleukin-6

The TG subfractions showed positive correlations with one another (figure 5.2), and clinical chemistry measured lipids. For example, TG in extremely large VLDL was positively correlated with 12 subfractions with r_s in the range of 0.33 to 0.91. Various participant characteristics were correlated with TG-subfractions (r_s range -0.01 to 0.32), specifically BMI was positively correlated 14 TG subfractions r_s in the range 0.10 to 0.32. With the exception of TG in large HDL, the remaining TG subfractions (mainly in the VLDL subclass) exhibited a negative correlation with HDL-C (r_s range -0.42 to -0.33).

Figure 5.1 Correlation matrix heat map of 14 triglyceride subfractions and study variables



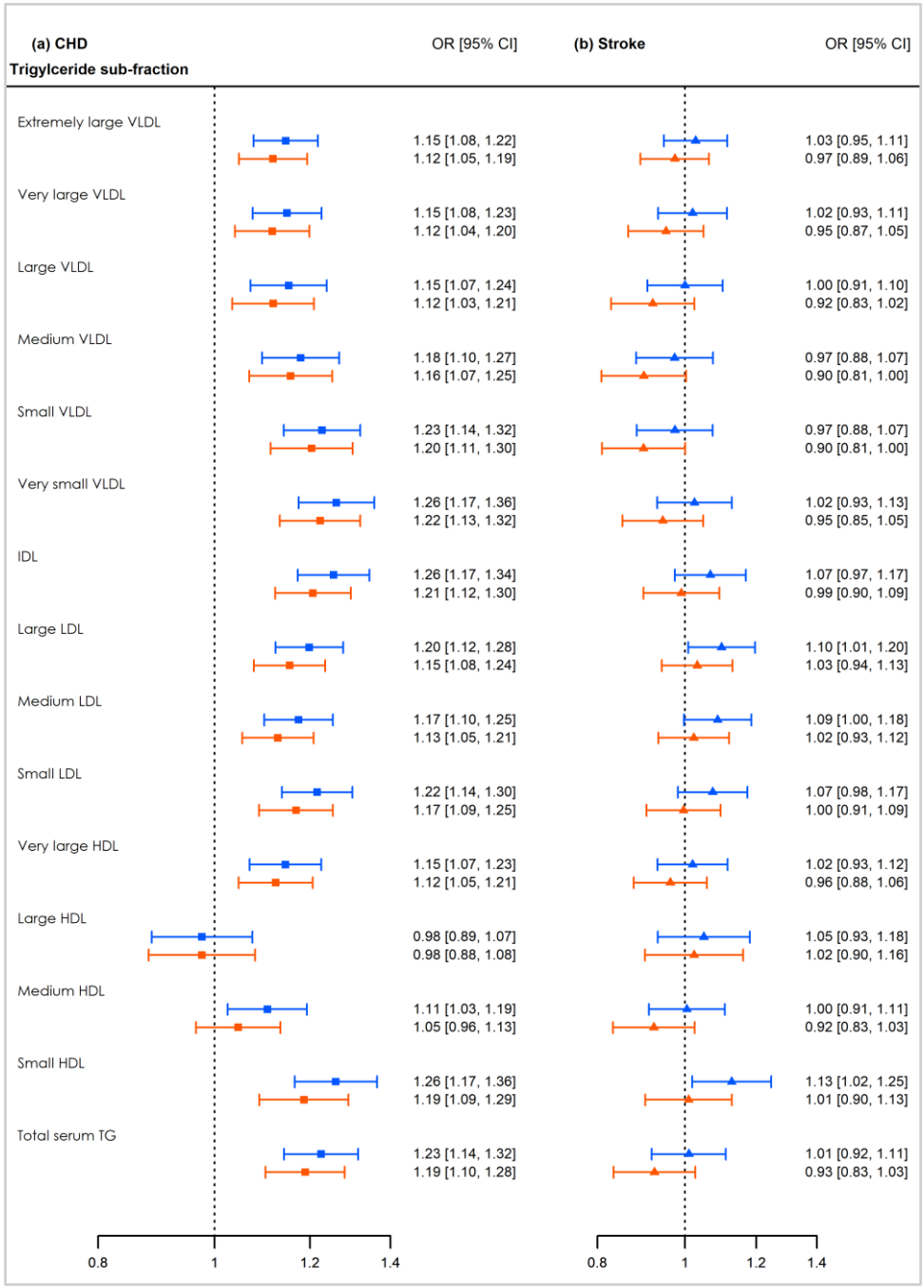
N.B. Heat map matrix are Spearman's correlations

5.3.1 Associations of TG subfractions with coronary heart disease and stroke

In age and sex adjusted analysis total TG and all TG subfractions, with the exception of TG in large HDL, were associated with an increased risk of CHD (figure 5.3**)**. There was low to moderate (I^2 <50%) heterogeneity for 13 subfraction estimates with CHD and stroke (table 5.2) and high heterogeneity (I^2 >70%) in estimates for the large HDL subfraction. After additional adjustments for BMI, systolic blood pressure, smoking and type 2 diabetes (model 2), effect estimates attenuated towards the null, with ORs in the range of 0.98 to 1.22. Triglycerides in the small and medium HDL subclass were associated with an increased risk of CHD, whereas TG in large HDL was inversely associated. Medium and large HDL subfractions did not exclude a neutral-effect. The extent to which LDL-C and HDL-C explained the independent association of TG subfraction with CHD was assessed. Compared to model 2, additionally conditioning on LDL-C and HDL-C attenuated CHD ORs, with none of the 14 estimates showing convincing evidence of an independent effect (Figure 5.4).

Associations of total and 14 TG subfractions with stroke were smaller than those observed for CHD (figure 5.3), with considerable attenuation in both direction and magnitude, after adjustments for SBP, BMI, smoking and type 2 diabetes (model 2) none was statistically significant.

Figure 5.2 Total triglyceride and 14 triglyceride subfraction associations with CHD and stroke



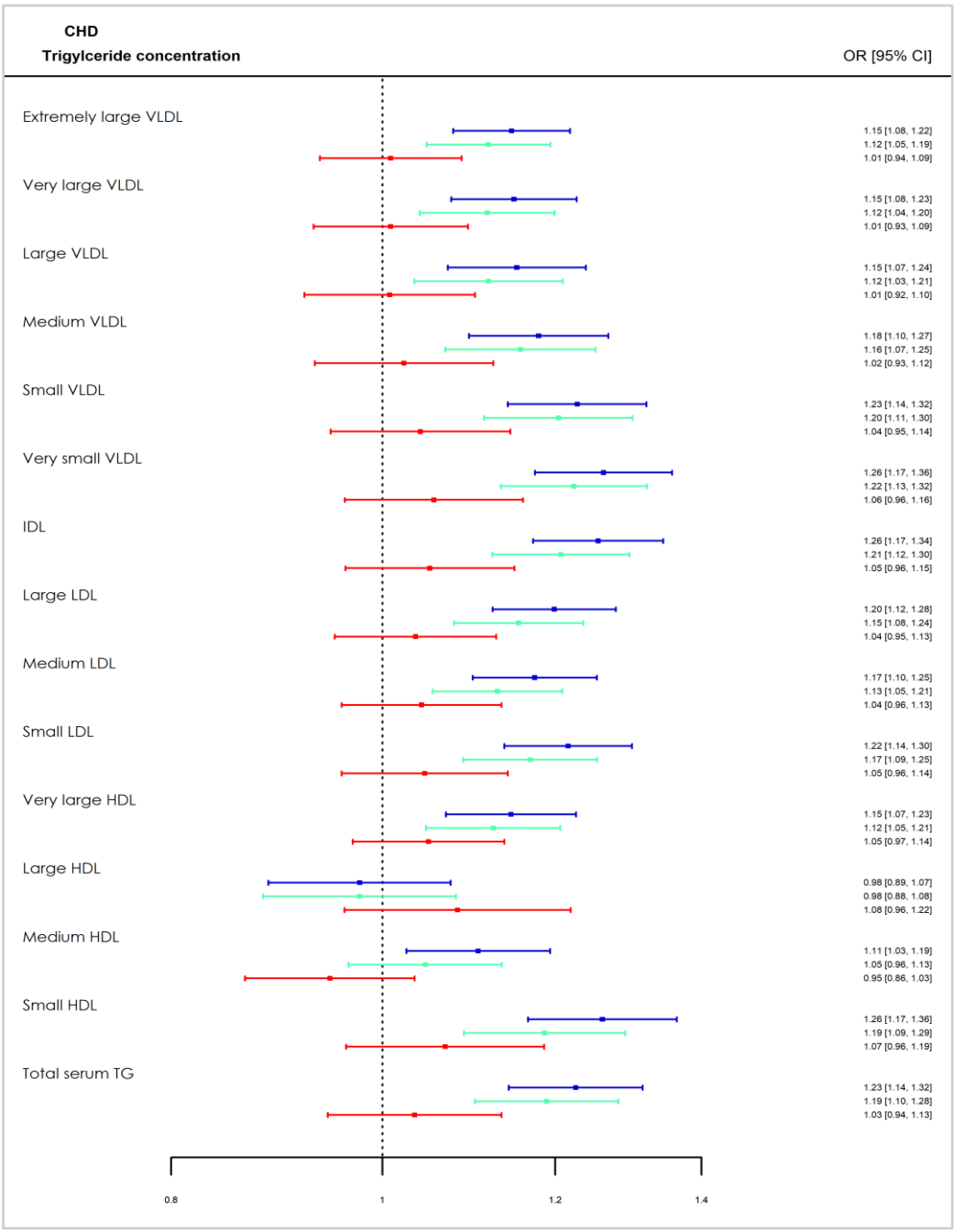
N.B. Effect estimates are presented as odds ratios (OR) with 95% confidence intervals (CI) per 1 standard deviation increase in the analyte for CHD (a) and stroke (b). Models are adjusted for; age and sex (model 1 denoted by blue bar), model 1 with additional correction for smoking status, BMI, systolic blood pressure and type 2 diabetes (model 2 denoted by orange bar)

Table 5.2 I2 statistics for CHD and stroke

	CHD I2 [95% CI]	Stroke I2 [95% CI]
Extremely large VLDL	0.00 [0.00, 86.90]	27.42 [0.00, 97.20]
Very large VLDL	0.00 [0.00, 95.70]	0.00 [0.00, 93.81]
Large VLDL	6.66[0.00, 99.24]	0.00 [0.00, 87.79]
Medium VLDL	6.49 [0.00, 98.82]	0.00 [0.00, 89.13]
Small VLDL	0.00 [0.00, 98.07]	0.00 [0.00, 90.01]
Very small VLDL	14.71 [0.00, 96.29]	39.88 [0.00, 95.21]
IDL	24.11 [0.00, 96.01]	25.98 [0.00, 93.71]
Large LDL	20.75 [0.00, 95.25]	42.49 [0.00, 95.21]
Medium LDL	23.79 0.00, 93.93]	5.35 [0.00, 93.32]
Small LDL	20.62 [1.00, 94.51]	40.18 [0.00, 93.15]
Very large HDL	31.05 [0.92, 99.00]	0.00 [0.00, 94.44]
Large HDL	70.53 [24.79, 99.37]	33.87 [0.00, 99.20]
Medium HDL	0.00 [0.78, 94.99]	26.91 [0.00, 93.85]
Small HDL	38.94 [0.94, 95.86]	14.13 [0.00, 90.83]
Total TG	15.58 [0.00, 97.90]	9.51 [0.00, 93.64]

Estimates are heterogeneity statistics I2 (%) and 95% confidence intervals (CI)
I2 adjusted for; age, sex, body mass index, smoking, systolic blood pressure, type 2 diabetes.
CHD = Coronary heart disease; VLDL = Very-low density lipoprotein; IDL = Intermediate-density lipoprotein; LDL = Low-density lipoprotein; HDL = High-density lipoprotein.

Figure 5.4 Total triglyceride and 14 triglyceride subfraction associations with CHD



N.B. Effect estimates are presented as odds ratios (OR) with 95% confidence intervals (CI) per 1 standard deviation increase in the analyte for CHD. Models are adjusted for; age and sex (model 1 denoted by blue bar), model 1 with additional correction for smoking status, BMI, systolic blood pressure and type 2 diabetes (model 2 denoted by green bar). Model 2 with additional correction for HDL-C and LDL-C (model 3 denoted by red bar)

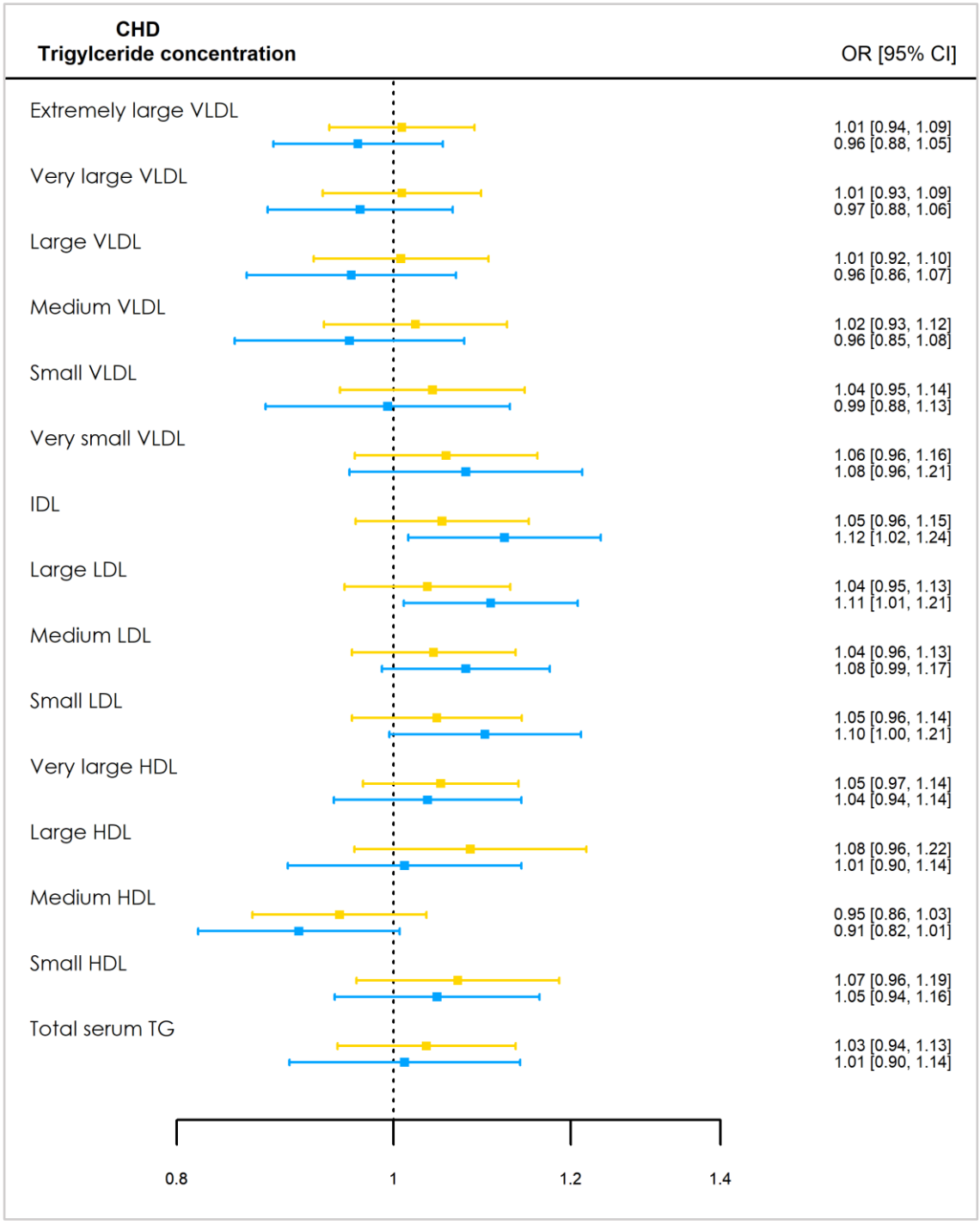
5.3.2 Sensitivity analysis

In a sensitivity analysis, the influence of replacing LDL-C and HDL-C by apoB and apoA1, which are considered more precise measurement were explored. However, adjustment for ApoA1 and apoB, rather than for HDL-C and LDL-C respectively, did not meaningfully change the results (Figure 5.5), with one possible exception of TG in the VLDL subclass for which the direction of association changed. For example, accounting for LDL-C and HDL-C, ORs were in the range of 1.01 to 1.04, whereas adjustment of apoA1 and apoB yielded negative associations in the range of 0.96 to 0.99.

Effect estimates of 14 TG subfractions with CHD were comparable in direction and pattern of association using TG measures quantified in the fasting and non-fasting state (Appendix figure 5.1).

The effect estimates of total serum TG with CHD measured using NMR methods vs clinical chemistry methods were comparable in direction and magnitude of effect (NMR measured total TG: OR 1.19, 95% CI 1.10 to 1.28, vs clinical chemistry measured total TG: OR 1.11, 95% CI 1.07 to 1.21 (Appendix table 5.1)

Figure 5.5 The HDL-C and LDL-C indpendent association of total triglyceride and 14 triglyceride subfraction with CHD



N.B. Effect estimates are presented as odds ratios (OR) with 95% confidence intervals (CI) per 1 standard deviation increase in the analyte for CHD. Models are adjusted for; age, sex, smoking status, BMI, systolic blood pressure, type 2 diabetes and; HDL-C and LDL-C (denoted by yellow bar), ApoA1 and ApoB (denoted by blue bar).

5.3.3 CHD associations in participants with and without type 2 diabetes.

The differences in effect of the 14 TG sub0fractions with CHD were explored in a T2DM stratified analysis (1,071 T2DM participants vs 10,347 non-T2DM participants). There was an absence of evidence of an interaction between TG subfractions and diabetes with CHD risk (table 5.3). TG concentration in 13 subfractions were positively associated with CHD in participants with and without type 2 diabetes, although effect estimates for the non-type 2 diabetes subgroup had wide 95% confidence intervals that frequently included unity likely due to the comparatively small sample size.

Table 5.3 Effect estimates and upper and lower 95% confidence interval in subgroup participant population of type 2 diabetes and risk of CHD

	Type 2 diabetes (N =1,071) CHD OR [95% CI]	No type 2 diabetes (N=10,347) CHD OR [95% CI]	P value for interaction
Extremely large VLDL	1.05 [0.92, 1.20]	1.14 [1.06, 1.22]	0.286
Very large VLDL	1.04 [0.91, 1.18]	1.13 [1.04, 1.22]	0.262
Large VLDL	1.02 [0.88, 1.18]	1.13 [1.04, 1.23]	0.201
Medium VLDL	1.05 [0.91, 1.22]	1.17 [1.08, 1.28]	0.168
Small VLDL	1.11 [0.94, 1.30]	1.22 [1.13, 1.33]	0.261
Very small VLDL	1.15 [0.96, 1.36]	1.25 [1.15, 1.36]	0.364
IDL	1.22 [1.04, 1.43]	1.21 [1.12, 1.31]	0.903
Large LDL	1.19 [1.02, 1.39]	1.15 [1.06, 1.24]	0.667
Medium LDL	1.11 [0.95, 1.29]	1.13 [1.05, 1.22]	0.812
Small LDL	1.16 [1.00, 1.36]	1.17 [1.08, 1.26]	0.954
Very large HDL	1.08 [0.92, 1.25]	1.11 [1.03, 1.20]	0.685
Large HDL	0.91 [0.73, 1.14]	0.94 [0.85, 1.04]	0.814
Medium HDL	0.94 [0.78, 1.13]	1.05 [0.96, 1.14]	0.264
Small HDL	1.12 [0.94, 1.33]	1.21 [1.11, 1.33]	0.380

Estimates are odds ratios (OR) and upper and lower bands (UB, LB)).
Odds ratios are adjusted for; age, sex, body mass index, smoking status, alcohol, systolic blood pressure.
CHD = Coronary heart disease; VLDL = Very-low density lipoprotein; IDL = Intermediate-density lipoprotein; LDL = Low-density lipoprotein; HDL = High-density lipoprotein.

5.4 Discussion

Total TG and most TG subfractions (with the exception of TG in the HDL subclass) were associated with an increased risk of CHD. These associations persisted after accounting for differences in age, sex, systolic blood pressure, BMI, smoking and type 2 diabetes, ranging from an CHD OR of 1.12 to 1.22. Importantly, TG within the VLDL subclass had the strongest association with CHD (OR in the range of 1.12 to 1.22), which is consistent with the view that VLDL lipoprotein particles are particularly atherogenic[21,22]. Accounting for HDL-C and LDL-C, attenuated the CHD associations towards the null, suggesting TG associations are not independent of LDL-C and HDL-C. While such analyses may suggest an absence of a direct pathway, i.e. independent of LDL-C and HDL-C, an alternative explanation for this attenuation may be found in the different variability of TG measurements compared with HDL-C[23]. Biological variability in TG measurements, has previously been reported with a median variation of 23.5% and is a consideration when evaluating the role of TG in CVD risk[22]. We did not observe any interaction between T2DM status and the subfraction associations with CHD.

Consistent with a previous study[9], concentrations of TG within large and medium HDL particles did not show an association with risk of CHD. Measures in the HDL subfraction likely represents the role of HDL particles that mediates reverse cholesterol transport, a different process to that indexed by the other subfractions. In relation to reverse cholesterol transport, variants in the CETP gene and treatment with a potent, target-specific CETP inhibitor for a sufficient duration, both result in reduced triglycerides, LDL-C and apoB, elevated HDL-C and apoA1, and reduced CHD risk[24].

The apparent inconsistency of associations of total TG and CHD[2] and uncertainty on causal relationships may be due in part to the complexity of TG absorption and metabolism, and a potentially heterogeneous role of different TG containing lipoproteins in atherogenesis. Conventional measures of total serum TG levels do not take into account the differing lipoprotein compositions with the same detailed precision as NMR technology and thus, standard clinical chemistry measurement processes may not sufficiently delineate the relationship of individual lipoproteins with CHD. To evaluate this, the association of total serum TG measured using clinical chemistry methods and NMR methods with CHD was assessed to find comparable effects despite the absolute difference in total serum TG concentration between the two methods.

TG are most abundant in chylomicrons transporting fatty acids from the intestine, and very large-density lipoproteins (VLDL), transporting TG from the liver[6]. In general, the TG-content of lipoprotein subfractions decreases as the lipid pools of these lipoproteins are hydrolysed and thus, different lipoproteins may be associated with differential CHD risk[7]. Increases in total TG may reflect an increase in the TG content of lipoproteins, or an increase in the total number of triglyceride rich (predominantly VLDL) particles. In addition, an emerging view is that the remnant cholesterol content (that can be proxied by total cholesterol – HDL-C – LDL-C) in triglyceride-rich lipoprotein (TRL) as well as triglycerides, have pro-atherogenic actions[6,25,26]. A suggested mechanism is that triglycerides can penetrate the arterial intima and trigger inflammation and are subsequently very rapidly metabolised whereas, the cholesterol remnants are not, promoting foam cell formation, atherosclerotic plaques, and ultimately CHD[27].

It has been suggested that apoB (which increases the total number of atherogenic particles) is more robustly associated with CHD than other measures[28]. However, the observed associations were comparable in direction and magnitude for CHD, regardless of whether adjustments were for LDL-C and HDL-C or apoA1 or apoB, with the possible exception of TG in the VLDL subclass.

In contrast to associations with CHD, there was no convincing evidence of an association between TG subfractions and risk of stroke. The magnitude of unadjusted and adjusted associations for stroke were all weaker than for CHD and showed greater variation in direction between the TG subfractions. However, it is possible that large studies with more stroke events might find evidence of a modest association of certain subfractions with stroke risk.

5.4.1 Research in context

The results presented here are compared to existing knowledge in this area. A study using Finnish participant data to examine the association of circulating metabolites and CVD reported age and sex adjusted associations of triglycerides and incident CVD[10] (800 cases/7,256 participants; hazard ratio (HR): 1.25, 95% CI: 1.13-1.35). Similar to the findings presented here, Wurtz et al report TG associations with CHD attenuate towards the null when performing additional adjustment for HDL-C. A recent study in participants from a Chinese cohort (912 cases/4,662 participants) demonstrated consistent associations of TG subfractions and incident myocardial infarction (MI) after controlling for age, sex, fasting hours, region, smoking status, and educational attainment (OR per 1 SD increase in the metabolite

in the range 1.03-1.31)[9], however, effect estimates attenuated towards the null with additional adjustment for SBP and BMI. The present study, although not controlling for fasting hours or educational attainment identified comparable effect estimates which remained robust to adjustments for SBP and BMI.

5.4.2 Strengths and limitations

This study has several strengths. There was an no observed large variation in the distribution of the 14 TG subfractions when combining study specific distributions, indicating homogeneity in the spread of the 14 TG subfractions and no discernible differences in the contributing populations from which the samples were derived. This study includes a tri-ethnic cohort and reports a similar distribution of each analyte irrespective of ethnic group. The findings here were consistent with those reported by Holmes *et al*[9] and Wurtz *et al*[10] but included a larger sample size for CHD (CHD 1,031cases/10,467 controls, Stroke 582 cases/10,479 controls) compared to Holmes *et al (*MI 912 cases/1,466 controls and ischaemic stroke 1,146 cases /1,466 controls) and Wurtz *et al (*CVD 800 cases/7256 controls), thereby increasing precision in the effect estimates presented here. This study extends the investigation of the relationship between TG-containing lipoprotein subfractions beyond ethnically distinct Finnish and Chinese populations, to demonstrate consistent findings using data from multiple cohorts across multiple ethnicities consisting of European, South-Asian and African-Caribbean descent, and across differing age groups.

Total TG and TG-containing lipoprotein subfraction concentration in plasma samples are affected by food intake[6]. Before 2009, in accordance with guidelines and

statements, lipid profiles were measured using fasted samples (defined as blood samples drawn after an 8-hour fast) mainly due to the increase seen in TG during a fat tolerance test. More recent evidence from 2018 suggests that consumption of food is usually evenly distributed throughout the day and thus most people find themselves in the non-fasting state for the majority of a 24 hour period, perhaps with the exception of morning hours, and therefore lipid profiles change minimally in response to normal food intake in individuals in the general population and may be clinically unimportant[29]. For example, evidence from four large prospective studies found maximal mean changes were +0.3 mmol/L for triglycerides, -0.2 mmol/L for total cholesterol, -0.2 mmol/L for LDL-C, and -0.1 mmol/L for HDL-C[30]. This view is further supported by a recent meta-analysis that found fasting and non-fasting TG levels were equally as good at predicting increased risk of CHD[31,32]. Nonetheless, fasting status was assessed in the present study. In a stratified analysis, there were similar associations of TG containing lipoprotein subfractions with CHD and stroke among fasted and non-fasted subjects from the SABRE study. Further supporting and contributing to the possible shift away from using fasting measures for clinical lipid profiling.

In this study stroke was defined as a composite of ischaemic and haemorrhagic stroke. Lower TG concentrations have been shown to be associated with decreased risk of haemorrhagic stroke, whereas higher TG associated with increased risk of ischaemic stroke[33]. We were unable to differentiate types of stroke in this study (in which typically there is a 4:1 ratio of ischaemic to haemorrhagic stroke events), as such the null effect estimates presented here may not reflect the true association of TG and ischaemic stroke.

In addition, we were unable to account for lipid lowering medication or socio-economic position as has been conducted in previous work[10]. Lipid lowering medication may modify the effect of TG subfraction associations with CHD, therefore it is possible that observed associations may be explained by residual confounding due to factors not included in this study.

5.5 Conclusions

In conclusion, we demonstrate risk-increasing associations of 13 triglyceride subfractions with CHD, with varying effect estimates between subfractions, the strongest being for the triglycerides in the VLDL subfraction. By contrast, we did not observe similar associations for stroke. Further studies, for example using Mendelian randomisation, could assess which (if any) of the associations is causal, and if any causal associations might be modifiable by new or existing drugs.

5.6 References

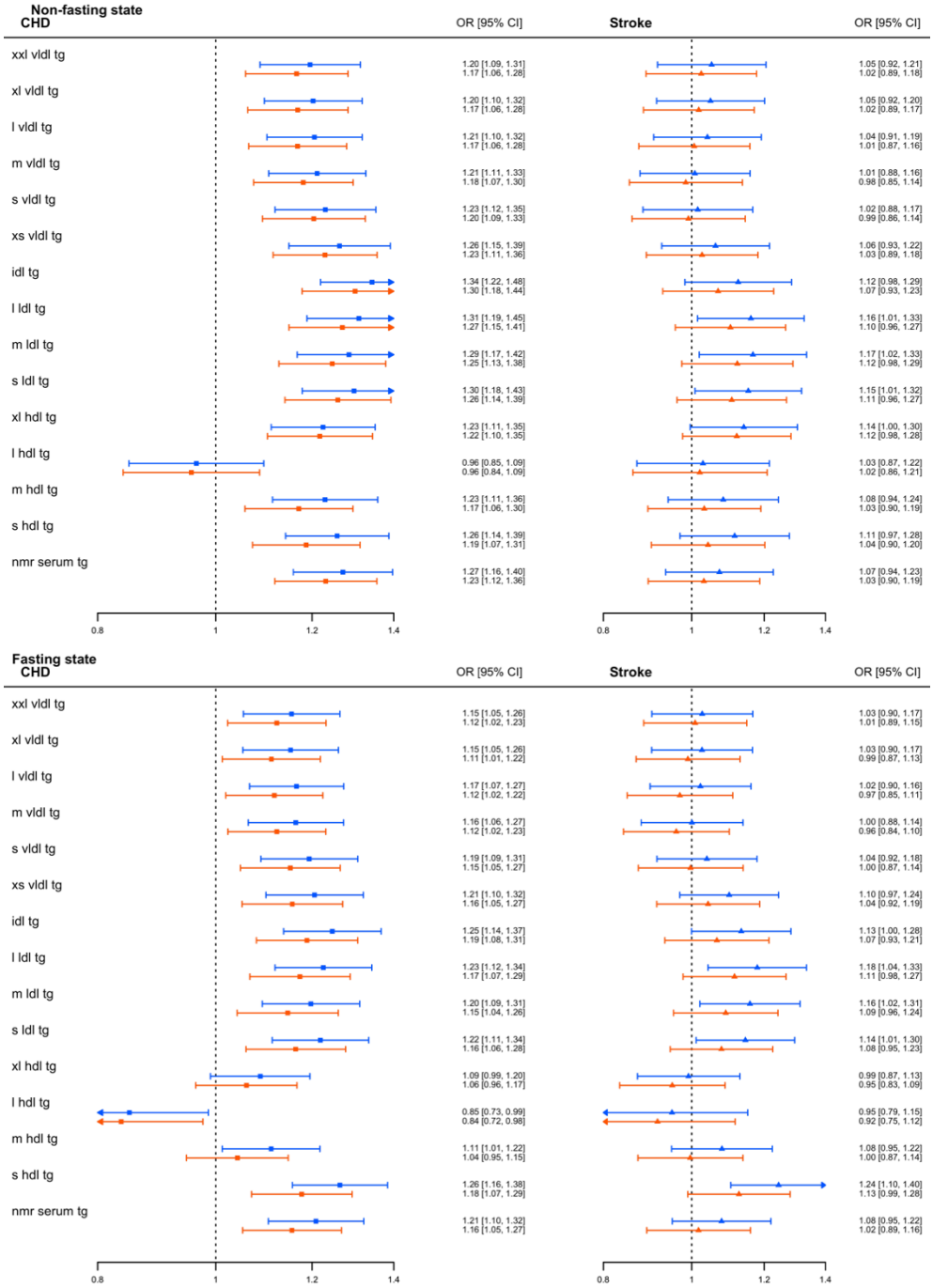
1. Steinberg, D. Thematic review series: The Pathogenesis of Atherosclerosis. An interpretive history of the cholesterol controversy, part III: mechanistically defining the role of hyperlipidemia. *J. Lipid Res.* **46**, 2037–2051 (2005).
2. The Emerging Risk Factors Collaboration*, T. E. R. F. Major Lipids, Apolipoproteins, and Risk of Vascular Disease. *JAMA* **302**, 1993 (2009).
3. Cholesterol Treatment Trialists' (CTT) Collaboration *et al.* Efficacy and safety of LDL-lowering therapy among men and women: meta-analysis of individual data from 174 000 participants in 27 randomised trials. *Lancet* **385**, 1397–1405 (2015).
4. White, J. *et al.* Association of Lipid Fractions With Risks for Coronary Artery Disease and Diabetes. *JAMA Cardiol.* **1**, 692 (2016).
5. Bhatt, D. L. *et al.* Cardiovascular Risk Reduction with Icosapent Ethyl for Hypertriglyceridemia. *N. Engl. J. Med.* NEJMoa1812792 (2018) doi:10.1056/NEJMoa1812792.
6. Nordestgaard, B. G. Triglyceride-Rich Lipoproteins and Atherosclerotic Cardiovascular Disease: New Insights From Epidemiology, Genetics, and Biology. *Circ. Res.* **118**, 547–63 (2016).
7. Dron, J. S. & Hegele, R. A. Genetics of Triglycerides and the Risk of Atherosclerosis. *Current Atherosclerosis Reports* (2017) doi:10.1007/s11883-017-0667-9.
8. Soininen, P., Kangas, A. J., Würtz, P., Suna, T. & Ala-Korpela, M. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. *Circ. Cardiovasc. Genet.* **8**, 192–206 (2015).
9. Holmes, M. V. *et al.* Lipids, Lipoproteins, and Metabolites and Risk of Myocardial Infarction and Stroke. *J. Am. Coll. Cardiol.* (2018) doi:10.1016/j.jacc.2017.12.006.
10. Würtz, P. *et al.* Metabolite Profiling and Cardiovascular Event Risk. *Circulation* **131**, 774–785 (2015).
11. Shah, T. *et al.* Population Genomics of Cardiometabolic Traits: Design of the University College London-London School of Hygiene and Tropical Medicine-Edinburgh-Bristol (UCLEB) Consortium. *PLoS One* **8**, e71345 (2013).
12. Shaper, A. G. *et al.* British Regional Heart Study: cardiovascular risk factors in middle-aged men in 24 towns. *BMJ* **283**, 179–186 (1981).
13. Tillin, T. *et al.* Southall And Brent REvisited: Cohort profile of SABRE, a UK population-based comparison of cardiovascular disease and diabetes in people of European, Indian Asian and African Caribbean origins. *Int. J. Epidemiol.* **41**, 33–42 (2012).
14. Ala-Korpela, M., Kangas, A. J. & Soininen, P. Quantitative high-throughput

metabolomics: a new era in epidemiology and genetics. *Genome Med.* **4**, 36 (2012).
15. Soininen, P. *et al.* High-throughput serum NMR metabonomics for cost-effective holistic studies on systemic metabolism. *Analyst* **134**, 1781 (2009).
16. Friedewald, W. T., Levy, R. I. & Fredrickson, D. S. Estimation of the Concentration of Low-Density Lipoprotein Cholesterol in Plasma, Without Use of the Preparative Ultracentrifuge. *Clin. Chem.* **18**, (1972).
17. Lee, J. S. *et al.* Triglyceride and HDL-C Dyslipidemia and Risks of Coronary Heart Disease and Ischemic Stroke by Glycemic Dysregulation Status: The Strong Heart Study. *Diabetes Care* **40**, 529–537 (2017).
18. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. *Vienna, Austria* (2018) doi:10.1108/eb003648.
19. Wei, T. & Simko, V. R package 'corrplot': visualization of a correlation matrix (version 0.84). *R Found. Stat. Comput., Vienna. https://github. com/taiyun/corrplot* (2017).
20. Viechtbauer, W. Metafor Package. *J. Stat. Softw.* (2010) doi:10.1002/wics.10.
21. Sacks, F. M. *et al.* VLDL, Apolipoproteins B, CIII, and E, and Risk of Recurrent Coronary Events in the Cholesterol and Recurrent Events (CARE) Trial. *Circulation* **102**, 1886–1892 (2000).
22. Miller, M. *et al.* Triglycerides and Cardiovascular Disease. *Circulation* **123**, 2292–2333 (2011).
23. März, W. *et al.* HDL cholesterol: reappraisal of its clinical relevance. *Clin. Res. Cardiol.* **106**, 663–675 (2017).
24. Ference, B. A. *et al.* Association of genetic variants related to CETP inhibitors and statins with lipoprotein levels and cardiovascular risk. *JAMA - J. Am. Med. Assoc.* **318**, 947–956 (2017).
25. Varbo, A. & Nordestgaard, B. G. Remnant cholesterol and ischemic heart disease. *Current Opinion in Lipidology* vol. 25 266–273 (2014).
26. Balling, M. *et al.* A third of nonfasting plasma cholesterol is in remnant lipoproteins: Lipoprotein subclass profiling in 9293 individuals. *Atherosclerosis* **286**, 97–104 (2019).
27. Nordestgaard, B. G. & Varbo, A. Triglycerides and cardiovascular disease. *Lancet* **384**, 626–635 (2014).
28. Benn, M., Nordestgaard, B. G., Jensen, G. B. & Tybjærg-Hansen, A. Improving Prediction of Ischemic Cardiovascular Disease in the General Population Using Apolipoprotein B. *Arterioscler. Thromb. Vasc. Biol.* **27**, 661–670 (2007).
29. Langsted, A. & Nordestgaard, B. G. Nonfasting versus fasting lipid profile for cardiovascular risk prediction. *Pathology* vol. 51 131–141 (2019).
30. Langsted, A., Freiberg, J. J. & Nordestgaard, B. G. Fasting and nonfasting lipid levels influence of normal food intake on lipids, lipoproteins, apolipoproteins, and cardiovascular risk prediction. *Circulation* **118**, 2047–2056 (2008).

31. Nordestgaard, B. G. *et al.* Fasting is not routinely required for determination of a lipid profile: clinical and laboratory implications including flagging at desirable concentration cut-points—a joint consensus statement from the European Atherosclerosis Society and European Federation of Clinical Chemistry and Laboratory Medicine. *Eur. Heart J.* **37**, 1944–1958 (2016).
32. Sarwar, N. *et al.* Clinical perspective. *Circulation* **115**, 450–458 (2007).
33. Bonaventure, A. *et al.* Triglycerides and risk of hemorrhagic stroke vs. ischemic vascular events: The Three-City Study. *Atherosclerosis* **210**, 243–248 (2010).

5.7 Chapter 5 appendices

Supplmentary figure 5.1 2 Total and 14 triglyceride sub-fraction measured in the fasting state and non-fasting state associations with CHD and stroke



N.B. TG sub-fraction measures in; top: non-fasting state, bottom: fasting state. Effect estimates are presented as odds ratios (OR) with 95% confidence intervals (CI) per 1 standard deviation increase in the analyte for CHD and stroke. Models are adjusted for; age and sex (model 1 denoted by blue bar), model 1 with additional correction for smoking status, BMI, systolic blood pressure and type 2 diabetes (model 2 denoted by orange bar

Appendix table 5.1 Effect estimates for NMR measured and clinical chemistry measured lipids with CHD

Lipid measure method	CHD OR	LB	UB
NMR measured TG	1.19	1.10	1.28
Clinical chemistry measured TG	1.14	1.07	1.21
NMR measured total cholesterol	1.10	1.02	1.19
Clinical chemistry measured total cholesterol	1.22	1.14	1.31
Estimates are odds ratios and 95% confidence intervals (CI) **OR adjusted for; age, sex, body mass index, smoking, systolic blood pressure, type 2 diabetes.** **TG = triglycerides** **CHD = coronary heart disease** **OR = odds ratio** **LB = lower band** **UB = upper band**			

6 Evaluation of triglyceride and cholesterol content in fourteen lipoprotein subfractions with coronary heart disease: An observational and genetic analysis

Related publication

Joshi, Roshni, et al, In preparation "Evaluating the causal relevance of triglyceride and cholesterol content in 14 NMR measured lipoprotein subfractions with risk of coronary heart disease: An observational and genetic analysis", 2021

Data sources

UCLEB Consortium studies

- British Regional Heart Study (BRHS)
- Whitehall II study (WHII)
- Southall Brent REvisted study (SABRE)

Genetic instruments

- Kettunen and de novo GWAS of UCLEB measurements
- CARDIoGRAMplusC4D

Abstract

Background

Triglycerides (TG) and cholesterol are carried in varying quantities in chylomicrons, very-low, intermediate, low, and high-density lipoproteins (VLDL, LDL, IDL and HDL) subfractions. Questions have arisen about the potential atherogenicity of the cholesterol content in lipoprotein subfractions other than the LDL subfraction. It remains unclear which lipid component (TG or cholesterol) in which lipoprotein subfraction predominates the association with CHD. This study estimates the effects of TG and cholesterol in fourteen lipoprotein subfractions measured using NMR spectroscopy on CHD.

Methods

The TG and cholesterol content of fourteen lipoprotein subfractions were measured using the Nightingale NMR platform in cohort studies contributing to the UCL-Edinburgh-Bristol (UCLEB) consortium. Logistic regression was used to evaluate TG in fourteen lipoprotein subfractions accounting for; age, sex, BMI, smoking, systolic blood pressure (SBP) and type 2 diabetes, and additionally for the cholesterol in each subfraction on CHD. The same approach was taken to evaluate the association of cholesterol in each lipoprotein subfraction with CHD. Genetic instruments for TG and cholesterol content in each lipoprotein subfraction were identified from a de novo meta-analysis GWAS of UCLEB measurements and harmonised to CHD data from CARDIoGRAMplusC4D. Univariable and multivariable MR was used to estimate the total and direct effect of TG and cholesterol in each subfraction on CHD, respectively. A Rucker selection framework

was applied to decide between the inverse variance weighted or the pleiotropy robust Egger method.

Results

In age, sex, BMI, smoking, SBP and type 2 diabetes adjusted analysis, there was a positive association of TG content in 14 lipoprotein subfractions with CHD. With additional adjustment for the cholesterol content in each lipoprotein subfraction, TG content in 10 lipoprotein subfractions retained a positive association with CHD (OR in the range 1.08 to 1.25). Using the same adjustment approach, cholesterol content in 13 lipoprotein subfractions had a mixed positive and inverse association with CHD. With additional adjustment for the TG content in each lipoprotein subfraction, cholesterol in four VLDL lipoprotein subfractions retained a robust positive association with CHD (OR in the range 1.26 to 1.63). Cholesterol in three HDL subfractions retained inverse associations with CHD (OR in the range 0.65 to 0.92)

There was a total causal association of TG content in five lipoprotein subfractions, and cholesterol content in 10 lipoprotein subfractions with CHD. In MVMR analysis there was a direct association of TG in four lipoprotein subfractions and cholesterol in 10 lipoprotein subfractions in CHD. Cholesterol content in TRL displayed the largest effects (MVMR OR in the range 2.73 to 14.31), an association that was not observed for TG in TRL.

Broadly speaking, in MVMR analysis the inverse, null estimates for the TG content, and positive effect estimates for the cholesterol content in the VLDL lipoprotein subfractions yielded point estimates with imprecise confidence intervals. It is likely this imprecision is representative of multi-collinearity of TG and

cholesterol content in each of the VLDL lipoprotein subfractions included in the same analysis model, making it difficult to deduce an independent effect of each lipid trait, rather than assume a true absence of effect. This may be especially true for the TG content in the VLDL subfractions.

Conclusion

This study provides strong evidence that the cholesterol content of VLDL and IDL subfractions is the predominant trait that is causally related with CHD, independent of the TG content in these lipoprotein subfractions.

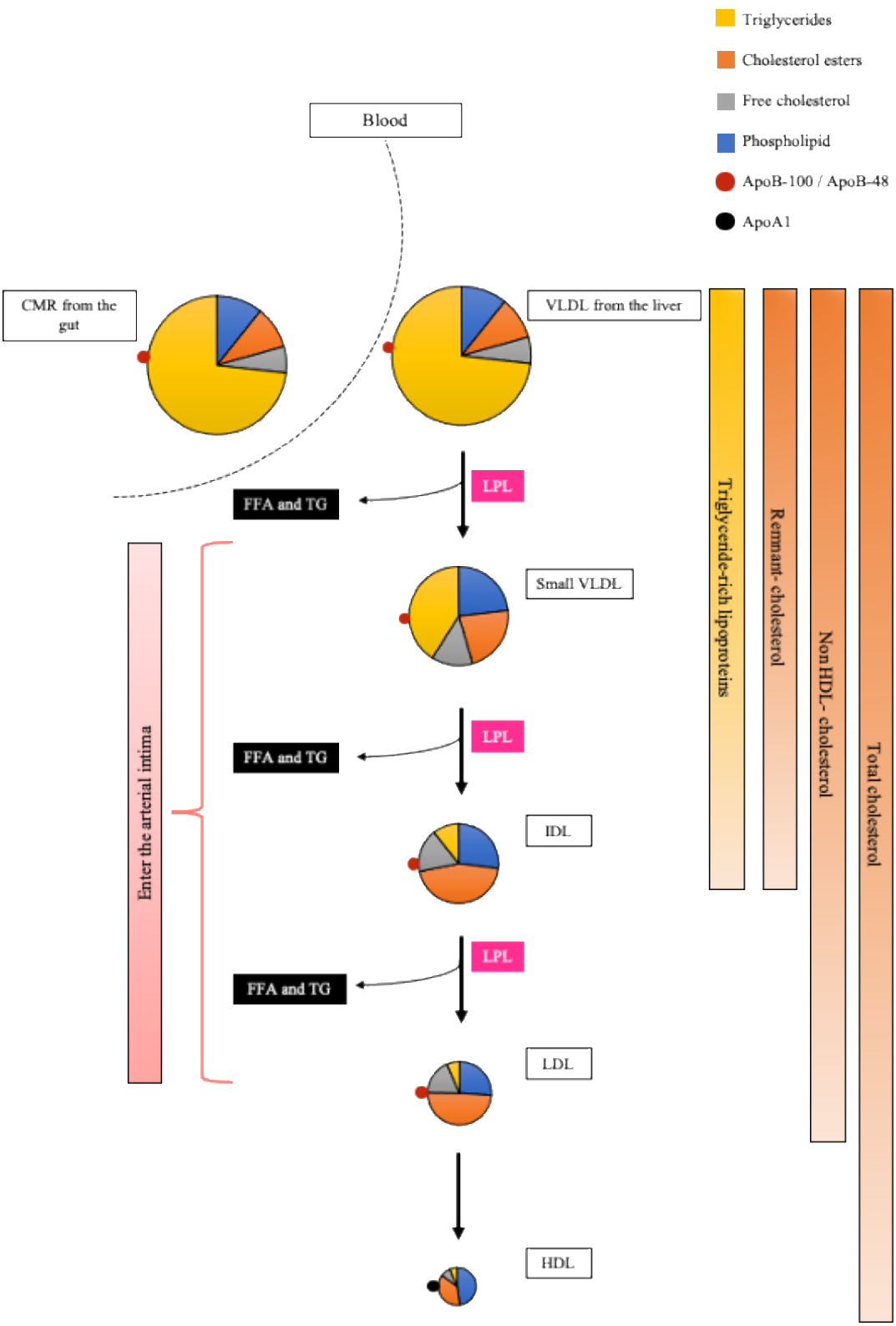
6.1 Introduction

The major blood lipid components, free cholesterol, cholesteryl-esters (collectively cholesterol), and triglycerides (TG) are transported in the core of membrane bound lipoprotein particles which can be separated by size and density[1]. Each of these lipoproteins differ in their cholesterol and TG content. Large lipoprotein particles, which encompass chylomicrons (CMR) derived from dietary fat, as well as very-low density lipoproteins synthesised in the liver, are TG-rich[2]. These particles express a single apolipoprotein B (Apo-B) on the surface (Apo-B 48 for chylomicrons and Apo-B 100 otherwise) and are progressively depleted of TG following hydrolysis by lipoprotein lipase (LPL), becoming smaller, denser and proportionately richer in cholesterol[2,3]. The ApoB expressing lipoproteins, which are involved in the process of transporting cholesterol to peripheral tissues, are generally classified as (VLDL) very-low-density-, (IDL) intermediate-density- and (LDL) low-density-lipoproteins. Reverse cholesterol transport, from tissues to liver, is mediated by high-density lipoprotein (HDL) particles that are synthesised and released from the liver, and which express membrane-bound apolipoproteinA1 (Apo-A1)[4], see figure 6.1.

Much of our knowledge of the relationships of these blood lipids and lipoprotein components with coronary heart disease (CHD) has been coloured by the way in which they have been measured. Early epidemiological studies demonstrated associations between total cholesterol content (in all lipoprotein subfractions) and CHD[5]. Later, when it was possible to measure the cholesterol content of low-density lipoproteins (LDL-C) and of high-density lipoproteins (HDL-C) separately, it became clear that LDL-C was positively, and HDL-C was negatively associated with

CHD[6]. In parallel, it was observed that the total concentration of TG in lipoproteins was also positively associated with CHD[3,7].

Figure 6.1 Graphical depiction of lipoprotein subfractions and lipid content



LPL = lipoprotein lipase; FFA = Free fatty acids; TG = triglycerides

Of the three commonly measured lipid fractions (LDL-C, HDL-C and TG), it is now clear that LDL-C is causally related to CHD[8]. The evidence comes from monogenic disorders familial hypercholesteremia (FH), Mendelian randomisations studies and RCTs of LDL-C lowering drugs[8–10]. However, evidence on any causal role for TG and HDL-C is less clear. Observational associations can be affected by confounding and reverse causation, Mendelian randomisation studies have been equivocal, and RCTs of agents to lower TG (niacin, fibrates) or raised HDL-C (niacin, CETP-I) have been inconsistent[6,11].

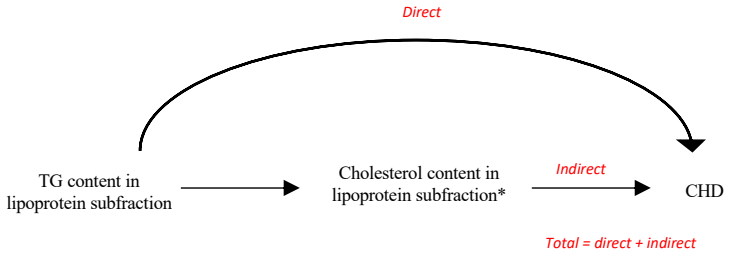
Recently questions have also arisen about the potential causal role of the cholesterol content of lipoprotein particles other than LDL. One way in which this question has been addressed is to investigate the relationship of two measures that can be derived from standard clinical chemistry measures: non-HDL-C, and remnant cholesterol[12,13]. Non-HDL-C is derived as the difference between total cholesterol and HDL-C, and represents the cholesterol content of LDL-C but also IDL, sVLDL, VLDL and CMR[14]. Remnant cholesterol is defined as total cholesterol minus LDL-C minus HDL-C[15]. It comprises the cholesterol content of triglyceride-rich lipoproteins (TRL) namely, CMR, VLDL, sVLDL and IDL, see figure 6.1.

It is possible that both TG and cholesterol in lipoprotein subfractions play a causal role in disease progression or, that one lipid dominates and accounts for the relationship of any particular lipoprotein subfraction with CHD. Advances in NMR spectroscopy now allow interrogation of the lipid content (both cholesterol and TG) of each lipoprotein subfraction individually[16]. For example, it is possible to measure both the cholesterol and TG content of IDL. This provides new opportunities to

explore relationships of individual lipoprotein subfractions and both their TG and cholesterol content on CHD risk. By identifying genetic variants that associate with individual lipoproteins and their TG or cholesterol content, it also becomes possible to investigate whether any such relationships are causal using Mendelian randomisation (MR), while recognising certain limitations to this approach[17,18].

In this Chapter, I evaluate the relationship of the TG content and the cholesterol content in fourteen lipoprotein subfractions with CHD using observational and MR methods. In the context of MR, a recent genome-wide association study has identified a number of genetic variants for TG and cholesterol in the fourteen lipoprotein subfractions measured on the Nightingale NMR spectroscopy platform. The lipoprotein subfractions can be used as instrumental variables in MR analysis[16]. In this study, the genetic instruments for TG and cholesterol in the fourteen lipoprotein subfractions are used in MR analyses to estimate the 'total' and 'direct' effect of TG and cholesterol in each lipoprotein subfraction exposure on CHD. The total effect is obtained from univariable analysis and describes the change in CHD through all potential pathways by intervening on the exposures, TG or cholesterol content, in each lipoprotein subfraction. Multivariable MR (MVMR) is then used to estimate the 'direct effect' of the TG content on CHD adjusting for the cholesterol content in each lipoprotein subfraction, figure 6.2. The same approach was taken to estimate the direct effect of the cholesterol content of each of the fourteen lipoprotein subfractions on CHD. In MVMR analysis, genetic instruments for both TG and cholesterol content in the fourteen lipoprotein subfractions are included together as instruments in the analysis[19].

Figure 6.2 Directed acyclic graph, DAG showing the total and direct effect of triglycerides and cholesterol in lipoprotein subfractions on coronary heart disease



6.2 Methods

6.2.1 Study overview

In this study the observational associations of TG and cholesterol content in fourteen lipoprotein subfractions with CHD. The observational associations were compared to associations obtained from MR analysis using genetic instruments for TG and cholesterol content in the fourteen lipoprotein subfractions, in a two-sample Mendelian randomisation design. For clarity, and to describe the effect of TG and cholesterol content in each lipoprotein subfraction with CHD, I will use 'unadjusted and adjusted effects' in the context of observational analysis, and 'total and direct effects' in MR and MVMR analysis respectively.

6.2.2 Observational data

Individual participant data were available for 14,990 participants enrolled in the UCL-Edinburgh-Bristol (UCLEB) consortium, previously described[20]. The Nightingale high-throughput NMR metabolomics platform was used to quantify TG and cholesterol concentrations in fourteen lipoprotein subfractions. These were extremely large (XXL), extra-large (XL), large (L), medium (M), small (S), and

extra-small (XS) VLDL, intermediate-density lipoprotein (IDL), L, M and S LDL, and XL, L, M, and S HDL, for details see[16,21,22]. In addition to NMR measurements, data were collated on the following participant characteristics recorded at time of metabolite measurement: age (years), sex, lifestyle factors; smoking (categorised as ever/never) and alcohol (ever/never); BMI (kg/m$^{2)}$; diastolic blood pressure (mm Hg) and type 2 diabetes mellitus prevalence. Missingness was accounted for using listwise deletion. Incident CHD was defined as the first occurrence of fatal or nonfatal, myocardial infarction (MI), or coronary revascularisation; see[20] and chapter 3 for a description of disease ascertainment methods used by the contributing UCLEB studies.

6.2.3 Genetic data sources and variant selection

Genetic associations with NMR measurement were obtained through a meta-analyses of Kettunen[23] and a de novo GWAS of UCLEB measurements. In UCLEB genotyping was completed using the Illumina Cardio-Metabochip array (PMID: 22876189) and imputed based on 1000G phase I for SNPs with MAF<0.001. All metabolic measures were mapped to a form a normal distribution using an inverse rank normal transformation, and tested against genotypes adjusting for age and gender using a general linear mode. Genetic variants were selected as MR instruments based on their association with TG (N=147 SNPs) and cholesterol (N= 171 SNPs) at a p-value threshold of $<5 \times 10^{-8}$ for each of the 14 subfractions (see appendix table 6.1 for individual SNP associations with lipids and CHD). Selected instruments were harmonised to the genetic association with CHD. In order to perform MR, the effect of the instrument on the exposure and outcome must be harmonised to be relative to the same allele. Outcome data were sourced from

CARDIoGRAMplusC4D[24], a meta-analysis of GWAS studies of European and South Asian descent involving 63,746 cases and 130,681 controls and defined cases as having myocardial infarction, acute coronary syndrome, chronic stable angina, or coronary stenosis >50%. Variants without overlap between the exposure and outcome GWAS were replaced by proxy variants (r-squared > 0.80), if available.

6.2.4 Statistical analysis

Study level heterogeneity was assessed using the Cochran's Q statistic p value and subsequently, study-specific logistic regression effect estimates were synthesised across cohorts using the fixed-effect inverse variance weighted estimator (random effects meta-analysis yielded similar point estimates,). The observational association of the TG content of the 14 lipoprotein subfractions with CHD were evaluated with adjustment for age and sex only (unadjusted association; model 1a). Model 1a was then additionally adjusted for; BMI, smoking, systolic blood pressure (SBP) and type 2 diabetes; model 1b. Model 2a was the effect of TG in each lipoprotein subfraction adjusted for age, sex and cholesterol content in each lipoprotein subfraction. Model 2a was then additionally adjusted for BMI, smoking, SBP and type 2 diabetes; model 2b. The same approach was taken to evaluate the association of the cholesterol content of the same lipoprotein subfractions with CHD.

I next performed MR analyses to evaluate the causal relationship of TG and cholesterol content in the fourteen lipoprotein subfractions with CHD. MR is more robust to confounding compared to estimates obtained via observational studies. This is due to inaccurate measurement of confounders and the inability to account for the time varying effect of the confounder, both of which lead to residual confounding

despite adjustment. MR estimates may be biased by horizontal pleiotropy (whereby the genetic instrument(s) affect CHD through non-lipid pathways). I evaluate the MR total effect of TG and cholesterol content of each lipoprotein subfraction on CHD using univariable MR analysis and the inverse variance weighted (IVW) and MR Egger estimator. A model selection framework was applied to select the most appropriate estimator (IVW or MR-Egger) for each specific exposure – outcome relationship. The MR-Egger correction is unbiased even in an extreme setting where 100% of the selected variants affect disease through horizontal pleiotropy[25]. I next conducted multivariable MR (MVMR), which allows for multiple phenotypes to be incorporated in the analysis to estimate the direct effect of TG and cholesterol content of each lipoprotein subfraction on CHD, not mediated by any other factor in the model. In the context of this study, I fit a MVMR model with genetic instruments for the TG and cholesterol content of each lipoprotein subfraction. This helps to identify which lipid, the TG or cholesterol content of each lipoprotein subfraction predominates the effect on CHD. I also applied the same model selection framework described above to select between IVW-MVMR and MVMR-Egger. Multivariable methods such as MVMR, may fail when including (conditionally) multicollinear variables –inclusion of which leads to numerically unstable models with noticeably lower precision[26]. To identify these likely erroneous results, I assessed the phenotypic correlation between TG and cholesterol content of the fourteen lipoprotein subfractions. I also assess the genetic correlation between TG and cholesterol instruments used in MVMR analysis for each lipoprotein subfraction effect estimate. I include precision estimates (the multiplicative inverse of the standard errors), where any sudden drop in precision (towards zero) is indicative of model instability and multicollinearity.

In this study the observational models 1b and 2b are considered to be analogous to the MR univariable total effect and multivariable direct effect models, respectively. The effect estimates obtained from these models are presented graphically in the results section. The point estimates and 95% confidence intervals from all observational and MR models are available in the appendix tables 6.2 and 6.3.

Where appropriate results are presented as correlation coefficients and odds ratio (OR) with 95% confidence interval (95% CI) per 1 standard deviation (SD). Analysis was conducted using R studio version 1.1.423 using 'TwoSampleMR' and 'ggplot' packages for visulisations[27,28].

6.3 Results

The composition and lipid distribution of and association of TG and cholesterol content of the 14 lipoprotein subfractions with CHD were assessed in a sample of 14,990 participants, of which 1,291 experienced CHD (table 6.1). The mean age among men (N= 7738) was 61.2 (SD: 9.9) years, 7,738 mean BMI was 26.4 (SD: 4.1) kg/m^2 and mean systolic blood pressure (SBP) was 136.2 (SD: 24.2) mmHg. Among women, mean age was 62.1 (9.0) years, BMI was 26.8 (4.5) kg/m^2 and SBP 135.5 (24.3) mmHg see Table 6.1 and appendix table 6.4 for median lipid concentrations in 14 subfractions.

Table 6.1 Description of study sample

	Men N= 7738 (52)	Women N= 7252 (48)
Age, years	61.2 (9.9)	62.1 (9.0)
BMI, kg/m^2	26.4 (3.8)	26.8 (4.5)
Smoking, ever	2671 (34.5)	2745 (37.0)
SBP, mmHg	136.2 (24.2)	135.5 (24.3)
Lipids		
TG, mmol/L	1.2 (0.7)	1.2 (0.8)
Total cholesterol, mmol/L	4.4 (1.3)	5.3 (1.5)
LDL-C, mmol/L	1.6 (0.7)	2.0 (0.8)
HDL-C, mmol/L	1.1 (0.4)	1.5 (0.5)
CHD	862/6209 (13.8)	429/6587 (6.6)

Values are mean ± SD or %.
BMI = Body mass index; SBP = systolic blood pressure; DBP = diastolic blood pressure; LDL-C = low-density lipoprotein cholesterol; HDL-C = high-density lipoprotein cholesterol

The results obtained from correlation, observational and MR analyses showed three lipoprotein subfraction groupings. These were, 1) extremely large VLDL to medium VLDL subfractions, 2) small VLDL to small LDL subfractions, and 3) very large HDL to small HDL subfractions. The association of TG and cholesterol content with CHD from observational models 1b and 2b, and MR total and direct effects are discussed based on these lipoprotein subfraction groupings. See appendix tables 2 and 3 for effect estimates and 95% CI intervals for all observartional and genetic models.

6.3.1 Evaluation of TG and cholesterol content with CHD in Extremely large VLDL, very large VLDL, large VLDL and medium VLDL subfractions with CHD

Observational association, see figure 6.4

In age, sex, BMI, smoking, SBP and type 2 diabetes adjusted analysis, there was a positive association of TG content in extremely large, very large, large and medium VLDL subfractions with CHD (OR in the range 1.13 to 1.19). The point estimates in these lipoprotein subfractions reversed with additional adjustment for the cholesterol content of each subfraction, and no longer excluded the null. The association of cholesterol content in these lipoprotein subfractions with CHD adjusted for age, sex, BMI, smoking SBP and type 2 diabetes were comparable to the associations observed for the TG content, with the OR for cholesterol content in the range 1.12 to 1.16. Cholesterol in two of the four VLDL subfractions retained a positive association with CHD following additional adjustment for the TG content of these lipoprotein subfractions, an effect that was not observed for TG content under the same adjustment. More specifically, these were cholesterol in extremely large VLDL (OR 1.46; 95% CI 1.12 to 1.92) and large VLDL (OR 1.63; 95% CI 1.01 to

2.64), adjusted for age, sex, BMI, smoking, SBP, type 2 diabetes and TG content in the lipoprotein subfractions.

Mendelian randomisation association, see figure 6.5

Univariable MR analysis estimating the total effect of TG content in extremely large, very large and medium had positive point estimates that did not exclude the null. TG in large VLDL had a positive total effect on CHD (OR 1.45; 95% CI 1.02 to 2.05). The direct effect obtained from MVMR analysis, taking into account the cholesterol content of the same lipoprotein subfractions, yielded inverse associations with CHD that did not exclude the null for all lipoprotein subfractions in this group. By comparison, the univariable MR association of the total effect of cholesterol in the lipoprotein subfractions yielded positive point estimates with CHD (OR in the range 1.39 to 8.34). The largest effect was observed for cholesterol in the very large VLDL subfraction (OR 8.34; 95% CI 3.87 to 17.94). In MVMR analysis, the direct effect of cholesterol in two lipoprotein subfractions had positive and imprecise association with CHD. These were the cholesterol in extremely large VLDL (OR 14.31; 95% CI 1.81 to 113.54) and cholesterol in medium VLDL (OR 2.73; 95% CI 1.14 to 6.54).

6.3.2 Evaluation of TG and cholesterol content in small, extra small VLDL, IDL, and large, medium and small LDL subfractions with CHD

Observational association

There was a positive association of the TG content in lipoprotein subfractions ranging from small VLDL to small LDL, with CHD (OR in the range 1.20 to 1.25) adjusted for age, sex, BMI, smoking SBP and type 2 diabetes. The largest effect was

observed for TG in extra small VLDL (OR 1.25; 95% CI 1.15 to 1.31). TG content in extra small VLDL, IDL and three LDL subfractions remained robust to additional adjustment for cholesterol content in the lipoprotein subfractions, to yield a positive effect with CHD (OR in the range 1.08 to 1.25). Comparable effect estimates were found for the cholesterol content in the lipoprotein subfractions in this group. The association of cholesterol content with CHD adjusted for age, sex, BMI, smoking, SBP and type 2 diabetes was in the range OR 1.17 to 1.25. The largest effect was observed for cholesterol in the medium LDL subfraction (OR 1.26; 95% CI 1.02 to 1.54). Additional adjustment for the TG content in the lipoprotein subfractions yielded point estimates for cholesterol that no longer excluded the null.

Mendelian randomisation association

In univariable MR analysis there was a positive association of the total effect of the TG content in small VLDL (OR 1.40; 95% CI 1.17 to 1.65) and very small VLDL (OR 2.03; 95% CI 1.62 to 2.54) with CHD. TG content in the large LDL subfraction yielded an inverse association with CHD (OR 0.70; 95% CI 0.54 to 0.89). TG in the remaining lipoprotein subfractions yielded point estimates that did not exclude the null. In MVMR analysis, accounting for cholesterol content in the lipoprotein subfractions, TG in in small VLDL and IDL had inverse associations with CHD. The remaining lipoprotein subfractions yielded positive associations that did not exclude the null. By comparison, cholesterol content in the lipoprotein subfractions in this group had positive associations with CHD in univariable MR analysis (OR in the range 1.62 to 1.74). The largest total effect was observed for cholesterol in small LDL (OR 1.74; 95% CI 1.46 to 2.08). In MVMR analysis adjusting for the TG content in the lipoprotein subfraction, there was a positive

association for the direct effect of cholesterol in IDL and large, medium and small LDL subfractions on CHD (OR in the range 1.60 to 1.80)

6.3.3 Evaluation of TG and cholesterol content in very large, large, medium, and small HDL subfractions with CHD

Observational association

Triglyceride in three of the HDL subfractions yielded positives associations with CHD, adjusted for age, sex, smoking SBP and type 2 diabetes. These were TG in extra large, medium and small HDL subfractions (OR in the range 1.07 to 1.22). Additional adjustment for the cholesterol content in these lipoprotein subfractions yielded positive associations of all four HDL subfractions with CHD (OR in the range 1.12 to 1.23). By comparison, the cholesterol content in all four HDL subfractions yielded inverse associations with CHD adjusted for age, sex, BMI, smoking, SBP and type 2 diabetes (OR in the range 0.70 to 0.93). The association of cholesterol content in the HDL subfractions with CHD remained robust with additional adjustment for the TG content to yield OR in the range 0.65 to 0.92. The largest inverse effect was observed for cholesterol in large HDL with CHD (OR 0.65; 95% CI 0.54 to 0.78).

Mendelian randomisation association

In univariable MR analysis, the total effect of TG in large, medium and small HDL had positive associations with CHD (OR in the range 1.29 to 1.42). In MVMR analysis, adjusted for the cholesterol content in each of the HDL subfractions, TG content in large and small HDL retained a positive association with CHD. By comparison, the univariable association of the total effect of cholesterol in the very

large, large, and medium HDL subfractions yielded inverse associations with CHD (OR in the range 0.67 to 0.99). This association remained robust when accounting for the TG content of the same lipoprotein subfractions in MVMR analysis (OR in the range 0.66 to 0.75)

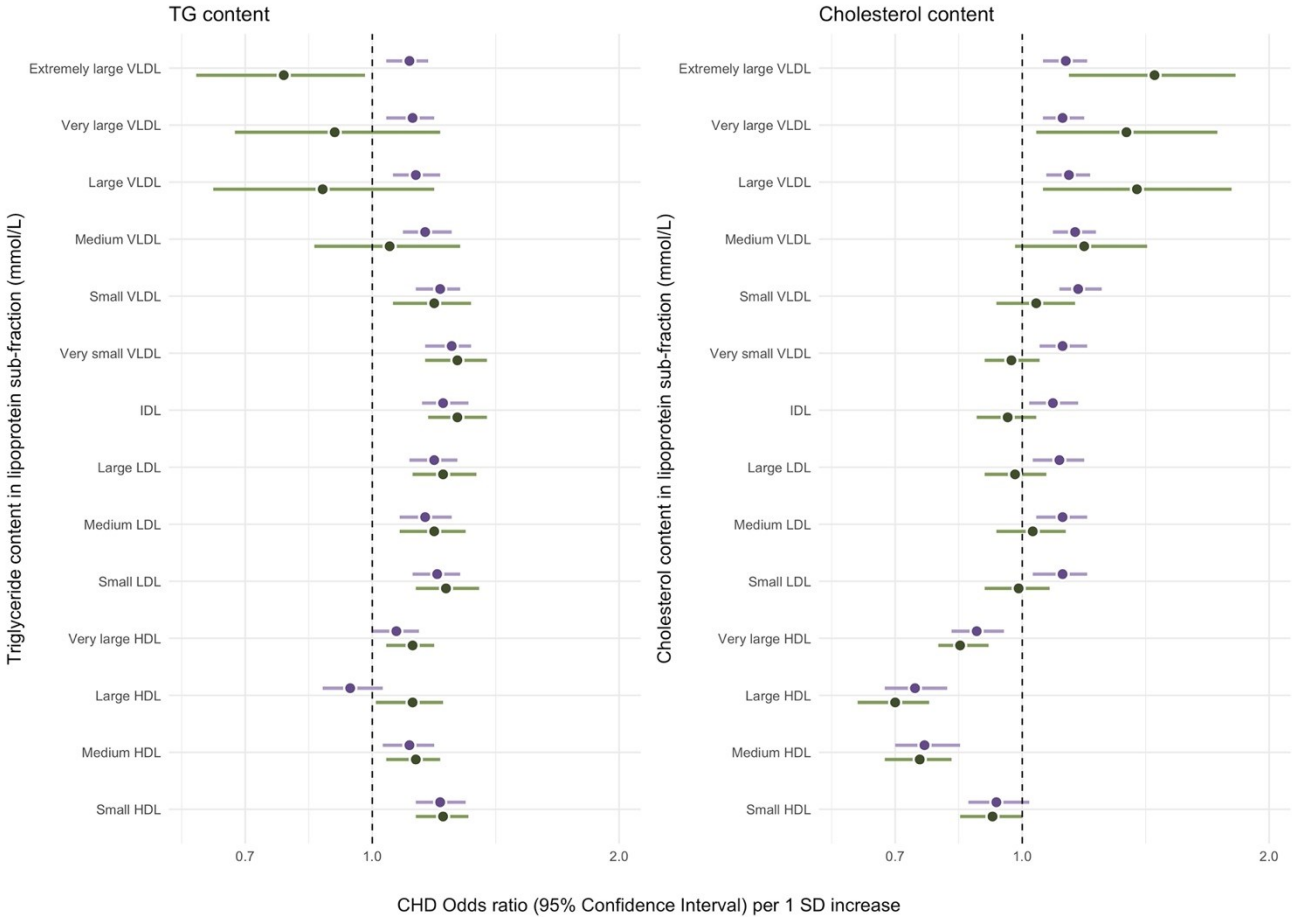
6.3.4 Correlation of the triglyceride and cholesterol content of each lipoprotein subfraction

The relative properties of TG vs cholesterol in each of the 14 lipoprotein subfractions varied depending on size and density of the lipoprotein subfraction. In four of the VLDL subfractions, specifically extremely large to medium VLDL, TG content exceeded the cholesterol content. In the small and very small VLDL subfractions, the TG and cholesterol content were more comparable. In the remaining eight lipoprotein subfractions, IDL, three LDL and four HDL subfractions, the relative cholesterol content was greater than the TG content, see figure 7 for and appendix table 4 for median TG and cholesterol concentrations.

The variation between individuals in each of the lipoprotein subfractions is observed in the correlation of TG and cholesterol content. Specifically, the TG and cholesterol content in the VLDL subfraction subclass displayed strong correlation in the range r = 0.96 to 0.54. The correlation between TG and cholesterol content in the LDL subfractions were in the range r = 0.71 to 0.73 and were weaker in HDL subclass range r = 0.29 to 0.59, figure 7, appendix table 6.5). The strong TG-cholesterol content correlation may make it difficult to separate the independent effects of TG or cholesterol content in the lipoprotein subfractions with CHD. This was observed for the MVMR analysis TG and cholesterol in the VLDL subfractions

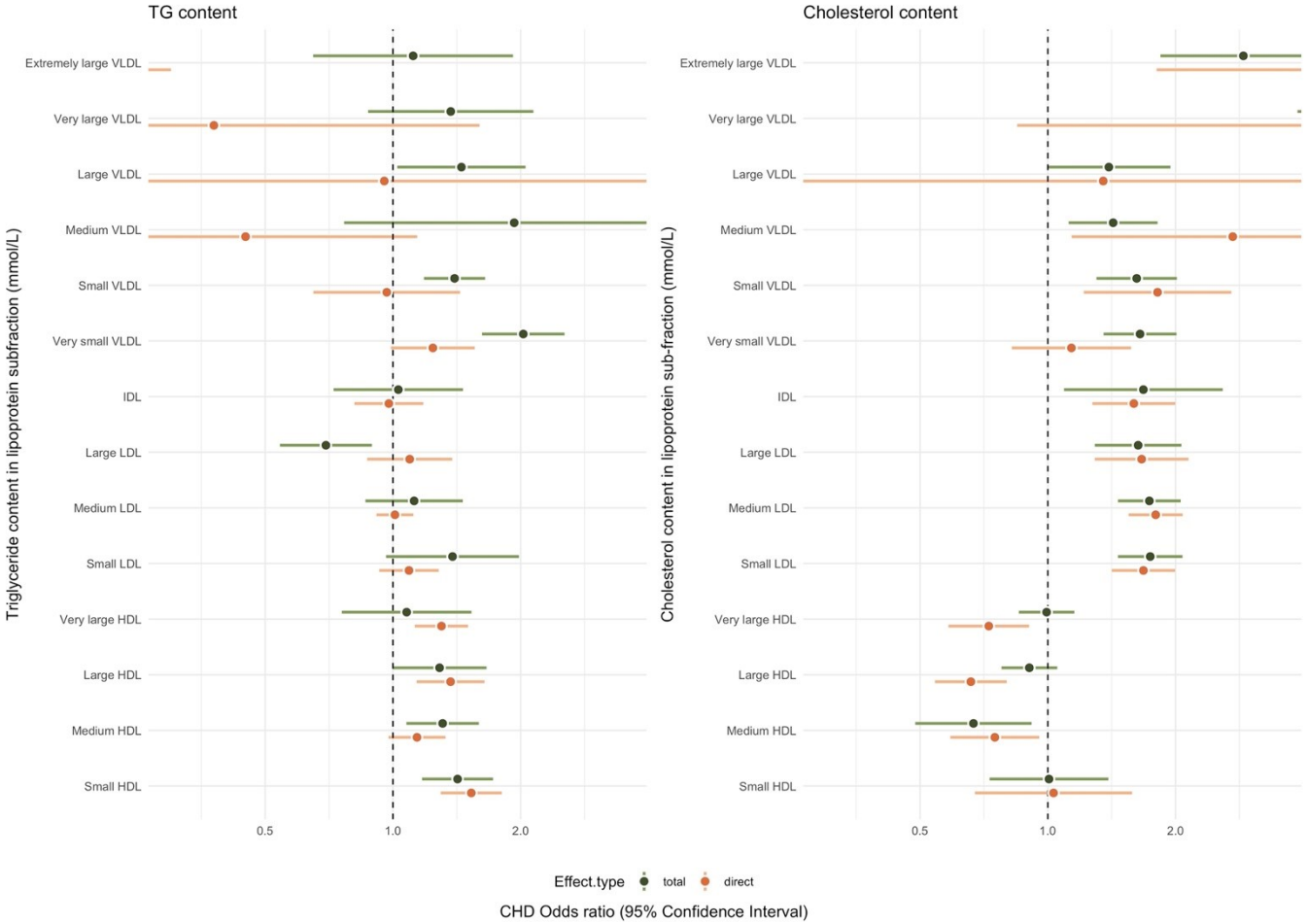
with CHD. For example, the MVMR association of the cholesterol content in the extremely large VLDL subfraction with CHD, for which the genetic correlation of TG and cholesterol was r = 0.98, yielded point estimates with wide, imprecise confidence intervals (OR 14.31, 95% CI 1.81 to 113.54). This is likely representative of the strong correlation between each lipid trait and indicates multicollinearity, causing model instability and a drop in precision, rather than an absence of effect.

Figure 6.4 Observational estimates of the effect of triglyceride and cholesterol content in fourteen lipoprotein subfraction on CHD



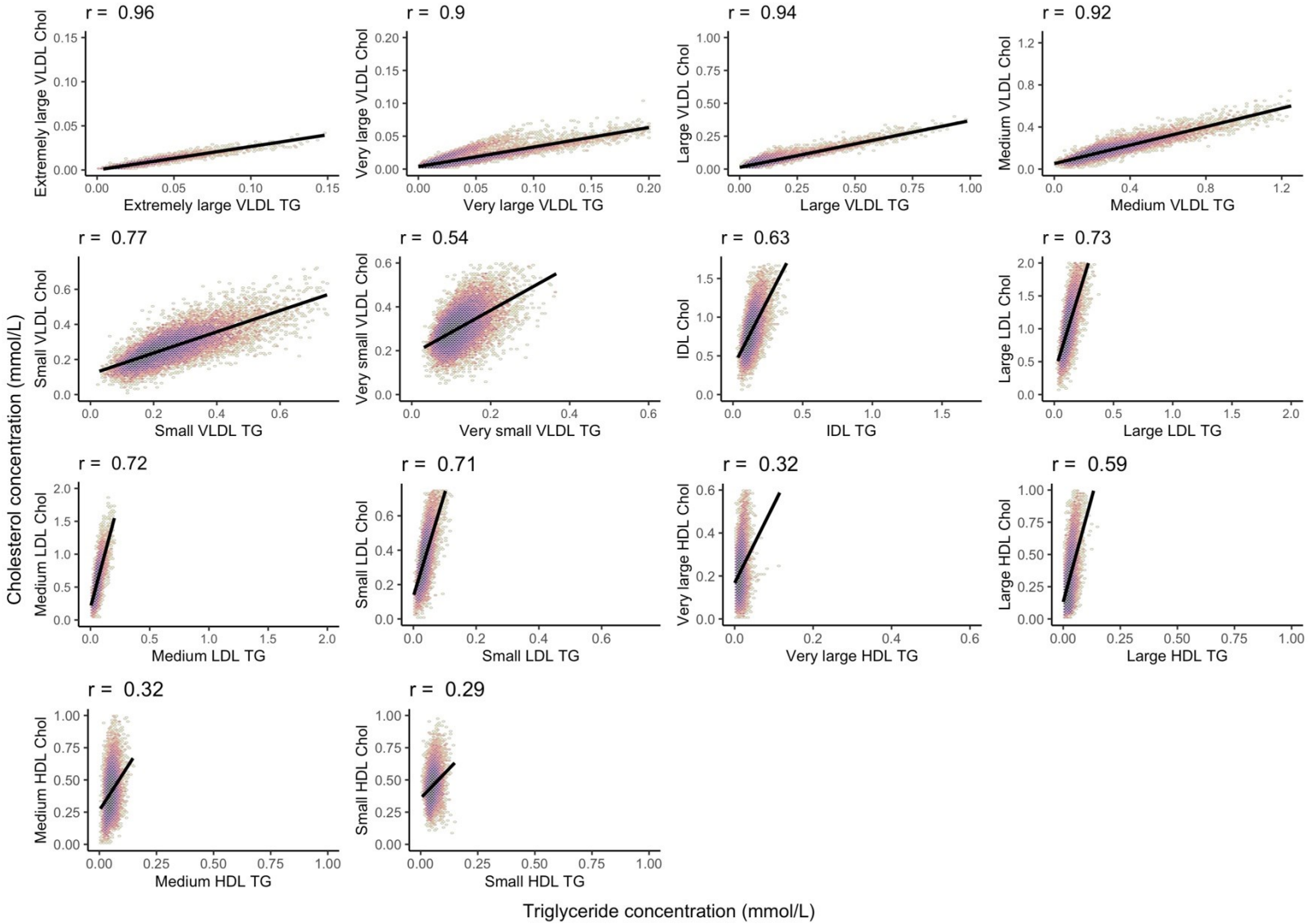
N.B. Effect estimates are adjusted for age and sex, BMI, smoking, SBP and type 2 diabetes (purple), and additional mutual lipid adjustment in each lipoprotein subfraction (green).

Figure 6.5 Univariable and MVMR estimates of the causal effect of triglyceride and cholesterol content in fourteen lipoprotein subfractions on CHD



N.B. Univariable total effect estimates (green) MVMR direct effect estimates (orange). A Rucker selection framework was applied to select between the IVW and Egger estimators, see appendix table 6.3

Figure 6.6 Correlation between triglyceride and cholesterol lipids in 14 lipoprotein subfractions



6.4 Discussion

There is equivocal evidence from observational and genetic studies that suggest a relationship between plasma TG, which represents that sum of TG across all lipoproteins, and CHD. Questions have also arisen about the potential atherogenicity of the cholesterol content in lipoprotein subfractions other than the LDL subfraction. This study explored the observational and MR association of TG and cholesterol content in fourteen lipoprotein subfractions with CHD and, evaluated which lipid trait in which lipoprotein subfractions predominates as causal. The results in this chapter are presented and discussed in three distinct groups however, they represent a continuum of lipoprotein and lipid content flux and metabolism and should be a consideration when interpreting these findings. The distinct groupings are especially relevant when discussing the potential atherogenicity of cholesterol in TRL later in this chapter.

In observational analysis, TG content in nine lipoprotein subfractions had positive associations with CHD. The positive associations were in the small VLDL, IDL, LDL and three HDL subfractions, adjusted for age, sex, BMI, smoking, type 2 diabetes and cholesterol. Using the same adjustment approach, the cholesterol content in 6 lipoprotein subfractions had mixed positive and inverse associations with CHD. These were, cholesterol in three VLDL lipoprotein subfractions had a robust, positive association with disease, an effect that was not observed for TG in the same lipoprotein subtractions. Cholesterol in three HDL displayed inverse associations with CHD. In MR analysis to ascertain causality of each lipid trait with CHD, the total effect of TG content in five lipoprotein subfractions had positive

causal associations with disease. In MVMR analysis, there was a mixed positive and inverse attenuated association of TG content in all lipoprotein subfractions with CHD. In particular, TG content VLDL subfractions yielded effect estimates with wide, imprecise confidence intervals that did not exclude the null. By comparison, univariable total effect estimates of the cholesterol content in nine lipoprotein subfractions displayed causal associations with CHD. These were cholesterol in extremely large VLDL, IDL and LDL subfractions had positive associations and two HDL subfractions had inverse associations with disease. In MVMR analysis, the cholesterol content in two VLDL subfractions retained a positive direct association with disease, an effect that was not observed for TG content in VLDL. Cholesterol content in the remaining VLDL subfractions had positive direct effect estimates that did not exclude the null. There was a direct causal effect of cholesterol in IDL and LDL subfractions with CHD, further corroborating the established association of cholesterol in LDL subfractions and risk of CHD. Cholesterol in HDL subfractions retained an inverse association with disease, also confirming established inverse association HDL-C and CHD.

Broadly speaking, in MVMR analysis the inverse, null estimates for the TG content, and positive effect estimates for the cholesterol content in the VLDL lipoprotein subfractions yielded point estimates with imprecise confidence intervals. It is likely this imprecision is representative of multi-collinearity of TG and cholesterol content in each of the VLDL lipoprotein subfractions included in the same analysis model. Collinearity between TG and cholesterol makes it difficult to deduce an independent effect of each lipid trait, rather than assume a true absence of

effect. This may be especially true for the TG content in the VLDL subfractions and should be a consideration when interpreting the results presented here.

6.4.1 Research in context

The results in this study suggest TG and cholesterol in different lipoprotein subfractions are associated with increased risk of CHD. This study provides evidence that recapitulates data reported in previous observational and MR studies. Previous studies largely investigate total TG and cholesterol concentrations measured as a sum across all lipoproteins. This study goes further to evaluate and compare the observational and MR association of TG and cholesterol content in 14 lipoprotein subfractions with CHD. The results presented here identify the cholesterol content in lipoprotein subfractions as the predominate causal lipid associated with disease. This association was found to be independent of the TG content in the lipoprotein subfractions. More specifically, cholesterol in TRL (extremely large and very large VLDL subfractions) and TRL remnants (large VLDL to IDL subfractions), confer the more prominent association with CHD, compared to TG content in the same lipoprotein subfractions. However, the association of TG content in the lipoprotein subfractions with CHD cannot be discounted. The role and underlying mechanism of TG in atherosclerosis and contribution to CHD remains uncertain. MR studies provide evidence of causality for TG-mediated pathways, yet there is largely an absence of clinical trial evidence showing CHD benefit by TG lowering[7,29]. An explanation for this may be that TG most abundant in TRL subfractions, are unlike cholesterol, too large to enter the arterial intima[30]. TG are readily degraded by most cells in the body and therefore do not accumulate in atherosclerotic plaque. Hydrolysis of TG in TRL leads to smaller,

cholesterol-rich, TRL remnants[29]. TRL remnants, differing in atherogenic potential, are able to accumulate in the arterial wall, induce local low-grade inflammation, endothelial dysfunction, and contribute to atherogenesis[31–33]. Recent findings suggest that the cholesterol carried in TRL remnants are more potent inducers of macrophage foam cells, are more atherogenic than LDL, and do not require structural modification to trigger uptake of cholesterol[34]. Evidence also suggests approximately 50% of the cholesterol found in atherosclerotic plaque is derived from cholesterol in TRL remnants[33,35]. TRL remnants in the small VLDL and IDL subfraction range carry approximately 30% cholesterol by weight and may contain up to four times more cholesterol than LDL subfractions. TRL remnants are also enriched in apoE and apoCIII, protein molecules implicated in binding and retention in the artery wall[34]. These factors likely enhance deposition of cholesterol in TRL remnants. Therefore, it is possible that TG are not disease causing per se, and instead represent a proxy marker for increased cholesterol concentrations and increased risk of CHD. This may in turn explain the absence of CHD benefit observed in clinical trials of TG lowering, with the possible exception of the REDUCE-IT trial[36].

The findings in this study are important when considering the residual risk of CHD in people already taking lipid lowering therapy for LDL-C lowering, and in the context of drug development to modify atherogenic lipid concentrations. Elevated TG is accompanied by a myriad of lipoprotein changes such as elevated non–HDL-C levels (due to increased cholesterol in TRL and their remnants, small and total LDL particles, and total apoC3[33]. All of these changes are associated with increased risk of CHD and which parameters are causal is debated. Under increased hepatic lipogenesis, the liver secretes enriched VLDL, delaying peripheral lipolysis and

clearance of TRL subfractions[34,37]. This further delays the conversion of VLDL to LDL[32]. Delayed VLDL conversion causes an increase in cholesterol in TRL, TRL remnants, and LDL subfractions that are also enriched in TG. Studies have shown non-HDL-C levels, which recapitulates cholesterol content of all apoB-containing lipoproteins (VLDL, IDL and LDL subfractions), as a better predictor of coronary disease risk[38,39]. Triglycerides and cholesterol are both carried in atherogenic lipoproteins containing an apoB molecule and recent reviews indicate that apoB is necessary for atherosclerosis to occur. This may be via the 'response to retention' hypothesis[40], in which apoB containing particles become trapped in the arterial wall[41,42]. Such studies evaluate the contribution of the total number of atherogenic lipoproteins particles quantified by measuring plasma apoB irrespective of lipoprotein lipid composition, to CHD risk. Some studies suggest the primary focus of lipid lowering therapies ought to focus on the reduction of atherogenic lipoproteins measured by apoB, rather than a reduction in TG or cholesterol. This is because the apoB protein molecule does not appear in the circulation without lipids and has led to the view that TG, LDL-C and cholesterol in TRL remnants all appear on the causal pathway to CHD[39]. Controversy exists regarding the utility of apoB and whether there is an atherogenicity gradient across apoB-containing lipoproteins. Nevertheless, multiple studies indicate that TRL remnants are at least as if not more, atherogenic than LDL subfractions and all the discussed lipoprotein lipid changes are associated with increased risk of CHD. Therefore, therapeutic interventions to lower TG will result in an accompanying reduction in cholesterol carried in TRL remnants, and LDL subfractions, reduction in and apoB particle number, and overall reduction in CHD [33,43].

Drug target MR studies have shown angiopoietin-like proteins 3 and 4 (ANGPLT3/4) inhibition as effective in lowering TG and represent emerging drug targets to lower TG and reduce CHD risk[44,45]. ANGPTL3 and ANGPTL4 are negative regulators of lipoprotein lipase (LPL), the enzyme involved in clearing circulating TG. Loss of function mutations in *ANGPTL3* are associated with lower concentration of TG, LDL and HDL-C, and lower risk of CHD[44]. Similarly, loss of function mutations in *ANGPTL4* are associated with lower TG, higher HDL-C and lower CHD risk[46]. A recent study investigating the effects of ANGPTL3 and ANGPLT4 inhibition, and LPL enhancement on NMR measured TG and cholesterol in 14 subfraction metabolites and total subfraction particle concentration, found evidence to suggest that enhancing LPL activity (either directly or via upstream effects) associates with lower TG and lower coronary artery disease (CAD) risk (CAD OR: *LPL* 0.68, [95% CI 0.56 to 0.83]; *ANGPTL4* 0.52, [95% 0.35 to 0.77], *ANGPTL3* 0.81 [95% CI -.59 to 1.10)[47].

6.4.2 Strengths and limitations

This study is the first to evaluate the observational and MR causal independent association of TG and cholesterol content in fourteen lipoprotein subfractions measured using the NMR platform, with CHD. A fundamental challenge in this study is the correlation of the lipoprotein lipids. Here we identify and utilise genetic instruments associated with TG and cholesterol across the 14 lipoprotein subfractions as discrete entities assessed in separate analyses, inferring causal effect estimates for TG and cholesterol in individual lipoprotein subfractions. Lipid lipoprotein metabolism is a continuum in the circulation and concentrations are in a

constant state of flux. Therefore, the complex physiological interrelationship is an important consideration when interpreting the findings presented here.

In the context of MR analysis, the use of a less stringent criteria for instrument selection, e.g. a more relaxed P value cut off or a more relaxed LD threshold for clumping, may have identified more variants for use as genetic instruments. However, this could also potentially have had an impact on the sensitivity and specificity of the analyses due to the potential of including weak or invalid instruments. To select the method with the most reliable causal estimate I employed a framework to select between the IVW and Egger approach, the latter considered more robust in presence of unbalanced horizontal pleiotropy. An effect that includes the null, as observed in the MVMR VLDL subfraction estimates for TG and cholesterol content with CHD, should not be interpreted as absence of effect. Rather, it is likely a reflection of model instability induced by correlated TG and cholesterol instruments included in the model. While it could be argued that alternative MR models could be applied to multicollinear settings, such methods do not allow simultaneous adjustment for multiple traits and assume the absence of horizontal pleiotropy, which may result in bias less easily recognised than numerical instability[48,49].

This study contributes to the growing body of evidence implicating cholesterol in TRL and remnant cholesterol in CHD. The findings here in particular support the recent European Society of Atherosclerosis (EAS) consensus publication, which reviews TRL and TRL remnants in atherosclerotic cardiovascular disease.

6.5 Conclusions

In summary, this study reports the cholesterol content in the 14 lipoprotein subfractions have TG independent effects on CHD, suggesting that cholesterol is the predominate lipid causally associated with CHD.

6.6 Appendices

Appendix table 6.1. Association of genetic instruments for triglycerides and cholesterol in 14 lipoprotein subfractions on coronary heart disease

Exposure	SNP	Chromosome	Effect allele	Other allele	EAF	Beta	SE	P value	Outcome	Beta outcome	SE outcome	P value outcome
XXL_VLDL_TG	rs1168002	11	A	G	0.01	-0.07	0.01	0.00	Coronary heart disease	0.02	0.01	0.14
XXL_VLDL_TG	rs1260326	11	C	T	0.86	-0.08	0.01	0.00	Coronary heart disease	-0.02	0.01	0.10
XXL_VLDL_TG	rs1729410	11	C	G	0.02	-0.05	0.01	0.00	Coronary heart disease	-0.02	0.02	0.15
XXL_VLDL_TG	rs174574	11	A	C	0.04	0.05	0.01	0.00	Coronary heart disease	0.00	0.02	0.83
XXL_VLDL_TG	rs2954027	1	A	T	0.17	-0.05	0.01	0.00	Coronary heart disease	-0.05	0.01	0.00
XXL_VLDL_TG	rs33951980	15	C	T	0.98	0.10	0.01	0.00	Coronary heart disease	-0.06	0.03	0.03
XXL_VLDL_TG	rs55730499	15	C	T	0.21	0.15	0.02	0.00	Coronary heart disease	-0.28	0.04	0.00
XXL_VLDL_TG	rs58542926	1	C	T	0.67	0.13	0.01	0.00	Coronary heart disease	0.11	0.03	0.00
XXL_VLDL_TG	rs75278536	16	G	T	0.31	-0.13	0.01	0.00	Coronary heart disease	-0.07	0.02	0.00
XS_VLDL_TG	rs11902417	19	A	G	0.03	-0.11	0.01	0.00	Coronary heart disease	-0.04	0.02	0.02
XS_VLDL_TG	rs1260326	20	C	T	0.22	-0.08	0.01	0.00	Coronary heart disease	-0.02	0.01	0.10
XS_VLDL_TG	rs1532085	2	A	G	0.23	0.09	0.01	0.00	Coronary heart disease	0.01	0.01	0.65
XS_VLDL_TG	rs17410962	2	A	G	0.38	-0.17	0.01	0.00	Coronary heart disease	-0.06	0.02	0.00
XS_VLDL_TG	rs1838504	6	A	T	0.94	-0.05	0.01	0.00	Coronary heart disease	0.00	0.02	0.76
XS_VLDL_TG	rs247617	7	A	C	0.12	-0.10	0.01	0.00	Coronary heart disease	-0.04	0.02	0.08

XS_VLDL_TG	rs261334	7	C	G	0.87	-0.14	0.01	0.00	Coronary heart disease	-0.01	0.02	0.41
XS_VLDL_TG	rs2899624	8	A	G	0.48	0.08	0.01	0.00	Coronary heart disease	0.00	0.02	1.00
XS_VLDL_TG	rs34437827	8	A	G	0.02	0.08	0.01	0.00	Coronary heart disease	0.00	0.02	0.79
XS_VLDL_TG	rs4350231	8	A	G	0.24	-0.08	0.01	0.00	Coronary heart disease	0.02	0.01	0.16
XS_VLDL_TG	rs7100409	8	A	T	0.98	0.04	0.01	0.00	Coronary heart disease	-0.01	0.01	0.50
XS_VLDL_TG	rs964184	8	C	G	0.11	-0.22	0.01	0.00	Coronary heart disease	-0.13	0.02	0.00
XL_VLDL_TG	rs1168002	11	A	G	0.86	-0.07	0.01	0.00	Coronary heart disease	0.02	0.01	0.14
XL_VLDL_TG	rs1260326	11	C	T	0.02	-0.09	0.01	0.00	Coronary heart disease	-0.02	0.01	0.10
XL_VLDL_TG	rs174577	11	A	C	0.04	0.06	0.01	0.00	Coronary heart disease	0.00	0.01	0.74
XL_VLDL_TG	rs2296065	11	A	G	0.36	-0.06	0.01	0.00	Coronary heart disease	-0.03	0.02	0.16
XL_VLDL_TG	rs2954027	1	A	T	0.59	-0.05	0.01	0.00	Coronary heart disease	-0.05	0.01	0.00
XL_VLDL_TG	rs397923	1	A	T	0.67	-0.04	0.01	0.00	Coronary heart disease	0.01	0.02	0.44
XL_VLDL_TG	rs55730499	16	C	T	0.31	0.14	0.02	0.00	Coronary heart disease	-0.28	0.04	0.00
XL_VLDL_TG	rs58542926	17	C	T	0.03	0.13	0.01	0.00	Coronary heart disease	0.11	0.03	0.00
XL_VLDL_TG	rs6586884	19	C	T	0.93	-0.14	0.01	0.00	Coronary heart disease	-0.08	0.02	0.00
XL_VLDL_TG	rs964184	19	C	G	0.44	-0.17	0.01	0.00	Coronary heart disease	-0.13	0.02	0.00
XL_HDL_TG	rs1532085	11	A	G	0.81	0.21	0.01	0.00	Coronary heart disease	0.01	0.01	0.65

XL_HDL_TG	rs2070118	11	A	G	0.00	0.06	0.01	0.00	Coronary heart disease	-0.01	0.02	0.39
XL_HDL_TG	rs2070895	1	A	G	0.00	0.25	0.01	0.00	Coronary heart disease	0.01	0.02	0.60
XL_HDL_TG	rs2250900	15	C	T	0.87	0.06	0.01	0.00	Coronary heart disease	0.03	0.02	0.06
XL_HDL_TG	rs28595548	15	A	T	0.36	0.06	0.01	0.00	Coronary heart disease	0.02	0.02	0.14
XL_HDL_TG	rs2881925	15	A	G	0.21	-0.04	0.01	0.00	Coronary heart disease	0.01	0.01	0.48
XL_HDL_TG	rs35529421	19	A	T	0.05	-0.07	0.01	0.00	Coronary heart disease	0.02	0.01	0.19
XL_HDL_TG	rs57217136	19	C	T	0.85	-0.09	0.01	0.00	Coronary heart disease	-0.13	0.03	0.00
XL_HDL_TG	rs8029919	2	A	G	0.85	-0.05	0.01	0.00	Coronary heart disease	-0.04	0.02	0.02
XL_HDL_TG	rs952275	10	G	T	0.49	0.06	0.01	0.00	Coronary heart disease	0.01	0.01	0.40
XL_HDL_TG	rs964184	11	C	G	0.77	-0.14	0.01	0.00	Coronary heart disease	-0.13	0.02	0.00
S_VLDL_TG	rs10119	6	A	G	0.02	0.05	0.01	0.00	Coronary heart disease	0.00	0.02	0.91
S_VLDL_TG	rs1042034	6	C	T	0.94	-0.11	0.01	0.00	Coronary heart disease	-0.03	0.02	0.06
S_VLDL_TG	rs1260326	7	C	T	0.12	-0.09	0.01	0.00	Coronary heart disease	-0.02	0.01	0.10
S_VLDL_TG	rs174530	7	A	G	0.87	-0.05	0.01	0.00	Coronary heart disease	0.01	0.01	0.46
S_VLDL_TG	rs2954027	8	A	T	0.48	-0.07	0.01	0.00	Coronary heart disease	-0.05	0.01	0.00
S_VLDL_TG	rs3764261	8	A	C	0.98	-0.09	0.01	0.00	Coronary heart disease	-0.04	0.02	0.08
S_VLDL_TG	rs4846920	8	A	G	0.24	0.07	0.01	0.00	Coronary heart disease	0.03	0.02	0.16

S_VLDL_TG	rs6065904	8	A	G	0.90	0.07	0.01	0.00	Coronary heart disease	-0.03	0.02	0.14
S_VLDL_TG	rs6586884	11	C	T	0.01	-0.21	0.01	0.00	Coronary heart disease	-0.08	0.02	0.00
S_VLDL_TG	rs6983170	11	C	T	0.86	0.17	0.02	0.00	Coronary heart disease	0.14	0.04	0.00
S_VLDL_TG	rs7005265	1	A	T	1.00	-0.05	0.01	0.00	Coronary heart disease	0.00	0.02	0.99
S_VLDL_TG	rs7547965	15	A	G	0.98	0.08	0.01	0.00	Coronary heart disease	-0.02	0.01	0.13
S_VLDL_TG	rs7916868	15	A	T	0.39	0.05	0.01	0.00	Coronary heart disease	-0.01	0.01	0.35
S_VLDL_TG	rs964184	15	C	G	0.79	-0.24	0.01	0.00	Coronary heart disease	-0.13	0.02	0.00
S_LDL_TG	rs12239736	1	A	T	0.32	-0.07	0.01	0.00	Coronary heart disease	0.02	0.01	0.19
S_LDL_TG	rs1260326	19	C	T	0.07	-0.08	0.01	0.00	Coronary heart disease	-0.02	0.01	0.10
S_LDL_TG	rs2070895	19	A	G	0.78	0.23	0.01	0.00	Coronary heart disease	0.01	0.02	0.60
S_LDL_TG	rs541041	19	A	G	0.44	0.10	0.01	0.00	Coronary heart disease	0.06	0.02	0.00
S_LDL_TG	rs583104	2	G	T	0.23	-0.08	0.01	0.00	Coronary heart disease	-0.10	0.02	0.00
S_LDL_TG	rs964184	2	C	G	0.38	-0.17	0.01	0.00	Coronary heart disease	-0.13	0.02	0.00
S_HDL_TG	rs10096633	15	C	T	0.86	0.17	0.01	0.00	Coronary heart disease	0.06	0.02	0.01
S_HDL_TG	rs10401969	15	C	T	0.35	-0.09	0.01	0.00	Coronary heart disease	-0.11	0.03	0.00
S_HDL_TG	rs1177541	15	A	G	0.79	0.05	0.01	0.00	Coronary heart disease	-0.02	0.01	0.16
S_HDL_TG	rs11902417	19	A	G	0.85	-0.07	0.01	0.00	Coronary heart disease	-0.04	0.02	0.02

S_HDL_TG	rs1260326	2	C	T	0.47	-0.07	0.01	0.00	Coronary heart disease	-0.02	0.01	0.10
S_HDL_TG	rs1532085	2	A	G	0.85	0.07	0.01	0.00	Coronary heart disease	0.01	0.01	0.65
S_HDL_TG	rs261334	2	C	G	1.00	-0.09	0.01	0.00	Coronary heart disease	-0.01	0.02	0.41
S_HDL_TG	rs2954027	8	A	T	0.30	-0.07	0.01	0.00	Coronary heart disease	-0.05	0.01	0.00
S_HDL_TG	rs3764261	11	A	C	0.86	-0.15	0.01	0.00	Coronary heart disease	-0.04	0.02	0.08
S_HDL_TG	rs5880	11	C	G	0.04	0.16	0.02	0.00	Coronary heart disease	0.02	0.05	0.62
S_HDL_TG	rs6073966	11	C	T	0.37	-0.09	0.01	0.00	Coronary heart disease	0.04	0.02	0.03
S_HDL_TG	rs964184	1	C	G	0.84	-0.21	0.01	0.00	Coronary heart disease	-0.13	0.02	0.00
M_VLDL_TG	rs10455872	15	A	G	0.17	0.11	0.02	0.00	Coronary heart disease	-0.28	0.04	0.00
M_VLDL_TG	rs1260326	15	C	T	0.99	-0.09	0.01	0.00	Coronary heart disease	-0.02	0.01	0.10
M_VLDL_TG	rs174568	15	C	T	0.30	-0.06	0.01	0.00	Coronary heart disease	0.00	0.01	0.90
M_VLDL_TG	rs2144300	15	C	T	0.21	0.05	0.01	0.00	Coronary heart disease	0.03	0.01	0.06
M_VLDL_TG	rs2954027	16	A	T	0.33	-0.06	0.01	0.00	Coronary heart disease	-0.05	0.01	0.00
M_VLDL_TG	rs6065904	18	A	G	0.66	0.06	0.01	0.00	Coronary heart disease	-0.03	0.02	0.14
M_VLDL_TG	rs673548	18	A	G	0.02	-0.08	0.01	0.00	Coronary heart disease	-0.03	0.02	0.06
M_VLDL_TG	rs6983170	20	C	T	0.18	0.15	0.02	0.00	Coronary heart disease	0.14	0.04	0.00
M_VLDL_TG	rs7547965	1	A	G	0.78	0.07	0.01	0.00	Coronary heart disease	-0.02	0.01	0.13

M_VLDL_TG	rs7916868	11	A	T	0.81	0.05	0.01	0.00	Coronary heart disease	-0.01	0.01	0.35
M_VLDL_TG	rs964184	1	C	G	0.02	-0.20	0.01	0.00	Coronary heart disease	-0.13	0.02	0.00
M_LDL_TG	rs181403065	4	A	T	0.99	0.14	0.01	0.00	Coronary heart disease	0.05	0.02	0.01
M_LDL_TG	rs2070895	8	A	G	0.58	0.30	0.01	0.00	Coronary heart disease	0.01	0.02	0.60
M_LDL_TG	rs583104	8	G	T	0.89	-0.09	0.01	0.00	Coronary heart disease	-0.10	0.02	0.00
M_LDL_TG	rs7177289	11	C	T	0.28	-0.25	0.01	0.00	Coronary heart disease	-0.01	0.01	0.40
M_HDL_TG	rs10401969	2	C	T	0.03	-0.17	0.02	0.00	Coronary heart disease	-0.11	0.03	0.00
M_HDL_TG	rs1077834	2	C	T	0.87	0.09	0.01	0.00	Coronary heart disease	0.01	0.02	0.59
M_HDL_TG	rs1260326	2	C	T	0.38	-0.08	0.01	0.00	Coronary heart disease	-0.02	0.01	0.10
M_HDL_TG	rs12679834	4	C	T	0.99	-0.17	0.02	0.00	Coronary heart disease	-0.09	0.02	0.00
M_HDL_TG	rs1305521	4	A	G	0.01	-0.12	0.01	0.00	Coronary heart disease	0.02	0.01	0.15
M_HDL_TG	rs1532085	4	A	G	0.99	0.07	0.01	0.00	Coronary heart disease	0.01	0.01	0.65
M_HDL_TG	rs173539	4	C	T	0.99	0.18	0.01	0.00	Coronary heart disease	0.04	0.02	0.08
M_HDL_TG	rs5880	4	C	G	0.01	0.19	0.02	0.00	Coronary heart disease	0.02	0.05	0.62
M_HDL_TG	rs58847685	4	C	T	0.02	0.09	0.01	0.00	Coronary heart disease	-0.04	0.02	0.03
M_HDL_TG	rs964184	4	C	G	0.01	-0.25	0.01	0.00	Coronary heart disease	-0.13	0.02	0.00
L_VLDL_TG	rs10455872	19	A	G	0.81	0.13	0.02	0.00	Coronary heart disease	-0.28	0.04	0.00

L_VLDL_TG	rs1168040	19	C	T	0.13	0.07	0.01	0.00	Coronary heart disease	-0.02	0.02	0.16
L_VLDL_TG	rs1260326	19	C	T	0.04	-0.09	0.01	0.00	Coronary heart disease	-0.02	0.01	0.10
L_VLDL_TG	rs157581	19	C	T	0.02	0.07	0.01	0.00	Coronary heart disease	0.03	0.03	0.23
L_VLDL_TG	rs174578	19	A	T	0.84	0.06	0.01	0.00	Coronary heart disease	0.00	0.02	0.79
L_VLDL_TG	rs2296065	19	A	G	0.03	-0.07	0.01	0.00	Coronary heart disease	-0.03	0.02	0.16
L_VLDL_TG	rs2954027	21	A	T	0.60	-0.05	0.01	0.00	Coronary heart disease	-0.05	0.01	0.00
L_VLDL_TG	rs58542926	2	C	T	0.47	0.12	0.01	0.00	Coronary heart disease	0.11	0.03	0.00
L_VLDL_TG	rs6586884	2	C	T	0.53	-0.18	0.01	0.00	Coronary heart disease	-0.08	0.02	0.00
L_VLDL_TG	rs673548	2	A	G	0.98	-0.06	0.01	0.00	Coronary heart disease	-0.03	0.02	0.06
L_VLDL_TG	rs964184	2	C	G	0.03	-0.22	0.01	0.00	Coronary heart disease	-0.13	0.02	0.00
L_LDL_TG	rs181403065	16	A	T	0.31	0.15	0.02	0.00	Coronary heart disease	0.05	0.02	0.01
L_LDL_TG	rs261334	16	C	G	0.82	-0.31	0.01	0.00	Coronary heart disease	-0.01	0.02	0.41
L_LDL_TG	rs28690720	17	G	T	0.02	-0.21	0.02	0.00	Coronary heart disease	0.00	0.02	0.88
L_LDL_TG	rs2980888	17	C	T	0.97	-0.07	0.01	0.00	Coronary heart disease	-0.08	0.02	0.00
L_LDL_TG	rs583104	17	G	T	0.01	-0.09	0.01	0.00	Coronary heart disease	-0.10	0.02	0.00
L_HDL_TG	rs10468017	15	C	T	0.04	-0.20	0.01	0.00	Coronary heart disease	-0.01	0.02	0.52
L_HDL_TG	rs2070895	15	A	G	0.98	0.22	0.01	0.00	Coronary heart disease	0.01	0.02	0.60

L_HDL_TG	rs3764261	1	A	C	0.98	0.12	0.01	0.00	Coronary heart disease	-0.04	0.02	0.08
L_HDL_TG	rs6507934	1	A	G	0.31	0.07	0.01	0.00	Coronary heart disease	-0.02	0.02	0.13
IDL_TG	rs10401845	1	C	T	0.78	-0.09	0.01	0.00	Coronary heart disease	-0.06	0.03	0.03
IDL_TG	rs1168041	11	C	T	0.86	0.07	0.01	0.00	Coronary heart disease	-0.01	0.02	0.34
IDL_TG	rs12471982	12	A	C	0.36	-0.07	0.01	0.00	Coronary heart disease	-0.04	0.02	0.04
IDL_TG	rs1260326	1	C	T	0.02	-0.05	0.01	0.00	Coronary heart disease	-0.02	0.01	0.10
IDL_TG	rs17091891	1	C	T	0.03	-0.10	0.01	0.00	Coronary heart disease	-0.06	0.02	0.00
IDL_TG	rs1838504	1	A	T	0.02	-0.08	0.01	0.00	Coronary heart disease	0.00	0.02	0.76
IDL_TG	rs2043085	15	C	T	0.16	-0.20	0.01	0.00	Coronary heart disease	-0.01	0.01	0.40
IDL_TG	rs2259816	15	G	T	0.87	-0.05	0.01	0.00	Coronary heart disease	-0.05	0.01	0.00
IDL_TG	rs247617	15	A	C	0.99	-0.06	0.01	0.00	Coronary heart disease	-0.04	0.02	0.08
IDL_TG	rs261334	15	C	G	0.97	-0.24	0.01	0.00	Coronary heart disease	-0.01	0.02	0.41
IDL_TG	rs28690720	15	G	T	0.49	-0.15	0.01	0.00	Coronary heart disease	0.00	0.02	0.88
IDL_TG	rs34335269	15	A	G	0.03	0.05	0.01	0.00	Coronary heart disease	0.01	0.01	0.32
IDL_TG	rs583104	15	G	T	0.40	-0.08	0.01	0.00	Coronary heart disease	-0.10	0.02	0.00
IDL_TG	rs9302635	15	C	T	0.97	-0.06	0.01	0.00	Coronary heart disease	0.04	0.02	0.04
IDL_TG	rs952275	15	G	T	0.78	0.10	0.01	0.00	Coronary heart disease	0.01	0.01	0.40

IDL_TG	rs964184	15	C	G	0.01	-0.16	0.01	0.00	Coronary heart disease	-0.13	0.02	0.00
Exposure	SNP	Chromosome	effect allele	other allele	EAF	Beta	SE	P value	Outcome	Beta outcome	SE outcome	P value outcome
IDL_C	rs1077835	2	A	G	0.95	-0.10	0.01	0.00	Coronary heart disease	-0.01	0.02	0.64
IDL_C	rs1168041	2	C	T	0.09	0.05	0.01	0.00	Coronary heart disease	-0.01	0.02	0.34
IDL_C	rs12721051	2	C	G	0.33	-0.18	0.01	0.00	Coronary heart disease	-0.11	0.03	0.00
IDL_C	rs12916	5	C	T	0.58	0.08	0.01	0.00	Coronary heart disease	0.04	0.01	0.01
IDL_C	rs174553	7	A	G	0.76	0.08	0.01	0.00	Coronary heart disease	0.00	0.02	0.94
IDL_C	rs2043085	9	C	T	0.20	-0.08	0.01	0.00	Coronary heart disease	-0.01	0.01	0.40
IDL_C	rs207154	11	C	T	0.86	-0.08	0.01	0.00	Coronary heart disease	-0.03	0.02	0.27
IDL_C	rs387976	11	A	C	0.33	0.08	0.01	0.00	Coronary heart disease	0.02	0.02	0.21
IDL_C	rs532436	1	A	G	0.69	0.08	0.01	0.00	Coronary heart disease	0.09	0.02	0.00
IDL_C	rs533617	12	C	T	0.49	-0.14	0.02	0.00	Coronary heart disease	-0.10	0.05	0.03
IDL_C	rs579826	12	C	T	0.14	0.08	0.01	0.00	Coronary heart disease	0.04	0.03	0.19
IDL_C	rs646776	1	C	T	0.84	-0.12	0.01	0.00	Coronary heart disease	-0.09	0.02	0.00
IDL_C	rs6511720	15	G	T	0.63	0.21	0.01	0.00	Coronary heart disease	0.13	0.03	0.00
IDL_C	rs72654473	15	A	C	0.72	-0.39	0.01	0.00	Coronary heart disease	-0.10	0.04	0.02

IDL_C	rs73066442	15	A	G	0.77	-0.06	0.01	0.00	Coronary heart disease	0.01	0.02	0.52
IDL_C	rs7575840	16	G	T	0.31	-0.10	0.01	0.00	Coronary heart disease	-0.04	0.01	0.02
IDL_C	rs8107974	16	A	T	0.05	0.12	0.01	0.00	Coronary heart disease	0.11	0.03	0.00
IDL_C	rs964184	16	C	G	0.88	-0.07	0.01	0.00	Coronary heart disease	-0.13	0.02	0.00
L_HDL_C	rs1042034	18	C	T	0.17	0.07	0.01	0.00	Coronary heart disease	-0.03	0.02	0.06
L_HDL_C	rs1077835	20	A	G	0.03	-0.19	0.01	0.00	Coronary heart disease	-0.01	0.02	0.64
L_HDL_C	rs1689791	20	A	G	0.22	0.06	0.01	0.00	Coronary heart disease	-0.01	0.01	0.43
L_HDL_C	rs174544	2	A	C	0.77	-0.08	0.01	0.00	Coronary heart disease	-0.01	0.01	0.72
L_HDL_C	rs1800961	8	C	T	0.90	0.14	0.02	0.00	Coronary heart disease	-0.03	0.04	0.49
L_HDL_C	rs1883025	9	C	T	0.23	0.06	0.01	0.00	Coronary heart disease	0.01	0.02	0.41
L_HDL_C	rs261291	1	C	T	0.78	0.15	0.01	0.00	Coronary heart disease	0.01	0.01	0.40
L_HDL_C	rs3764261	11	A	C	0.86	0.22	0.01	0.00	Coronary heart disease	-0.04	0.02	0.08
L_HDL_C	rs440183	11	A	G	0.67	-0.05	0.01	0.00	Coronary heart disease	-0.01	0.02	0.36
L_HDL_C	rs4765611	1	A	G	0.91	0.05	0.01	0.00	Coronary heart disease	0.01	0.01	0.37
L_HDL_C	rs4939884	1	C	T	0.32	0.08	0.01	0.00	Coronary heart disease	0.00	0.02	0.93
L_HDL_C	rs5880	19	C	G	0.11	-0.24	0.02	0.00	Coronary heart disease	0.02	0.05	0.62
L_HDL_C	rs6065904	19	A	G	0.49	-0.13	0.01	0.00	Coronary heart disease	-0.03	0.02	0.14

L_HDL_C	rs612577	19	C	T	0.29	-0.07	0.01	0.00	Coronary heart disease	0.06	0.02	0.02
L_HDL_C	rs67053123	19	A	T	0.77	0.09	0.01	0.00	Coronary heart disease	0.00	0.02	0.98
L_HDL_C	rs75278536	2	G	T	0.18	0.16	0.01	0.00	Coronary heart disease	-0.07	0.02	0.00
L_HDL_C	rs964184	5	C	G	0.59	0.11	0.01	0.00	Coronary heart disease	-0.13	0.02	0.00
L_HDL_C	rs9923854	7	G	T	0.76	0.07	0.01	0.00	Coronary heart disease	0.03	0.04	0.34
L_LDL_C	rs12037659	8	C	T	0.89	0.05	0.01	0.00	Coronary heart disease	-0.02	0.01	0.13
L_LDL_C	rs12916	11	C	T	0.86	0.08	0.01	0.00	Coronary heart disease	0.04	0.01	0.01
L_LDL_C	rs12976241	11	C	T	0.62	0.09	0.01	0.00	Coronary heart disease	0.04	0.02	0.09
L_LDL_C	rs142130958	1	A	G	0.34	-0.20	0.01	0.00	Coronary heart disease	-0.13	0.03	0.00
L_LDL_C	rs174555	16	C	T	0.31	-0.06	0.01	0.00	Coronary heart disease	0.00	0.02	0.94
L_LDL_C	rs207154	19	C	T	0.93	-0.08	0.01	0.00	Coronary heart disease	-0.03	0.02	0.27
L_LDL_C	rs2126263	19	A	G	0.33	0.07	0.01	0.00	Coronary heart disease	0.02	0.03	0.40
L_LDL_C	rs2965156	2	C	G	0.77	-0.05	0.01	0.00	Coronary heart disease	-0.04	0.02	0.01
L_LDL_C	rs548145	2	C	T	0.38	0.14	0.01	0.00	Coronary heart disease	0.07	0.02	0.00
L_LDL_C	rs629301	6	G	T	0.06	-0.13	0.01	0.00	Coronary heart disease	-0.11	0.02	0.00
L_LDL_C	rs73066442	8	A	G	0.54	-0.06	0.01	0.00	Coronary heart disease	0.01	0.02	0.52
L_LDL_C	rs8106814	8	C	T	0.90	-0.09	0.01	0.00	Coronary heart disease	0.00	0.02	0.97

L_LDL_C	rs964184	11	C	G	0.86	-0.07	0.01	0.00	Coronary heart disease	-0.13	0.02	0.00
L_VLDL_C	rs10401969	1	C	T	0.30	-0.15	0.01	0.00	Coronary heart disease	-0.11	0.03	0.00
L_VLDL_C	rs1042034	1	C	T	0.31	-0.07	0.01	0.00	Coronary heart disease	-0.03	0.02	0.06
L_VLDL_C	rs1168002	16	A	G	0.31	-0.08	0.01	0.00	Coronary heart disease	0.02	0.01	0.14
L_VLDL_C	rs1260326	20	C	T	0.03	-0.08	0.01	0.00	Coronary heart disease	-0.02	0.01	0.10
L_VLDL_C	rs174530	20	A	G	0.16	-0.05	0.01	0.00	Coronary heart disease	0.01	0.01	0.46
L_VLDL_C	rs2001945	2	C	G	0.31	-0.06	0.01	0.00	Coronary heart disease	-0.04	0.01	0.00
L_VLDL_C	rs3764261	8	A	C	0.90	-0.09	0.01	0.00	Coronary heart disease	-0.04	0.02	0.08
L_VLDL_C	rs439401	8	C	T	0.89	0.07	0.01	0.00	Coronary heart disease	0.01	0.03	0.76
L_VLDL_C	rs55730499	1	C	T	0.78	0.12	0.02	0.00	Coronary heart disease	-0.28	0.04	0.00
L_VLDL_C	rs75278536	1	G	T	0.75	-0.16	0.01	0.00	Coronary heart disease	-0.07	0.02	0.00
L_VLDL_C	rs964184	1	C	G	0.10	-0.22	0.01	0.00	Coronary heart disease	-0.13	0.02	0.00
M_HDL_C	rs12239737	19	A	T	0.11	-0.05	0.01	0.00	Coronary heart disease	0.02	0.01	0.19
M_HDL_C	rs1367117	19	A	G	0.49	-0.06	0.01	0.00	Coronary heart disease	0.04	0.02	0.02
M_HDL_C	rs1800961	19	C	T	0.77	0.13	0.02	0.00	Coronary heart disease	-0.03	0.04	0.49
M_HDL_C	rs247617	2	A	C	0.18	0.16	0.01	0.00	Coronary heart disease	-0.04	0.02	0.08
M_HDL_C	rs4240624	5	A	G	0.59	0.09	0.01	0.00	Coronary heart disease	0.02	0.03	0.40

M_HDL_C	rs6073966	11	C	T	0.86	-0.07	0.01	0.00	Coronary heart disease	0.04	0.02	0.03
M_HDL_C	rs75278536	1	G	T	0.59	0.12	0.01	0.00	Coronary heart disease	-0.07	0.02	0.00
M_HDL_C	rs867772	1	A	G	0.33	-0.05	0.01	0.00	Coronary heart disease	-0.02	0.02	0.22
M_HDL_C	rs964184	2	C	G	0.23	0.08	0.01	0.00	Coronary heart disease	-0.13	0.02	0.00
M_LDL_C	rs12916	2	C	T	0.38	0.08	0.01	0.00	Coronary heart disease	0.04	0.01	0.01
M_LDL_C	rs2965156	5	C	G	0.56	-0.05	0.01	0.00	Coronary heart disease	-0.04	0.02	0.01
M_LDL_C	rs562338	7	A	G	0.88	-0.14	0.01	0.00	Coronary heart disease	-0.07	0.02	0.00
M_LDL_C	rs565436	8	A	G	0.48	0.05	0.01	0.00	Coronary heart disease	0.02	0.02	0.23
M_LDL_C	rs629301	8	G	T	0.10	-0.13	0.01	0.00	Coronary heart disease	-0.11	0.02	0.00
M_LDL_C	rs6511720	1	G	T	0.77	0.19	0.01	0.00	Coronary heart disease	0.13	0.03	0.00
M_LDL_C	rs75552326	11	A	G	0.86	-0.08	0.01	0.00	Coronary heart disease	-0.03	0.02	0.27
M_LDL_C	rs8106814	15	C	T	0.39	-0.09	0.01	0.00	Coronary heart disease	0.00	0.02	0.97
M_VLDL_C	rs1260326	1	C	T	0.36	-0.07	0.01	0.00	Coronary heart disease	-0.02	0.01	0.10
M_VLDL_C	rs2678379	16	A	G	0.33	-0.10	0.01	0.00	Coronary heart disease	-0.03	0.02	0.06
M_VLDL_C	rs2954027	19	A	T	0.12	-0.07	0.01	0.00	Coronary heart disease	-0.05	0.01	0.00
M_VLDL_C	rs3735964	19	A	C	0.92	-0.16	0.01	0.00	Coronary heart disease	-0.07	0.02	0.00
M_VLDL_C	rs3846662	19	A	G	0.16	-0.06	0.01	0.00	Coronary heart disease	-0.03	0.01	0.06

M_VLDL_C	rs4350231	2	A	G	0.18	-0.09	0.01	0.00	Coronary heart disease	0.02	0.01	0.16
M_VLDL_C	rs4846914	2	A	G	0.68	-0.05	0.01	0.00	Coronary heart disease	-0.03	0.01	0.04
M_VLDL_C	rs80189144	5	C	T	0.60	-0.12	0.01	0.00	Coronary heart disease	0.05	0.02	0.02
M_VLDL_C	rs964184	8	C	G	0.89	-0.22	0.01	0.00	Coronary heart disease	-0.13	0.02	0.00
S_HDL_C	rs563290	2	A	G	0.82	0.08	0.01	0.00	Coronary heart disease	0.06	0.02	0.00
S_HDL_C	rs60223219	8	A	G	0.89	-0.10	0.02	0.00	Coronary heart disease	0.08	0.02	0.00
S_LDL_C	rs111617668	1	C	T	0.78	0.12	0.01	0.00	Coronary heart disease	0.12	0.04	0.00
S_LDL_C	rs12721051	11	C	G	0.86	-0.19	0.01	0.00	Coronary heart disease	-0.11	0.03	0.00
S_LDL_C	rs12916	11	C	T	0.33	0.08	0.01	0.00	Coronary heart disease	0.04	0.01	0.01
S_LDL_C	rs174549	1	A	G	0.10	-0.05	0.01	0.00	Coronary heart disease	0.00	0.01	0.84
S_LDL_C	rs34042070	16	C	G	0.80	-0.07	0.01	0.00	Coronary heart disease	-0.04	0.02	0.03
S_LDL_C	rs387976	19	A	C	0.11	0.07	0.01	0.00	Coronary heart disease	0.02	0.02	0.21
S_LDL_C	rs562338	19	A	G	0.67	-0.13	0.01	0.00	Coronary heart disease	-0.07	0.02	0.00
S_LDL_C	rs629301	19	G	T	0.10	-0.13	0.01	0.00	Coronary heart disease	-0.11	0.02	0.00
S_LDL_C	rs6511720	19	G	T	0.78	0.18	0.01	0.00	Coronary heart disease	0.13	0.03	0.00
S_LDL_C	rs71352247	19	G	T	0.49	-0.05	0.01	0.00	Coronary heart disease	-0.02	0.02	0.25
S_LDL_C	rs72654473	2	A	C	0.18	-0.38	0.01	0.00	Coronary heart disease	-0.10	0.04	0.02

S_LDL_C	rs75552326	2	A	G	0.07	-0.08	0.01	0.00	Coronary heart disease	-0.03	0.02	0.27
S_LDL_C	rs964184	5	C	G	0.59	-0.08	0.01	0.00	Coronary heart disease	-0.13	0.02	0.00
S_VLDL_C	rs11204085	1	C	T	0.78	-0.05	0.01	0.00	Coronary heart disease	-0.06	0.01	0.00
S_VLDL_C	rs11329061	11	A	G	0.86	0.05	0.01	0.00	Coronary heart disease	0.01	0.01	0.39
S_VLDL_C	rs1168041	1	C	T	0.29	0.08	0.01	0.00	Coronary heart disease	-0.01	0.02	0.34
S_VLDL_C	rs1260326	15	C	T	0.40	-0.07	0.01	0.00	Coronary heart disease	-0.02	0.01	0.10
S_VLDL_C	rs17671591	15	C	T	0.23	-0.07	0.01	0.00	Coronary heart disease	-0.04	0.01	0.01
S_VLDL_C	rs2043085	1	C	T	0.31	-0.04	0.01	0.00	Coronary heart disease	-0.01	0.01	0.40
S_VLDL_C	rs2070895	16	A	G	0.31	0.05	0.01	0.00	Coronary heart disease	0.01	0.02	0.60
S_VLDL_C	rs2119690	16	A	G	0.05	-0.08	0.01	0.00	Coronary heart disease	-0.06	0.02	0.00
S_VLDL_C	rs3764261	2	A	C	0.52	-0.13	0.01	0.00	Coronary heart disease	-0.04	0.02	0.08
S_VLDL_C	rs5880	2	C	G	0.38	0.12	0.02	0.00	Coronary heart disease	0.02	0.05	0.62
S_VLDL_C	rs646776	5	C	T	0.34	-0.07	0.01	0.00	Coronary heart disease	-0.09	0.02	0.00
S_VLDL_C	rs952275	8	G	T	0.28	0.09	0.01	0.00	Coronary heart disease	0.01	0.01	0.40
S_VLDL_C	rs964184	8	C	G	0.56	-0.17	0.01	0.00	Coronary heart disease	-0.13	0.02	0.00
XL_HDL_C	rs10438978	11	C	T	0.64	0.07	0.01	0.00	Coronary heart disease	0.01	0.02	0.69
XL_HDL_C	rs11865000	15	A	G	0.63	-0.09	0.01	0.00	Coronary heart disease	0.03	0.03	0.27

XL_HDL_C	rs174547	15	C	T	0.78	-0.07	0.01	0.00	Coronary heart disease	0.00	0.01	0.87
XL_HDL_C	rs1883025	16	C	T	0.10	0.06	0.01	0.00	Coronary heart disease	0.01	0.02	0.41
XL_HDL_C	rs247616	16	C	T	0.31	-0.17	0.01	0.00	Coronary heart disease	0.04	0.02	0.08
XL_HDL_C	rs261291	16	C	T	0.88	0.13	0.01	0.00	Coronary heart disease	0.01	0.01	0.40
XL_HDL_C	rs261334	18	C	G	0.18	-0.16	0.01	0.00	Coronary heart disease	-0.01	0.02	0.41
XL_HDL_C	rs4665710	20	A	C	0.22	0.05	0.01	0.00	Coronary heart disease	-0.03	0.02	0.06
XL_HDL_C	rs6065904	2	A	G	0.23	-0.14	0.01	0.00	Coronary heart disease	-0.03	0.02	0.14
XL_HDL_C	rs686030	8	A	C	0.90	0.07	0.01	0.00	Coronary heart disease	0.03	0.02	0.19
XL_HDL_C	rs75278536	9	G	T	0.23	0.10	0.01	0.00	Coronary heart disease	-0.07	0.02	0.00
XL_HDL_C	rs9923854	9	G	T	0.86	0.08	0.01	0.00	Coronary heart disease	0.03	0.04	0.34
XL_VLDL_C	rs11207994	11	C	T	0.86	0.08	0.01	0.00	Coronary heart disease	-0.02	0.02	0.26
XL_VLDL_C	rs739846	1	A	G	0.34	-0.13	0.02	0.00	Coronary heart disease	-0.07	0.03	0.01
XL_VLDL_C	rs964184	19	C	G	0.08	-0.14	0.01	0.00	Coronary heart disease	-0.13	0.02	0.00
XS_VLDL_C	rs1077834	11	C	T	0.06	0.09	0.01	0.00	Coronary heart disease	0.01	0.02	0.59
XS_VLDL_C	rs10889335	11	A	G	0.66	0.06	0.01	0.00	Coronary heart disease	-0.02	0.01	0.20
XS_VLDL_C	rs13392272	15	C	T	0.79	-0.09	0.01	0.00	Coronary heart disease	0.00	0.01	0.86
XS_VLDL_C	rs17242381	1	C	T	0.64	-0.13	0.02	0.00	Coronary heart disease	-0.13	0.04	0.00

XS_VLDL_C	rs174564	16	A	G	0.06	0.11	0.01	0.00	Coronary heart disease	0.00	0.01	0.90
XS_VLDL_C	rs2967668	19	A	G	0.87	0.11	0.02	0.00	Coronary heart disease	-0.06	0.06	0.29
XS_VLDL_C	rs3846662	19	A	G	0.87	-0.06	0.01	0.00	Coronary heart disease	-0.03	0.01	0.06
XS_VLDL_C	rs4299376	2	G	T	0.50	0.06	0.01	0.00	Coronary heart disease	0.05	0.02	0.00
XS_VLDL_C	rs5880	2	C	G	0.68	0.16	0.02	0.00	Coronary heart disease	0.02	0.05	0.62
XS_VLDL_C	rs7350481	5	C	T	0.57	-0.13	0.02	0.00	Coronary heart disease	-0.12	0.03	0.00
XXL_VLDL_C	rs11207994	11	C	T	0.86	0.09	0.01	0.00	Coronary heart disease	-0.02	0.02	0.26
XXL_VLDL_C	rs4665972	1	C	T	0.34	-0.09	0.01	0.00	Coronary heart disease	-0.02	0.01	0.10
XXL_VLDL_C	rs6586884	19	C	T	0.08	-0.17	0.02	0.00	Coronary heart disease	-0.08	0.02	0.00
XXL_VLDL_C	rs739846	2	A	G	0.40	-0.15	0.02	0.00	Coronary heart disease	-0.07	0.03	0.01
XXL_VLDL_C	rs964184	8	C	G	0.90	-0.20	0.01	0.00	Coronary heart disease	-0.13	0.02	0.00

Appendix table 6.2. Observational effect estimates for triglyceride and cholesterol in 14 subfractions meta-analysed across cohorts using the fixed effects estimator

Triglyceride subfraction effect estimates on CHD									
Lipoprotein subfraction	Numevents	NumObs	Point	LB	UB	Qpval	Qstat	model	Effect.type
xxl_vldl_tg	1225	13164	1.12	1.06	1.19	0.18	7.57	age + sex	1A
xl_vldl_tg	1206	12692	1.13	1.06	1.20	0.12	8.65	age + sex	1A
l_vldl_tg	1252	13348	1.14	1.07	1.21	0.14	8.26	age + sex	1A
m_vldl_tg	1273	13808	1.17	1.11	1.25	0.21	7.15	age + sex	1A
s_vldl_tg	1276	13829	1.22	1.15	1.30	0.38	5.33	age + sex	1A
xs_vldl_tg	1275	13826	1.27	1.19	1.35	0.79	2.39	age + sex	1A
idl_tg	1276	13830	1.26	1.17	1.34	0.69	3.09	age + sex	1A
l_ldl_tg	1276	13831	1.22	1.14	1.30	0.36	5.53	age + sex	1A
m_ldl_tg	1276	13828	1.20	1.12	1.27	0.16	7.91	age + sex	1A
s_ldl_tg	1276	13826	1.23	1.15	1.31	0.33	5.78	age + sex	1A
xl_hdl_tg	1240	13529	1.07	1.01	1.14	0.45	4.69	age + sex	1A
l_hdl_tg	1210	13256	0.95	0.88	1.03	0.42	4.99	age + sex	1A
m_hdl_tg	1274	13812	1.15	1.07	1.22	0.83	2.10	age + sex	1A
s_hdl_tg	1275	13824	1.26	1.17	1.34	0.34	5.62	age + sex	1A
xxl_vldl_tg	933	8840	1.11	1.04	1.17	0.15	8.13	age + sex + bmi + sbp + smoke + t2dm	1B
xl_vldl_tg	927	8633	1.12	1.04	1.19	0.16	7.97	age + sex + bmi + sbp + smoke + t2dm	1B
l_vldl_tg	948	8944	1.13	1.06	1.21	0.15	8.09	age + sex + bmi + sbp + smoke + t2dm	1B
m_vldl_tg	959	9141	1.16	1.09	1.25	0.17	7.73	age + sex + bmi + sbp + smoke + t2dm	1B
s_vldl_tg	959	9140	1.21	1.13	1.28	0.31	6.01	age + sex + bmi + sbp + smoke + t2dm	1B
xs_vldl_tg	959	9140	1.25	1.16	1.32	0.54	4.07	age + sex + bmi + sbp + smoke + t2dm	1B

idl_tg	959	9137	1.22	1.15	1.31	0.41	5.08	age + sex + bmi + sbp + smoke + t2dm	1B
l_ldl_tg	959	9138	1.19	1.11	1.27	0.29	6.22	age + sex + bmi + sbp + smoke + t2dm	1B
m_ldl_tg	959	9135	1.16	1.08	1.25	0.15	8.08	age + sex + bmi + sbp + smoke + t2dm	1B
s_ldl_tg	959	9133	1.20	1.12	1.28	0.20	7.24	age + sex + bmi + sbp + smoke + t2dm	1B
xl_hdl_tg	947	8996	1.07	1.00	1.14	0.25	6.67	age + sex + bmi + sbp + smoke + t2dm	1B
l_hdl_tg	916	8732	0.94	0.87	1.03	0.35	5.60	age + sex + bmi + sbp + smoke + t2dm	1B
m_hdl_tg	957	9119	1.11	1.03	1.19	0.51	4.28	age + sex + bmi + sbp + smoke + t2dm	1B
s_hdl_tg	958	9131	1.21	1.13	1.30	0.18	7.63	age + sex + bmi + sbp + smoke + t2dm	1B
xxl_vldl_tg	1225	13164	0.82	0.65	1.03	0.33	5.76	age + sex + cholesterol	2A
xl_vldl_tg	1206	12692	1.00	0.76	1.32	0.05	10.96	age + sex + cholesterol	2A
l_vldl_tg	1252	13348	0.95	0.70	1.28	0.07	10.06	age + sex + cholesterol	2A
m_vldl_tg	1273	13808	1.11	0.90	1.34	0.04	11.53	age + sex + cholesterol	2A
s_vldl_tg	1276	13829	1.23	1.11	1.36	0.20	7.32	age + sex + cholesterol	2A
xs_vldl_tg	1275	13826	1.31	1.21	1.42	0.57	3.87	age + sex + cholesterol	2A
idl_tg	1276	13830	1.34	1.23	1.43	0.68	3.12	age + sex + cholesterol	2A
l_ldl_tg	1276	13831	1.30	1.19	1.42	0.87	1.83	age + sex + cholesterol	2A
m_ldl_tg	1276	13828	1.25	1.15	1.36	0.98	0.68	age + sex + cholesterol	2A
s_ldl_tg	1276	13826	1.31	1.20	1.42	0.80	2.32	age + sex + cholesterol	2A
xl_hdl_tg	1240	13529	1.13	1.05	1.20	0.47	4.55	age + sex + cholesterol	2A
l_hdl_tg	1210	13256	1.14	1.04	1.25	0.08	9.81	age + sex + cholesterol	2A
m_hdl_tg	1274	13812	1.16	1.08	1.25	0.10	9.19	age + sex + cholesterol	2A
s_hdl_tg	1275	13824	1.26	1.17	1.35	0.29	6.15	age + sex + cholesterol	2A

xxl_vldl_tg	933	8840	0.78	0.61	0.98	0.41	5.02	age + sex + cholesterol + bmi + sbp + smoke + t2dm	2B
xl_vldl_tg	927	8633	0.90	0.68	1.21	0.01	15.56	age + sex + cholesterol + bmi + sbp + smoke + t2dm	2B
l_vldl_tg	948	8944	0.87	0.64	1.19	0.01	15.33	age + sex + cholesterol + bmi + sbp + smoke + t2dm	2B
m_vldl_tg	959	9141	1.05	0.85	1.28	0.01	15.93	age + sex + cholesterol + bmi + sbp + smoke + t2dm	2B
s_vldl_tg	959	9140	1.19	1.06	1.32	0.05	10.83	age + sex + cholesterol + bmi + sbp + smoke + t2dm	2B
xs_vldl_tg	959	9140	1.27	1.16	1.38	0.17	7.75	age + sex + cholesterol + bmi + sbp + smoke + t2dm	2B
idl_tg	959	9137	1.27	1.17	1.38	0.19	7.45	age + sex + cholesterol + bmi + sbp + smoke + t2dm	2B
l_ldl_tg	959	9138	1.22	1.12	1.34	0.38	5.31	age + sex + cholesterol + bmi + sbp + smoke + t2dm	2B
m_ldl_tg	959	9135	1.19	1.08	1.30	0.64	3.40	age + sex + cholesterol + bmi + sbp + smoke + t2dm	2B
s_ldl_tg	959	9133	1.23	1.13	1.35	0.28	6.32	age + sex + cholesterol + bmi + sbp + smoke + t2dm	2B
xl_hdl_tg	947	8996	1.12	1.04	1.19	0.08	9.90	age + sex + cholesterol + bmi + sbp + smoke + t2dm	2B
l_hdl_tg	916	8732	1.12	1.01	1.22	0.05	11.22	age + sex + cholesterol + bmi + sbp + smoke + t2dm	2B
m_hdl_tg	957	9119	1.13	1.04	1.21	0.10	9.21	age + sex + cholesterol + bmi + sbp + smoke + t2dm	2B
s_hdl_tg	958	9131	1.22	1.13	1.31	0.13	8.54	age + sex + cholesterol + bmi + sbp + smoke + t2dm	2B

Cholesterol subfraction effect estimates with CHD									
Lipoprotein subfraction	NumEvents	NumObs	Point	LB	UB	Qpval	Qstat	model	Effect.type
xxl_vldl_c	1225	13164	1.14	1.08	1.21	0.37	5.38	age + sex	1A
xl_vldl_c	1206	12692	1.14	1.07	1.20	0.28	6.24	age + sex	1A
l_vldl_c	1252	13348	1.15	1.08	1.21	0.20	7.25	age + sex	1A
m_vldl_c	1273	13808	1.19	1.12	1.25	0.26	6.56	age + sex	1A
s_vldl_c	1276	13829	1.19	1.12	1.26	0.22	7.04	age + sex	1A
xs_vldl_c	1275	13826	1.12	1.05	1.20	0.04	11.55	age + sex	1A
idl_c	1276	13830	1.07	1.00	1.15	0.00	17.15	age + sex	1A
l_ldl_c	1276	13831	1.09	1.02	1.16	0.00	19.51	age + sex	1A
m_ldl_c	1276	13828	1.09	1.02	1.17	0.00	19.17	age + sex	1A
s_ldl_c	1276	13826	1.08	1.01	1.17	0.00	19.96	age + sex	1A
xl_hdl_c	1240	13529	0.85	0.79	0.91	0.42	4.93	age + sex	1A
l_hdl_c	1210	13256	0.73	0.67	0.79	0.23	6.92	age + sex	1A

m_hdl_c	1274	13812	0.77	0.70	0.84	0.21	7.19	age + sex	1A
s_hdl_c	1275	13824	0.93	0.86	1.02	0.42	5.01	age + sex	1A
xxl_vldl_c	933	8840	1.13	1.06	1.20	0.25	6.65	age + sex + bmi + sbp + smoke + t2dm	1B
xl_vldl_c	927	8633	1.12	1.06	1.19	0.26	6.49	age + sex + bmi + sbp + smoke + t2dm	1B
l_vldl_c	948	8944	1.14	1.07	1.21	0.20	7.26	age + sex + bmi + sbp + smoke + t2dm	1B
m_vldl_c	959	9141	1.16	1.09	1.23	0.15	8.19	age + sex + bmi + sbp + smoke + t2dm	1B
s_vldl_c	959	9140	1.17	1.11	1.25	0.08	9.86	age + sex + bmi + sbp + smoke + t2dm	1B
xs_vldl_c	959	9140	1.12	1.05	1.20	0.01	14.15	age + sex + bmi + sbp + smoke + t2dm	1B
idl_c	959	9137	1.09	1.02	1.17	0.00	21.19	age + sex + bmi + sbp + smoke + t2dm	1B
l_ldl_c	959	9138	1.11	1.03	1.19	0.00	24.19	age + sex + bmi + sbp + smoke + t2dm	1B
m_ldl_c	959	9135	1.12	1.04	1.20	0.00	25.25	age + sex + bmi + sbp + smoke + t2dm	1B
s_ldl_c	959	9133	1.12	1.03	1.20	0.00	26.10	age + sex + bmi + sbp + smoke + t2dm	1B
xl_hdl_c	947	8996	0.88	0.82	0.95	0.30	6.10	age + sex + bmi + sbp + smoke + t2dm	1B
l_hdl_c	916	8732	0.74	0.68	0.81	0.26	6.54	age + sex + bmi + sbp + smoke + t2dm	1B

m_hdl_c	957	9119	0.76	0.70	0.84	0.14	8.33	age + sex + bmi + sbp + smoke + t2dm	1B
s_hdl_c	958	9131	0.93	0.86	1.02	0.43	4.88	age + sex + bmi + sbp + smoke + t2dm	1B
xxl_vldl_c	1225	13164	1.39	1.11	1.75	0.42	4.97	age + sex + triglyceride	2A
xl_vldl_c	1206	12692	1.23	0.96	1.58	0.08	9.76	age + sex + triglyceride	2A
l_vldl_c	1252	13348	1.31	1.02	1.70	0.09	9.40	age + sex + triglyceride	2A
m_vldl_c	1273	13808	1.14	0.95	1.36	0.02	13.34	age + sex + triglyceride	2A
s_vldl_c	1276	13829	1.01	0.91	1.12	0.03	12.38	age + sex + triglyceride	2A
xs_vldl_c	1275	13826	0.94	0.87	1.03	0.02	13.00	age + sex + triglyceride	2A
idl_c	1276	13830	0.92	0.85	1.00	0.00	19.08	age + sex + triglyceride	2A
l_ldl_c	1276	13831	0.93	0.85	1.02	0.00	18.52	age + sex + triglyceride	2A
m_ldl_c	1276	13828	0.98	0.89	1.07	0.00	16.97	age + sex + triglyceride	2A
s_ldl_c	1276	13826	0.94	0.86	1.03	0.00	21.01	age + sex + triglyceride	2A
xl_hdl_c	1240	13529	0.83	0.76	0.89	0.46	4.67	age + sex + triglyceride	2A
l_hdl_c	1210	13256	0.68	0.63	0.76	0.05	10.91	age + sex + triglyceride	2A

m_hdl_c	1274	13812	0.75	0.68	0.81	0.10	9.20	age + sex + triglyceride	2A
s_hdl_c	1275	13824	0.92	0.84	1.00	0.63	3.46	age + sex + triglyceride	2A
xxl_vldl_c	933	8840	1.45	1.14	1.82	0.54	4.10	age + sex + triglyceride + bmi + sbp + smoke + t2dm	2B
xl_vldl_c	927	8633	1.34	1.04	1.73	0.02	13.19	age + sex + triglyceride + bmi + sbp + smoke + t2dm	2B
l_vldl_c	948	8944	1.38	1.06	1.80	0.02	13.61	age + sex + triglyceride + bmi + sbp + smoke + t2dm	2B
m_vldl_c	959	9141	1.19	0.98	1.42	0.00	17.35	age + sex + triglyceride + bmi + sbp + smoke + t2dm	2B
s_vldl_c	959	9140	1.04	0.93	1.16	0.01	16.69	age + sex + triglyceride + bmi + sbp + smoke + t2dm	2B
xs_vldl_c	959	9140	0.97	0.90	1.05	0.00	17.83	age + sex + triglyceride + bmi + sbp + smoke + t2dm	2B
idl_c	959	9137	0.96	0.88	1.04	0.00	25.18	age + sex + triglyceride + bmi + sbp + smoke + t2dm	2B
l_ldl_c	959	9138	0.98	0.90	1.07	0.00	25.34	age + sex + triglyceride + bmi + sbp + smoke + t2dm	2B
m_ldl_c	959	9135	1.03	0.93	1.13	0.00	24.34	age + sex + triglyceride + bmi + sbp + smoke + t2dm	2B
s_ldl_c	959	9133	0.99	0.90	1.08	0.00	28.72	age + sex + triglyceride + bmi + sbp + smoke + t2dm	2B
xl_hdl_c	947	8996	0.84	0.79	0.91	0.15	8.09	age + sex + triglyceride + bmi + sbp + smoke + t2dm	2B
l_hdl_c	916	8732	0.70	0.63	0.77	0.06	10.44	age + sex + triglyceride + bmi + sbp + smoke + t2dm	2B

m_hdl_c	957	9119	0.75	0.68	0.82	0.06	10.61	age + sex + triglyceride + bmi + sbp + smoke + t2dm	2B
s_hdl_c	958	9131	0.92	0.84	1.00	0.51	4.26	age + sex + triglyceride + bmi + sbp + smoke + t2dm	2B
Triglyceride effect estimates on CHD									
Lipoprotein subfraction	NumEvents	NumObs	Point	LB	UB	model	Effect.type		
xxl_vldl_tg	1225	13164	1.15	1.05	1.26	age + sex	1A		
xl_vldl_tg	1206	12692	1.17	1.05	1.3	age + sex	1A		
l_vldl_tg	1252	13348	1.17	1.06	1.31	age + sex	1A		
m_vldl_tg	1273	13808	1.2	1.11	1.31	age + sex	1A		
s_vldl_tg	1276	13829	1.22	1.15	1.3	age + sex	1A		
xs_vldl_tg	1275	13826	1.27	1.19	1.35	age + sex	1A		
idl_tg	1276	13830	1.26	1.17	1.34	age + sex	1A		
l_ldl_tg	1276	13831	1.23	1.14	1.35	age + sex	1A		
m_ldl_tg	1276	13828	1.25	1.12	1.38	age + sex	1A		
s_ldl_tg	1276	13826	1.25	1.14	1.36	age + sex	1A		
xl_hdl_tg	1240	13529	1.07	1.01	1.14	age + sex	1A		
l_hdl_tg	1210	13256	0.95	0.86	1.04	age + sex	1A		
m_hdl_tg	1274	13812	1.15	1.07	1.22	age + sex	1A		
s_hdl_tg	1275	13824	1.26	1.17	1.34	age + sex	1A		
xxl_vldl_tg	933	8840	1.13	1.03	1.23	age + sex + bmi + sbp + smoke + t2dm	1B		
xl_vldl_tg	927	8633	1.14	1.03	1.26	age + sex + bmi + sbp + smoke + t2dm	1B		
l_vldl_tg	948	8944	1.15	1.04	1.28	age + sex + bmi + sbp + smoke + t2dm	1B		
m_vldl_tg	959	9141	1.19	1.07	1.31	age + sex + bmi + sbp + smoke + t2dm	1B		

s_vldl_tg	959	9140	1.21	1.13	1.28	age + sex + bmi + sbp + smoke + t2dm	1B
xs_vldl_tg	959	9140	1.25	1.16	1.32	age + sex + bmi + sbp + smoke + t2dm	1B
idl_tg	959	9137	1.22	1.15	1.31	age + sex + bmi + sbp + smoke + t2dm	1B
l_ldl_tg	959	9138	1.2	1.11	1.3	age + sex + bmi + sbp + smoke + t2dm	1B
m_ldl_tg	959	9135	1.21	1.07	1.35	age + sex + bmi + sbp + smoke + t2dm	1B
s_ldl_tg	959	9133	1.22	1.11	1.34	age + sex + bmi + sbp + smoke + t2dm	1B
xl_hdl_tg	947	8996	1.07	1	1.14	age + sex + bmi + sbp + smoke + t2dm	1B
l_hdl_tg	916	8732	0.94	0.87	1.03	age + sex + bmi + sbp + smoke + t2dm	1B
m_hdl_tg	957	9119	1.11	1.03	1.19	age + sex + bmi + sbp + smoke + t2dm	1B
s_hdl_tg	958	9131	1.22	1.11	1.35	age + sex + bmi + sbp + smoke + t2dm	1B
xxl_vldl_tg	1225	13164	0.82	0.61	1.12	age + sex + corresponding lipid	2A
xl_vldl_tg	1206	12692	0.87	0.53	1.42	age + sex + corresponding lipid	2A
l_vldl_tg	1252	13348	0.8	0.5	1.31	age + sex + corresponding lipid	2A
m_vldl_tg	1273	13808	0.93	0.64	1.36	age + sex + corresponding lipid	2A
s_vldl_tg	1276	13829	1.2	1.03	1.39	age + sex + corresponding lipid	2A
xs_vldl_tg	1275	13826	1.31	1.21	1.42	age + sex + corresponding lipid	2A
idl_tg	1276	13830	1.34	1.23	1.43	age + sex + corresponding lipid	2A
l_ldl_tg	1276	13831	1.3	1.19	1.42	age + sex + corresponding lipid	2A
m_ldl_tg	1276	13828	1.25	1.15	1.36	age + sex + corresponding lipid	2A
s_ldl_tg	1276	13826	1.31	1.2	1.42	age + sex + corresponding lipid	2A
xl_hdl_tg	1240	13529	1.13	1.05	1.2	age + sex + corresponding lipid	2A
l_hdl_tg	1210	13256	1.17	1.01	1.35	age + sex + corresponding lipid	2A
m_hdl_tg	1274	13812	1.19	1.06	1.32	age + sex + corresponding lipid	2A
s_hdl_tg	1275	13824	1.27	1.16	1.39	age + sex + corresponding lipid	2A

xxl_vldl_tg	933	8840	0.76	0.55	1.04	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
xl_vldl_tg	927	8633	0.7	0.38	1.27	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
l_vldl_tg	948	8944	0.64	0.34	1.2	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
m_vldl_tg	959	9141	0.79	0.49	1.27	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
s_vldl_tg	959	9140	1.11	0.9	1.35	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
xs_vldl_tg	959	9140	1.23	1.08	1.39	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
idl_tg	959	9137	1.25	1.11	1.39	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
l_ldl_tg	959	9138	1.21	1.07	1.35	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
m_ldl_tg	959	9135	1.19	1.06	1.31	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
s_ldl_tg	959	9133	1.21	1.07	1.38	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
xl_hdl_tg	947	8996	1.12	1.04	1.19	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
l_hdl_tg	916	8732	1.14	0.97	1.34	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
m_hdl_tg	957	9119	1.14	1.02	1.28	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
s_hdl_tg	958	9131	1.23	1.09	1.38	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B

Cholesterol effect estimates on CHD							
xxl_vldl_c	1225	13164	1.15	1.07	1.23	age + sex	1A
xl_vldl_c	1206	12692	1.15	1.06	1.25	age + sex	1A
l_vldl_c	1252	13348	1.17	1.08	1.28	age + sex	1A
m_vldl_c	1273	13808	1.2	1.12	1.3	age + sex	1A
s_vldl_c	1276	13829	1.21	1.11	1.31	age + sex	1A
xs_vldl_c	1275	13826	1.15	1.03	1.28	age + sex	1A
idl_c	1276	13830	1.14	0.99	1.31	age + sex	1A
l_ldl_c	1276	13831	1.17	1	1.38	age + sex	1A
m_ldl_c	1276	13828	1.2	1.01	1.4	age + sex	1A
s_ldl_c	1276	13826	1.19	1	1.4	age + sex	1A
xl_hdl_c	1240	13529	0.86	0.79	0.94	age + sex	1A
l_hdl_c	1210	13256	0.7	0.63	0.8	age + sex	1A
m_hdl_c	1274	13812	0.76	0.67	0.85	age + sex	1A
s_hdl_c	1275	13824	0.93	0.86	1.02	age + sex	1A
xxl_vldl_c	933	8840	1.14	1.06	1.22	age + sex + bmi + sbp + smoke + t2dm	1B
xl_vldl_c	927	8633	1.13	1.05	1.22	age + sex + bmi + sbp + smoke + t2dm	1B
l_vldl_c	948	8944	1.16	1.06	1.26	age + sex + bmi + sbp + smoke + t2dm	1B
m_vldl_c	959	9141	1.2	1.09	1.31	age + sex + bmi + sbp + smoke + t2dm	1B
s_vldl_c	959	9140	1.21	1.09	1.35	age + sex + bmi + sbp + smoke + t2dm	1B
xs_vldl_c	959	9140	1.16	1.03	1.32	age + sex + bmi + sbp + smoke + t2dm	1B
idl_c	959	9137	1.17	1	1.39	age + sex + bmi + sbp + smoke + t2dm	1B
l_ldl_c	959	9138	1.22	1.02	1.48	age + sex + bmi + sbp + smoke + t2dm	1B

m_ldl_c	959	9135	1.26	1.02	1.54	age + sex + bmi + sbp + smoke + t2dm	1B
s_ldl_c	959	9133	1.25	1.02	1.54	age + sex + bmi + sbp + smoke + t2dm	1B
xl_hdl_c	947	8996	0.9	0.81	0.99	age + sex + bmi + sbp + smoke + t2dm	1B
l_hdl_c	916	8732	0.7	0.62	0.81	age + sex + bmi + sbp + smoke + t2dm	1B
m_hdl_c	957	9119	0.75	0.65	0.85	age + sex + bmi + sbp + smoke + t2dm	1B
s_hdl_c	958	9131	0.93	0.86	1.02	age + sex + bmi + sbp + smoke + t2dm	1B
xxl_vldl_c	1225	13164	1.39	1.07	1.8	age + sex + corresponding lipid	2A
xl_vldl_c	1206	12692	1.3	0.87	1.93	age + sex + corresponding lipid	2A
l_vldl_c	1252	13348	1.45	0.97	2.14	age + sex + corresponding lipid	2A
m_vldl_c	1273	13808	1.31	0.93	1.82	age + sex + corresponding lipid	2A
s_vldl_c	1276	13829	1.09	0.9	1.32	age + sex + corresponding lipid	2A
xs_vldl_c	1275	13826	1	0.86	1.16	age + sex + corresponding lipid	2A
idl_c	1276	13830	0.99	0.84	1.17	age + sex + corresponding lipid	2A
l_ldl_c	1276	13831	1.01	0.84	1.22	age + sex + corresponding lipid	2A
m_ldl_c	1276	13828	1.05	0.87	1.27	age + sex + corresponding lipid	2A
s_ldl_c	1276	13826	1.03	0.84	1.27	age + sex + corresponding lipid	2A
xl_hdl_c	1240	13529	0.83	0.76	0.89	age + sex + corresponding lipid	2A
l_hdl_c	1210	13256	0.64	0.54	0.76	age + sex + corresponding lipid	2A
m_hdl_c	1274	13812	0.72	0.63	0.83	age + sex + corresponding lipid	2A
s_hdl_c	1275	13824	0.92	0.84	1	age + sex + corresponding lipid	2A
xxl_vldl_c	933	8840	1.46	1.12	1.92	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
xl_vldl_c	927	8633	1.49	0.93	2.36	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B

l_vldl_c	948	8944	1.63	1.01	2.64	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
m_vldl_c	959	9141	1.45	0.97	2.16	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
s_vldl_c	959	9140	1.17	0.93	1.46	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
xs_vldl_c	959	9140	1.05	0.88	1.26	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
idl_c	959	9137	1.06	0.86	1.3	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
l_ldl_c	959	9138	1.11	0.87	1.39	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
m_ldl_c	959	9135	1.15	0.9	1.46	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
s_ldl_c	959	9133	1.13	0.88	1.45	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
xl_hdl_c	947	8996	0.86	0.77	0.96	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
l_hdl_c	916	8732	0.65	0.54	0.78	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
m_hdl_c	957	9119	0.71	0.62	0.83	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B
s_hdl_c	958	9131	0.92	0.84	1	age + sex + corresponding lipid + bmi + sbp + smoke + t2dm	2B

Appendix table 6.3 Mendelian randomisation analyses of triglycerides and cholesterol in 14 subfractions on CHD

Lipoprotein subfraction	Triglycerides				Cholesterol			
	Point estimate	Lower bound	Upper bound	Egger selection	Point estimate	Lower bound	Upper bound	Egger selection
	Total effects							
VLDL								
Extremely large	1.12	0.65	1.92		2.89	1.84	4.53	*
Very large	1.37	0.87	2.14		8.34	3.87	17.94	*
Large	1.45	1.02	2.05	*	1.39	1.00	1.95	
Medium	1.93	0.77	4.85		1.43	1.12	1.81	
Small	1.40	1.18	1.65	*	1.62	1.30	2.01	
Very small	2.03	1.62	2.54	*	1.65	1.35	2.01	
IDL	1.03	0.72	1.46	*	1.68	1.09	2.58	*
LDL								
Large	0.70	0.54	0.89		1.63	1.29	2.06	
Medium	1.12	0.86	1.46	*	1.73	1.46	2.06	
Small	1.38	0.96	1.98	*	1.74	1.46	2.08	
HDL								
Very large	1.08	0.76	1.53		0.99	0.85	1.16	
Large	1.29	1.00	1.66		0.90	0.78	1.05	
Medium	1.31	1.08	1.59		0.67	0.49	0.92	
Small	1.42	1.17	1.72		1.01	0.73	1.39	
	Direct effects							
VLDL								
Extremely large	0.02	0.00	0.30		14.32	1.81	113.54	
Very large	0.38	0.09	1.60		4.92	0.85	28.59	
Large	0.95	0.16	5.68		1.35	0.24	7.52	

Medium	0.45	0.18	1.14	*		2.73	1.14	6.54
Small	0.97	0.65	1.44			1.81	1.22	2.71
Very small	1.24	0.99	1.56			1.14	0.82	1.57
IDL	0.98	0.81	1.18	*		1.60	1.27	2.00
LDL								
Large	1.10	0.87	1.38			1.66	1.29	2.15
Medium	1.01	0.91	1.12			1.80	1.55	2.08
Small	1.09	0.93	1.28	*		1.68	1.41	2.00
HDL								
Very large	1.30	1.13	1.50			0.73	0.58	0.90
Large	1.37	1.14	1.64	*		0.66	0.54	0.80
Medium	1.14	0.98	1.33			0.75	0.59	0.96
Small	1.53	1.30	1.81			1.03	0.67	1.58

Point estimates are odds ratios; * denotes the pleiotropy robust MR-Egger method was selected over the IVW method

Appendix table 6.4 Median, 25th and 75th percentiles of triglyceride and concentrations in 14 subfractions stratified by sex

Men (n = 7738)						
	TG concentration			Cholesterol concentration		
	25%	median	97.50%	250%	median	97.50%
VLDL						
Extremely large	0.01	0.02	0.03	0	0	0.01
Very large	0.01	0.03	0.06	0.01	0.01	0.02
Large	0.05	0.1	0.19	0.03	0.05	0.08
Medium	0.15	0.23	0.36	0.12	0.17	0.23
Small	0.17	0.22	0.29	0.22	0.28	0.34
Very small	0.09	0.11	0.13	0.27	0.32	0.38
IDL	0.1	0.12	0.14	0.74	0.87	1.02
LDL						
Large	0.08	0.1	0.12	0.89	1.07	1.29
Medium	0.04	0.05	0.06	0.49	0.6	0.74
Small	0.02	0.03	0.04	0.3	0.37	0.45
HDL						
Very large	0.01	0.01	0.02	0.16	0.22	0.29
Large	0.01	0.02	0.03	0.24	0.33	0.45
Medium	0.04	0.04	0.06	0.39	0.46	0.54
Small	0.04	0.05	0.06	0.41	0.47	0.55

Women (n = 7252)						
	25%	median	97.50%	25%	median	97.50%
VLDL						
Extremely large	0.01	0.02	0.03	0	0	0.01

Very large	0.01	0.03	0.05	0.01	0.01	0.02
Large	0.06	0.12	0.2	0.03	0.05	0.08
Medium	0.17	0.25	0.38	0.12	0.17	0.23
Small	0.18	0.24	0.33	0.22	0.28	0.34
Very small	0.1	0.12	0.16	0.27	0.32	0.38
IDL	0.11	0.14	0.17	0.74	0.87	1.02
LDL						
Large	0.1	0.12	0.15	0.89	1.07	1.29
Medium	0.04	0.06	0.07	0.49	0.6	0.74
Small	0.03	0.04	0.05	0.3	0.37	0.45
HDL						
Very large	0.01	0.01	0.02	0.16	0.22	0.29
Large	0.02	0.03	0.04	0.24	0.33	0.45
Medium	0.04	0.05	0.06	0.39	0.46	0.54
Small	0.04	0.05	0.06	0.41	0.47	0.55

Appendix table 6.5 Phenotypic correlations of triglycerides and cholesterol concentrations in 14 Subfractions

Trait one	Trait two	Pearson correlation	LB	UB	P-value
Extremely large VLDL triglycerides	Extremely large VLDL cholesterol	0.96	0.95	0.96	<0.001
Very large VLDL	Very large VLDL	0.90	0.90	0.91	<0.001
Large VLDL triglycerides	Large VLDL cholesterol	0.94	0.94	0.94	<0.001
Medium VLDL triglycerides	Medium VLDL cholesterol	0.92	0.92	0.92	<0.001
Small VLDL triglycerides	Small VLDL cholesterol	0.77	0.76	0.77	<0.001
Very small VLDL triglycerides	Very small VLDL cholesterol	0.54	0.53	0.55	<0.001
IDL triglycerides	IDL cholesterol	0.63	0.62	0.64	<0.001
Large LDL triglycerides	Large LDL cholesterol	0.73	0.73	0.74	<0.001
Medium LDL triglycerides	Medium LDL cholesterol	0.72	0.72	0.73	<0.001
Small LDL triglycerides	Small LDL cholesterol	0.71	0.71	0.72	<0.001
Very large HDL triglycerides	Very large HDL cholesterol	0.32	0.30	0.33	<0.001
Large HDL triglycerides	Large HDL cholesterol	0.59	0.58	0.60	<0.001
Medium HDL triglycerides	Medium HDL cholesterol	0.32	0.30	0.33	<0.001
Small HDL triglycerides	Small HDL cholesterol	0.29	0.27	0.30	3.18E-283

LB = lower bound, UB = upper bound

Appendix table 6.6. Correlations between genotypic effect estimates in multivariable MR analyses

Trait one	Trait two	Pearson correlation	LB	UB	P-value
Extremely large VLDL triglycerides	Extremely large VLDL cholesterol	0.98	0.92	1	6.30E-07
Very large VLDL	Very large VLDL	0.93	0.71	0.98	1.20E-04
Large VLDL triglycerides	Large VLDL cholesterol	0.99	0.97	1	1.10E-16
Medium VLDL triglycerides	Medium VLDL cholesterol	0.98	0.94	0.99	2.40E-13
Small VLDL triglycerides	Small VLDL cholesterol	0.91	0.81	0.96	6.10E-11
Very small VLDL triglycerides	Very small VLDL cholesterol	0.74	0.48	0.88	3.60E-05
IDL triglycerides	IDL cholesterol	0.69	0.44	0.84	2.00E-05
Large LDL triglycerides	Large LDL cholesterol	0.59	0.17	0.83	1.00E-02
Medium LDL triglycerides	Medium LDL cholesterol	0.53	-0.03	0.84	6.50E-02
Small LDL triglycerides	Small LDL cholesterol	0.46	0.01	0.76	4.50E-02
Very large HDL triglycerides	Very large HDL cholesterol	0.69	0.38	0.86	3.60E-04
Large HDL triglycerides	Large HDL cholesterol	0.83	0.62	0.93	3.40E-06
Medium HDL triglycerides	Medium HDL cholesterol	-0.56	-0.81	-0.15	1.20E-02
Small HDL triglycerides	Small HDL cholesterol	-0.51	-0.83	0.06	7.80E-02

LB = lower bound; UB = upper bound

1. McNamara JR, Warnick GR, Cooper GR. A brief history of lipid and lipoprotein measurements and their contribution to clinical chemistry. *Clin Chim Acta*. 2006;369(2):158-167. doi:10.1016/j.cca.2006.02.041
2. Feingold KR, Grunfeld C. *Introduction to Lipids and Lipoproteins*. MDText.com, Inc.; 2000. Accessed June 5, 2020. http://www.ncbi.nlm.nih.gov/pubmed/26247089
3. Guyton JR, Slee AE, Anderson T, et al. Relationship of Lipoproteins to Cardiovascular Events. *J Am Coll Cardiol*. 2013;62(17):1580-1584. doi:10.1016/j.jacc.2013.07.023
4. Nordestgaard BG, Varbo A. Triglycerides and cardiovascular disease. *Lancet*. 2014;384(9943):626-635. doi:10.1016/S0140-6736(14)61177-6
5. Tsao CW, Vasan RS. Cohort Profile: The Framingham Heart Study (FHS): overview of milestones in cardiovascular epidemiology. *Int J Epidemiol*. 2015;44(6):1800-1813. doi:10.1093/ije/dyv337
6. The Emerging Risk Factors Collaboration* TERF. Major Lipids, Apolipoproteins, and Risk of Vascular Disease. *JAMA*. 2009;302(18):1993. doi:10.1001/jama.2009.1619
7. Nordestgaard BG. Triglyceride-Rich Lipoproteins and Atherosclerotic Cardiovascular Disease: New Insights From Epidemiology, Genetics, and Biology. *Circ Res*. 2016;118(4):547-563. doi:10.1161/CIRCRESAHA.115.306249
8. Ebrahim S, Taylor FC, Brindle P. Statins for the primary prevention of cardiovascular disease. *BMJ*. 2014;348:g280. doi:10.1136/BMJ.G280
9. Lipid-Related Markers and Cardiovascular Disease Prediction. *JAMA*. 2012;307(23):2499-2506. doi:10.1001/jama.2012.6571
10. Bansilal S, Castellano JM, Fuster V. Global burden of CVD: Focus on secondary prevention of cardiovascular disease. *Int J Cardiol*. 2015;201:S1-S7. doi:10.1016/S0167-5273(15)31026-3
11. Rader DJ, Hovingh GK. HDL and cardiovascular disease. *Lancet*. 2014;384(9943):618-625. doi:10.1016/S0140-6736(14)61217-4
12. Varbo A, Nordestgaard BG. Remnant Cholesterol and Triglyceride-Rich Lipoproteins in Atherosclerosis Progression and Cardiovascular Disease. *Arterioscler Thromb Vasc Biol*. 2016;36(11):2133-2135. doi:10.1161/ATVBAHA.116.308305
13. Dallinga-Thie GM, Kroon J, Borén J, Chapman MJ. Triglyceride-Rich Lipoproteins and Remnants: Targets for Therapy? *Curr Cardiol Rep*. 2016;18(7). doi:10.1007/s11886-016-0745-6
14. Puri R, Nissen SE, Shao M, et al. Non-HDL Cholesterol and Triglycerides: Implications for Coronary Atheroma Progression and Clinical Events. *Arterioscler Thromb Vasc Biol*. 2016;36(11):2220-2228. doi:10.1161/ATVBAHA.116.307601
15. Varbo A, Nordestgaard BG. Remnant cholesterol and ischemic heart disease. *Curr Opin Lipidol*. 2014;25(4):266-273. doi:10.1097/MOL.0000000000000093

16. Soininen P, Kangas AJ, Würtz P, Suna T, Ala-Korpela M. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. *Circ Cardiovasc Genet*. 2015;8(1):192-206. doi:10.1161/CIRCGENETICS.114.000216
17. Swerdlow DI, Hingorani AD, Humphries SE. Genetic Risk Factors and Mendelian Randomization in Cardiovascular Disease. *Curr Cardiol Rep*. 2015;17(5):1-11. doi:10.1007/s11886-015-0584-x
18. Hingorani A, Humphries S. Nature's randomised trials. *Lancet (London, England)*. 2005;366(9501):1906-1908. doi:10.1016/S0140-6736(05)67767-7
19. Burgess S, Thompson DJ, Rees JMB, Day FR, Perry JR, Ong KK. Dissecting causal pathways using mendelian randomization with summarized genetic data: Application to age at menarche and risk of breast cancer. *Genetics*. 2017;207(2):481-487. doi:10.1534/genetics.117.300191
20. Shah T, Engmann J, Dale C, et al. Population Genomics of Cardiometabolic Traits: Design of the University College London-London School of Hygiene and Tropical Medicine-Edinburgh-Bristol (UCLEB) Consortium. Zeller T, ed. *PLoS One*. 2013;8(8):e71345. doi:10.1371/journal.pone.0071345
21. Ala-Korpela M, Kangas AJ, Soininen P. Quantitative high-throughput metabolomics: a new era in epidemiology and genetics. *Genome Med*. 2012;4(4):36. doi:10.1186/gm335
22. Soininen P, Kangas AJ, Würtz P, et al. High-throughput serum NMR metabonomics for cost-effective holistic studies on systemic metabolism. *Analyst*. 2009;134(9):1781. doi:10.1039/b910205a
23. Kettunen J, Tukiainen T, Sarin A-P, et al. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat Genet*. 2012;44(3):269-276. doi:10.1038/ng.1073
24. Nikpay M, Goel A, Won H-H, et al. A comprehensive 1000 Genomes–based genome-wide association meta-analysis of coronary artery disease. *Nat Genet*. 2015;47(10):1121.
25. Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan N, Thompson J. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Stat Med*. 2017;36(11):1783-1802. doi:10.1002/sim.7221
26. Farrar DE, Glauber RR. Multicollinearity in Regression Analysis: The Problem Revisited. *Rev Econ Stat*. 1967;49(1):92-107. doi:10.2307/1937887
27. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. *Vienna, Austria*. Published online 2018. doi:10.1108/eb003648
28. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife*. 2018;7:e34408.
29. Budoff M. Triglycerides and triglyceride-rich lipoproteins in the causal pathway of cardiovascular disease. In: *American Journal of Cardiology*. Vol 118. Elsevier Inc.; 2016:138-145. doi:10.1016/j.amjcard.2016.04.004
30. Talayero BG, Sacks FM. The role of triglycerides in atherosclerosis. *Curr*

Cardiol Rep. 2011;13(6):544-552. doi:10.1007/s11886-011-0220-3
31. Varbo A, Benn M, Smith GD, Timpson NJ, Tybjærg-Hansen A, Nordestgaard BG. Remnant cholesterol, low-density lipoprotein cholesterol, and blood pressure as mediators from obesity to ischemic heart disease. *Circ Res*. 2015;116(4). doi:10.1161/CIRCRESAHA.116.304846
32. Varbo A, Nordestgaard BG. Remnant lipoproteins. *Curr Opin Lipidol*. 2017;28(4):300-307. doi:10.1097/MOL.0000000000000429
33. Rosenson RS, Davidson MH, Hirsh BJ, Kathiresan S, Gaudet D. Genetics and causality of triglyceride-rich lipoproteins in atherosclerotic cardiovascular disease. *J Am Coll Cardiol*. 2014;64(23):2525-2540. doi:10.1016/j.jacc.2014.09.042
34. Ginsberg HN, Packard CJ, Chapman MJ, et al. Triglyceride-rich lipoproteins and their remnants: metabolic insights, role in atherosclerotic cardiovascular disease, and emerging therapeutic strategies—a consensus statement from the European Atherosclerosis Society. *Eur Heart J*. 2021;00:1-21. doi:10.1093/EURHEARTJ/EHAB551
35. Davidson MH. Triglyceride-rich lipoprotein cholesterol (TRL-C): The ugly stepsister of LDL-C. *Eur Heart J*. 2018;39(7):620-622. doi:10.1093/eurheartj/ehx741
36. Bhatt DL, Steg PG, Miller M, et al. Cardiovascular Risk Reduction with Icosapent Ethyl for Hypertriglyceridemia. *N Engl J Med*. Published online November 10, 2018:NEJMoa1812792. doi:10.1056/NEJMoa1812792
37. Chapman MJ, Ginsberg HN, Amarenco P, et al. Triglyceride-rich lipoproteins and high-density lipoprotein cholesterol in patients at high risk of cardiovascular disease: evidence and guidance for management. *Eur Heart J*. 2011;32(11):1345-1361. doi:10.1093/EURHEARTJ/EHR112
38. Richardson T, Sanderson E, Palmer T, et al. Apolipoprotein B underlies the causal relationship of circulating blood lipids with coronary heart disease. *Apolipoprotein B underlies causal Relatsh Circ blood lipids with Coron Hear Dis*. Published online August 29, 2019:19004895. doi:10.1101/19004895
39. Richardson TG, Sanderson E, Palmerid TM, et al. Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: A multivariable Mendelian randomisation analysis. *PLoS Med*. 2020;17(3):e1003062. doi:10.1371/JOURNAL.PMED.1003062
40. Hurt-Camejo E, Camejo G. ApoB-100 Lipoprotein Complex Formation with Intima Proteoglycans as a Cause of Atherosclerosis and Its Possible Ex Vivo Evaluation as a Disease Biomarker. *J Cardiovasc Dev Dis* . 2018;5(3). doi:10.3390/jcdd5030036
41. Borén J, Williams KJ. The central role of arterial retention of cholesterol-rich apolipoprotein-B-containing lipoproteins in the pathogenesis of atherosclerosis: a triumph of simplicity. *Curr Opin Lipidol*. 2016;27(5):473-483.
42. Ala-Korpela M. The culprit is the carrier, not the loads: cholesterol, triglycerides and apolipoprotein B in atherosclerosis and coronary heart

disease. *Int J Epidemiol*. 2019;48(5):1389-1392. doi:10.1093/ije/dyz068
43. Jørgensen AB, Frikke-Schmidt R, West AS, Grande P, Nordestgaard BG, Tybjærg-Hansen A. Genetically elevated non-fasting triglycerides and calculated remnant cholesterol as causal risk factors for myocardial infarction. *Eur Heart J*. 2013;34(24):1826-1833. doi:10.1093/eurheartj/ehs431
44. Wang Q, Oliver-Williams C, Raitakari OT, et al. Metabolic profiling of angiopoietin-like protein 3 and 4 inhibition: a drug-target Mendelian randomization analysis. *Eur Heart J*. Published online December 22, 2020. doi:10.1093/eurheartj/ehaa972
45. Stitziel NO, Khera A V, Wang X, et al. ANGPTL3 Deficiency and Protection Against Coronary Artery Disease. *J Am Coll Cardiol*. 2017;69(16):2054-2063. doi:10.1016/j.jacc.2017.02.030
46. Dewey FE, Gusarova V, O'Dushlaine C, et al. Inactivating variants in ANGPTL4 and risk of coronary artery disease. *N Engl J Med*. 2016;374(12):1123-1133.
47. Wang Q, Oliver-Williams C, Raitakari OT, et al. Metabolic profiling of angiopoietin-like protein 3 and 4 inhibition: a drug-target Mendelian randomization analysis. *Eur Heart J*. 2021;42(12):1160-1169. doi:10.1093/eurheartj/ehaa972
48. Burgess S, Bowden J, Dudbridge F, Thompson SG. Robust instrumental variable methods using multiple candidate instruments with application to Mendelian randomization. *arXiv Prepr arXiv160603729*. Published online 2016.
49. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015;44(2):512-525.

7 Discussion

The recently published consensus statement published by the European Atherosclerosis Society discusses the role of triglyceride-rich lipoproteins (TRL) and their remnants in atherosclerotic cardiovascular disease[1]. The consensus appraises the current understanding of the metabolism of TRL, and questions the atherogenicity of TRL, TRL remnant particles, and triglycerides as compared to LDL and LDL-C. The work in this thesis goes some way to answer the questions posed by the EAS consensus and contributes to the growing body of evidence implicating TG and cholesterol content in TRL subfractions in CHD. The next section discusses the main findings of this thesis and contextualises the contribution to the wider disease area.

7.1 Introduction

The work in the preceding chapters investigated the role of triglycerides (TG) in cardiovascular disease (CVD) using both observational and Mendelian randomisation (MR) approaches. Introductory chapter one describes atherosclerotic cardiovascular disease formation and progression. This chapter goes on to discuss the burden and risk factors for CVD, and the composition and role of lipoprotein lipids in disease formation. Chapter 1 also introduced the high-throughput proton (^{1}H) NMR metabolomics assay developed by Nightingale[2] for the quantification of fourteen lipoprotein subfractions based on size, density, and lipid content. Chapter 2 reviewed the current literature and understanding of the relationship between TG and CVD and discussed the principles of MR. Chapter 3 provided an overview of the datasets that have been used throughout this thesis and details the relevant exposure and outcome measures studied in the succeeding results chapters.

The work in Chapter 4 reports the distributions and determinants of TG and cholesterol content in the 14 lipoprotein subfractions. The reference intervals were established in a disease-free population and by sex, age, body mass index (BMI), smoking status, as well as in participants with CVD and Type 2 diabetes. The largest interval range for TG was observed in the medium VLDL subfraction (2.5^{th} 97.5^{th} percentile; 0.08 to 0.68 mmol/L), and for cholesterol in the large LDL subfraction (0.47 to 1.45 mmol/L). TG concentrations in all subclasses increased with increasing age and BMI. However, for cholesterol concentrations, increases were more gradual as compared to TG and were largely comparable between men and women. Lipid reference interval ranges are necessary to support decision making and apply analytical data in healthcare delivery. NMR methodology offers the potential for

more granular quantification of lipid content in lipoproteins, the utility of which might contribute to greater insights for the role of TG and cholesterol in disease. Due to the low cost, accuracy and additional information provided by NMR profiling over standard clinical chemistry-based lipid measures, it is likely NMR profiling of lipoprotein lipids will become available in clinical care in the future and it is envisioned the reference intervals defined in this chapter will aid future clinical decision making.

Observational data suggest higher concentration of total TG, which represents the sum of TG across all lipoproteins, associates with CHD. However, the association attenuates to the null when accounting for additional CVD risk factors. In Chapter 5 I investigated the association of TG content in fourteen NMR lipoprotein subfractions with CVD. The results from this chapter finds evidence to support the positive association of TG in 13 lipoprotein subfractions with CHD, with an attenuation of effects when accounting for LDL and HDL-cholesterol. There was no clear evidence of an association of TG in any lipoprotein subfraction on stroke.

Questions have arisen about the potential causal role of the cholesterol content of lipoprotein particles other than LDL. This refers to the cholesterol content of VLDL and IDL lipoprotein particles, collectively termed triglyceride-rich lipoproteins (TRL, and TRL remnants that become enriched in cholesterol following hydrolysis of TG in the very large VLDL particles). This question has recently been approached by evaluating the relationship of non-HDL-cholesterol and remnant cholesterol. Non-HDL-cholesterol is derived using total cholesterol minus HDL-C, and encompasses all lipoprotein particles containing a surface apoB particle (namely

VLDL, IDL and LDL). Remnant cholesterol is estimated using total cholesterol minus LDL-C minus HDL-C and represents the cholesterol content of VLDL and IDL. The results in Chapter 6 investigates the potential atherogenicity of TG and cholesterol in each of the fourteen lipoprotein subfractions and their association with CHD. In an observational and genetic approach using univariable and multivariable Mendelian randomisation methods, Chapter 6 identifies cholesterol in TRL as the predominate lipid causal in disease. However, despite this finding, the relevance of TG in CHD cannot be discounted. While TG may not be causal *per se* in CHD, it may still represent a proxy marker for elevated disease risk. This may be especially relevant in the context of ongoing drug development targeted at modifying TG or targeting TG-mediated pathways for disease reduction.

7.2 Research in context

More than 25 years ago, increased concentrations of TG were regarded as a cardiovascular risk factor, similar to high LDL-C concentrations[7,8]. Clinical practice of treating both lipid fractions to prevent CVD and reduce the risk of acute pancreatitis at one time were driven by clinicians and the Zilversmit hypothesis, postulating atherogenesis as a post prandial occurrence, and raised TG and TRL as a main cause of atherosclerosis[9]. The research focus shifted to raised LDL-C concentrations as the main target for CVD prevention. This was due to multiple scientific breakthroughs including, the LDL-oxidation hypothesis, identification of LDL-receptor mutations as a cause of familial hypercholesteremia, and discovery of statins as an inhibitor of HMG-CoA (3-hydroxy-3-methyl-glutaryl-coenzyme A) reductase as an effective way of reducing LDL-C concentrations and CVD risk[10]. The breakthrough 4S trial in 1994 reported reduced CVD and all-cause mortality

after LDL-C lowering with simvastatin, and cemented LDL-C lowering as the prime lipid target[11]. Trials of highly effective statins set the standard for intervening on CVD risk factors that show a large CVD benefit. By comparison to the LDL-C lowering trials, trials targeted at reducing TG for CVD benefit failed to meet expectations. Raised TG concentrations are strongly associated with low HDL-C, and research in the post stain era has focused on HDL-C investigations and to a lesser extent on TG. Despite observational associations of HDL-C with atheroprotection, more recent evidence from genetic studies oh HDL-C remain equivocal and clinical trials have been terminated early or have not shown CVD benefit[12–14]. Possible reasons for failure may be due to failure of the compound or failure of the biomarker. As such, causality remains uncertain but does not preclude the possibility that raising HDL-C through a target such as CETP may prove to be beneficial. Drug target MR paves the way to answering such questions. HDL-C nonetheless is associated with TG and may represent a bystander indicator of high TG levels, renewing the interest in raised TG concentration associations with CVD[8,14,15].

The evidence discussed above represents total TG concentrations summed across all lipoproteins, quantified using conventional methods in clinical laboratories. Triglyceride concentrations vary across the different lipoproteins. I hypothesised that the differential concentrations infer a diverse association with disease. This thesis aimed to investigate TG concentrations in 14 lipoprotein subfractions quantified using the NMR platform as a novel way to interrogate the relationship between TG and CHD beyond total TG measurements. In the remainder

of the Discussion, I review the main results from Chapters 4-6 in the context of existing observational, genetic, and randomised trial evidence.

Triglycerides are an efficient means of storing excess energy. In the blood, TG and cholesteryl esters (the cargo carried in lipoproteins) circulate in the core of spherical lipoproteins (the carrier), with apolipoproteins on the surface providing structural stabilisation[16]. Apolipoprotein B (ApoB) is the classifying apolipoprotein the surface of TRL and is present as its ApoB100 isoform on VLDL secreted from the liver, which are then metabolised to IDL and LDL in the circulation[7,17]. Chylomicrons from the intestine contain the apob48 isotope and are metabolised to remnant particles but not to IDL and LDL. The association between TG and CHD discussed earlier in this thesis is well established however, debate as to whether TG are causal in CHD persists. This is mainly due to two main questions, firstly are TG molecules (the cargo carried in lipoproteins) or TRL (the carriers of lipid cargo) causal in CHD? And secondly, for TRL, is it the TG or cholesterol that gives rise to risk? Several early animal studies showed higher cholesterol concentrations in low-density lipoprotein (LDL-C) are atherogenic, yet the same was not true for higher TG concentrations[8,9,18]. This has led to a view that TG are not disease causing per se, but represent a marker of high cholesterol concentrations in TRL[19]. This is thought to be due to the very large TRL and chylomicrons rich in TG, being too large to cross the endothelial barrier, penetrate the intima and cause atherosclerosis[20]. Biological insights into the mechanism by which TRL affect atherosclerosis and CHD have come from animal and human studies, and more recently genetic studies using the MR approach[19,21]. Partial lipolysis and liberation of TRL-TG by LPL, forms a cholesterol-rich population of particles, intermediate in size to the very large

VLDL and LDL, termed TRL remnants[1]. Unlike larger lipoproteins, TRL remnants can enter the intima at a slower speed compared to LDL, where they are preferentially trapped due to their larger molecular size making re-entry into the lumen more difficult[7]. Once in the arterial intima, LPL expressed in macrophages and foam cells, further degrades TG in TRL releasing free fatty acids (FA) and monoacylglycerols causing local inflammation[7]. This is evident with saturated FA but not polyunsaturated forms such as omega-3 free FA. TRL decrease in size following lipolysis and the core content of TG decreases and cholesterol content increases.[22] Inefficient lipolysis can cause TRL remnant accumulation and remodelling, acquiring more cholesterol. The remodelling can lead to remnant particles no longer susceptible to lipolysis and a prolonged residence time in the circulation[1]. The accumulating TRL remnants are directly taken up by macrophages via the VLDL receptor, turning these cells into macrophage foam cells rich in indigestible cholesterol[7,23]. Cholesterol content in TRL and TRL remnants are also referred to in the literature as remnant cholesterol and can be calculated using a standard lipid profile as remnant cholesterol (mmol/L): total cholesterol minus LDL-C minus HDL-C, or directly using the NMR platform[21,24]. The question remains as to whether cholesterol in TRL remnants leads to low-grade inflammation as detected by C-reactive protein in plasma. Genetic studies using the MR approach found a 1 mmol/L higher concentration of remnant cholesterol (defined as the cholesterol in TRL remnants namely; VLDL and IDL) was associated with 28% higher C-reactive protein level, the same study found an absence of causal association for LDL-C and CRP[25–27]. This suggests that elevated remnant cholesterol and thus TRL remnants are causally associated with low-grade inflammation whereas LDL-C is not, further

supporting the notion that cholesterol in TRL remnants promote atherogenesis independently of LDL-C via the inflammatory cascade described earlier.

While the extent to which TG promote atherosclerosis is disputed, evidence from MR studies strongly implicate TG-mediated pathways in development of CHD. Do et el[28] developed a statistical framework to dissect the causal influences among the correlated lipids TG, LDL-C and HDL-C. In an analytical approach of 185 common variants associating with lipid traits across the genome, 94 were associated with TG, of those, 7 associated with TG only, with the other 87 also associating with LDL-C or HDL-C. This study demonstrated firstly that SNPs with the same direction and a similar magnitude of association for both TG and LDL-C tend to associate with CHD risk. Second, SNPs that have an exclusive effect on TG also associate with CHD and thirdly, the strength of a SNP's effect on TG levels is correlated with the magnitude of its effect on CHD risk, even after accounting for the same SNP's effect on LDL-C and/or HDL-C levels[28]. A further study performed MR meta-analyses in 17 studies including 62,199 participants and 12,099 CHD events[29]. Using weighted allele scores based on multiple SNPs associated with TG from across the genome, this study reports unrestricted allele score (67 SNPs) and the restricted allele score (27 SNPs) were both associated with CHD (OR: 1.62; 95% CI: 1.24 to 2.11 and 1.61; 95 % CI: 1.00 to 2.59, respectively)[29]. The most difficult problem in understanding causality between TG and CHD is to deal with pleiotropy as nearly all TG-associated SNPs have additional effects on LDL-C and HDL-C, and can perturb estimates derived from MR analysis. This is not surprising as TG are carried in multiple lipoprotein subclasses in the blood and the measured TG concentrations reflect contributions a continuum of physiological processes.

Multivariable MR offers an approach to ascertain causality of TG more robustly than traditional univariate MR analysis, by allowing for adjustment of confounding variables affecting the exposure-outcome pathway. A study by Allara *et al* investigating the genetic determinants of lipid fractions from the Global Lipids Genetics Consortium and multiple cardiovascular outcomes from 367,703 participants in the UK Biobank, report evidence to support the independent causal role of TG on outcomes adjusting for correlated traits LDL-C and HDL-C. The study reports TG effect on CHD, OR: 1.25 (95% CI: 1.12-1.40). TG also associated with aortic stenosis (OR: 1.29, 95% CI: 1.04-1.61), and hypertension (OR 1.17, 95% CI: 1.07-1.27)[30]. A further study by Richardson *et al* identified 440 SNPs associated with TG from a GWAS conduced on lipid traits in over 440,000 participants in the UK Biobank, and used multivariable MR to disassociate the causal role of lipids in CHD. This study reports when the causal effect of TG was assessed using univariate MR, the odds ratio for CHD was 1.34; (95% CI: 1.25 to1.44), and when accounting for the other lipid traits in multivariable MR the association remained robust OR 1.12 (95% CI: 1.02–1.23)[31].

The studies described above utilise SNPs from across the genome, termed 'genome-wide MR'. This approach is useful to help determine causality of a biomarker such as TG, and has been the more relevant approach in the context of this thesis. The second approach is to select specific SNPs from a gene of interest, termed '*cis*-MR' and is common when the exposure of interest is a specific drug target such as a protein. The findings described above from genome-wide MR studies, concur with insights of specific genes predominantly related to TG concentrations that also affect risk for CHD. SNPs in genetic determinants of TG,

LPL, apolipoprotein A-V *(APOA5),* apolipoprotein CIII *(APOC3),* angiopoietin-like 3 *(ANGPTL3),* and *ANGPTL4* all share a common characteristic that they encode lipoprotein lipase or, encode regulators of lipoprotein lipase, the enzyme that hydrolyses TG in lipoprotein particles, and all consistently demonstrate associations with CHD events[8,28,32]. A loss of function mutation in the *LPL* and *APOA5* genes are associated with increased events[33]. Mutations in the gene encoding TG degrading enzyme lipoprotein lipase, *LPL*, leads to lifelong elevated TG concentrations and increased risk of CHD. Individuals heterozygous for LPL deficiency were 4.9 times more common in patients with CHD than those in the general population, and for 4 *LPL* versus 0 TG reducing alleles, a TG reduction of 36% resulted in a 46% reduction in risk of CHD[8,34–36]. For *APOA5* variation, a genetic doubling in plasma TG was associated with a corresponding 1.9 times causal and 1.6 times observational CHD risk, whereas TG reducing alleles led to a 35% reduction in plasma TG concentrations and a corresponding CHD risk of 24%[8,19]. Concordant findings were reported in a large MR study using a single *APOA5* genetic variant (OR 1.18; 95% CI 1.11 to 1.26), providing further evidence to support a causal association between triglyceride-mediated pathways and CHD[33]. Similarly, LOF mutations in *APOC3, ANGPTL3* and *ANGPTL4* are associated with a lower risk of CHD[37–39]. A study in 2008 reported *APOC3* loss of function heterozygosity reduced TG and remnant cholesterol concentrations, and reduced coronary artery calcification[40]. More recent evidence in 2014 from the Copenhagen general population found a 44% reduction of TG and a 41% reduction in CHD risk[39]. In a further study including 18 cohorts, *APOC3* loss of function caused a 39% and 40% reduction in TG and CVD risk, respectively[41]. In multivariable MR *cis*-analysis, accounting for instrument association with LDL-C and HDL-C, the effect of TG on CHD for instruments

derived from the *APOC3* was OR: 1.27 (95% CI: 1.09-1.47) and *LPL* regions OR:1.65 (95% CI: 1.41-1.93)[30]. Recently published findings investigating genetic variants in *ANGPTL3* and *ANGPTL4* associated lipid measures represent novel emerging drug targets to lower TG and reduce CHD risk[4]. The utility of variants in *ANGPTL3* and *ANGPTL4* are discussed in the Future Work section below.

Finally, what are the clinical implications of these data for drugs to lower TG aimed at reducing CHD risk? Multiple recent randomised control trials have tested if lowering TG with fibrates or fish oils leads to a reduced risk for CHD. Many of these trials have at best produced modest ambiguous results evident in post hoc analyses of patients with elevated TG levels, or when used without concomitant statin use or, have failed to show any CVD benefit[3,42,43]. Fibrates are currently the most potent agent to lower TG levels achieving up to 50% but have adverse effects on the liver and renal function[44–46]. Pemafibrate, a novel selective PPARα (peroxisome proliferator-activated receptor alpha) modulator was effective in managing atherogenic dyslipidaemia in clinical trials, either as monotherapy or as add-on to statin therapy, reports reduction of TG, remnant cholesterol and apolipoprotein CIII, thus far has the most favourable side effects, and is currently in a phase III outcome study[47–49]. Cholesteryl ester transfer protein (CETP) which mediates the exchange of TG from VLDL or LDL and cholesteryl esters from HDL was considered a possible therapeutic approach. Despite reductions in lipid concentrations, CETP inhibitors torcetrapib, dalcetrapib and evacetrapib had no effects on clinical cardiovascular outcomes[12]. The last of these agents, anacetrapib has been shown to have a modest benefit in the Randomized EValuation of the Effects of Anacetrapib Through Lipid-modification (REVEAL) trial at the end of four years, with no significant benefit in

years one and two and is not in current use due to safety issues and mediocre efficacy[12,13].

Omega 3 fatty acids have become a recent alternative approach for TG lowering. Long chain omega-3 fatty acids reduce plasma TG primarily from the decline in hepatic VLDL TG production and secondarily from the increase in VLDL clearance[50]. A Cochrane meta-analysis in 2018 (n = 112,059) found no effect of omega 3 on CVD risk[43]. Whereas the results from the REDUCE-IT trial evaluating n-3 eicosapentaenoic acid (EPA) in high dose (4 g/day) reported 25% reductions in major adverse cardiovascular events[51,52]. Despite the impressive reduction in clinical endpoint, the CVD benefit is unlikely due to the modest absolute TG reduction of 3.50 mmol/L, and is more likely due to the pleiotropic effects of the high dose EPA on anti-inflammatory and anti-thrombotic effects, as well as postulated benefits in endothelial function and membrane stabilising[53,54]. Following on from the REDUCE-IT trial, the STRENGTH study using 3g of EPA was discontinued due to its low likelihood of demonstrating a benefit to patients with mixed dyslipidaemia who are at increased risk of CVD, further supporting the possibility that EPA cannot be used as TG lowering intervention for CVD benefit[55]. Possible reason for the failed trials may be due to incorrect study populations. Most trials exclude participants with severely elevated plasma TG and therefore excluding those most at risk of CHD who may benefit from TG lowering. Excluding individuals with severe hypertriglyceridemia may also result in an insufficient TG lowering and therefore any TG reduction may not translate into clinical benefit. It is also difficult to ascertain if failed trials are due to failure of a compound or the biomarker. Drugs specifically target a single protein in a biochemical pathway regulating the level of

TG concentrations, and so it may be difficult to distinguish if negative findings from drug trials or meta-analyses are reflective of a failure of a compound (where the solution is to develop a more effective compound against the same protein target), or of a drug target, (where the solution is to develop a drug molecule that alters the same biomarker but through a different protein target), or failure of the biomarker itself (redirect efforts to a different biomarker).

7.2.1 Promising novel therapeutics for TG lowering

Human genetics offer novel therapeutic approaches in the absence of any existing successful treatments and to overcome limitations of pharmacological target validation. An emerging approach to lower TG concentrations and lower risk of CHD is through inhibition of LPL function. Angiopoietin-like proteins 3 and 4 (ANGPLT3/4) are negative regulators of LPL and have recently emerged as novel drug targets to manage dyslipidmia[38,56]. Loss of function variants in ANGPTL3 and ANGPTL4 are associated with lower concentrations of TG, LDL-C and HDL-C, as well as lower CHD risk[4,38]. A recent study using data from six cohorts, selected genetic instruments robustly associated with TG, the downstream target for ANGPTL3 inhibition and LPL enhancement, and an instrument affecting ANGPTL4 protein function, to assess reduction of cardiovascular risk[57]. Mendelian randomisation analyses were conducted for 61,240 participants across six cohorts and assessed against outcome associations obtained from CARDIoGRAMplusC4D. The associations scaled to OR per 1-SD genetically lowered TG were; ANGPTL3 OR: 0.81 (95% CI 0.59 to 1.10), ANGPTL4 OR: 0.51 (95% CI 0.35 to 0.77) and LPL OR 0.68 (95% CI 0.56 to 0.83)[4]. There was an overlap of the three genotypes when effect estimates were compared to a recent MR analysis that used 409 SNPs as

the TG instrument (OR: 0.75, 95% CI 0.69 to 0.80), suggesting all three genotypes were associated with a similar reduction in CHD in proportion to the TG reduction[58]. The genetic findings support early phase clinical trial results of the first ANGPTL3 blocking agent monoclonal antibody evinacumab. The phase 1 trial (n=83), reports no serious adverse events, no discontinuations, and promising reductions in TG and LDL-C[57]. Similarly, in a phase II proof of concept trial of nine participants with familial hypercholesteremia, a median 47% reduction of TG and LDL-C reduction of 25% to 90% was achieved when evinacumab was added to treatment with stains, ezetimibe and PCSK9 inhibitors[59]. The findings from the same study also report associations with lower concentrations of apolipoprotein B (ApoB). Recent genetic studies have found circulating apoB may account for the associations of TG with risk of CHD. ApoB is a protein that does not appear in the circulation without lipids and therefore, it is postulated that LDL-C, remnant cholesterol and TG (all ApoB containing lipoproteins) all appear on the causal pathway to CHD. It is further suggested that lowering the number of particles of ApoB can help explain the benefit of lowering TG by perturbing LPL-mediated lipolysis to provide cardiovascular benefit in addition to cholesterol lowering by statins[58,60].

Variation in *APOC3*, the gene for Apolipoprotein C-III (apoC-III) has emerged as an important regulator of TG transport and a novel therapeutic to reduce dyslipidaemia and CVD risk. Apoc-III is a small 79-amino acid glycosylated protein component of TRL, HDL, and is detectable in LDL[61]. The distribution of apoC-III between these lipoproteins varies dependant on the fasting and postprandial state. ApoC-III inhibits lipolysis of TRL resulting in atherogenic TRL in the plasma, enhances atherogenicity of LDL by increasing affinity for arterial wall

proteoglycans, and interferes with the binding of apoB to LDL-receptor, resulting delayed catabolism of atherogenic VLDL and chylomicron remnants[61]. Carriers of a null mutation in *APOC3* have been shown to have 50% lower apoC-III concentrations, 35% lower TG, lower coronary artery calcium score (OR: 0.35), and lower 10-year Framingham CHD risk (RR= 0.68) than non-carriers[62]. These findings were confirmed in MR studies investigating the effect of loss-of-function variants in the *APOC3* gene, which reports a 40% reduction in coronary artery disease[39]. A phase I study of potent agent volanesorsen, an antisense oligonucleotide (ASO), was completed in 2013 and showed promising reductions in apoC-III (up to 78.0%) and TG concentrations (up to 43.8%) in healthy subjects. Phase II trials provided evidence for LPL-independent TG-lowering effect of apoC-III but all studies showed dose-depended injection site reactions, fatigue, musculoskeletal pains and nausea in the treatment arm compared to placebo[6]. Recent phase III trials, APPROACH (n= 46) and COMPASS (n= 113) showed a reduction in plasma TG concentrations between 70-80%[63,64]. In combined analysis of these studies, acute pancreatitis was lower in the treatment arm when compared to the placebo group. In APPROACH however, volanesorsen was discontinued in five participants due to declines in platelet counts[62]. Despite being a potent TG-lowering agent, continued use of volanesorsen is unlikely due to its unfavourable side effects[62]. Preclinical studies show the dual apoC-II mimetic and apoC-III inhibiting peptide D6PV has the ability to activate LPL and rapidly reduce the TG concentrations in genetic or diet-induced mice models by up to 85%[6]. While the results are encouraging, anticipated future trials in humans will elicit the clinical applicability of D6PV in acute pancreatitis, hypertriglyceridemia, and CVD.

7.3 Thesis strengths and weaknesses

Studies contributing to University College, London School of Hygiene and Tropical Medicine, Edinburgh and Bristol (UCLEB) consortium have provided a rich data source to test the hypotheses of this thesis. The cohorts used in this thesis are UK based with wide geographic representation increasing generalisability to the UK population. Participants are almost exclusively of European ancestry except for the SABRE cohort, which includes individuals of Afro-Caribbean and South Asian descent. This has the potential to limit generalisability of the findings presented in this thesis to non-European populations. To address this, each ethnicity in the SABRE cohort was treated as a separate cohort (European, afro-Caribbean and south Indian studies) and assessed in Chapter 5. Chapter 5 reported no discernible differences in the distribution of TG in the fourteen lipoprotein subfractions, irrespective of the ethnic group of the contributing study population. Moreover, low heterogeneity was found when evaluating between study differences prior to pooling study-specific effect estimates, further providing support to combine studies in meta-analysis.

Each of the studies is of a prospective cohort design with the same sampling frame and clinic procedures to ascertain phenotypic outcomes. Uniform procedures enable the UCLEB studies to include richly phenotyped cardiometabolic traits such as lipids and lipoproteins, demographic and anthropometric factors with limited within, and between study heterogeneity, allowing pooling of study cohorts. This enabled a large sample size of participants sufficiently powered to address the aims of this thesis.

A further strength of this study is the availability and application of NMR metabolomic lipoprotein lipid data available for the UCLEB cohorts. Previous

studies have evaluated the association of total TG concentrations and CHD. Results from such studies have produced ambiguous and equivocal results. This thesis hypothesised that TG concentrations in the different lipoprotein subfractions may have a differential association with CHD. The Nightingale metabolic biomarker platform is based on high-throughput NMR, and provides detailed quantification of lipoproteins and lipid concentration, further to what is available using standard clinical chemistry measures. Lipoprotein lipid profiling by Nightingale Health's NMR platform provides consistent, reliable, and repeatable measurements. Quantified samples undergo multiple quality assurance stages to verify sample integrity and limit contamination.

All contributing UCLEB study cohorts included in this thesis have a reliable DNA repository with published genetic analyses, making it a rich data source for the genetic analyses performed in Chapter 6. The strength of cohort-based analyses is that genetic loci can be identified for every quantitative trait recorded in sufficiently large numbers. The meta-analysis of de novo GWAS of UCLEB measures and Kettunen et al[65] allows for the identification of genetic variants for TG and cholesterol in the fourteen lipoprotein subfractions. A major advantage of the genetic epidemiological approach used in this thesis is to overcome the limitations of observational epidemiology. These include reverse causation and confounding, as genetic variants are fixed at conception. This supports causal inferences made about the effects of TG and cholesterol content in the fourteen lipoprotein subfractions on CHD.

There are limitations of the studies presented in this thesis that deserve consideration when interpreting the results. First, the age of recruitment in the

cohorts used in this thesis spans the 5th to 9th decade of life. Age was included as a continuous variable in observational analyses. The prevalence and incidence of CHD has been shown to increase with increasing age in both men and women. The American Heart Association reports that the incidence of CVD in men and women is ~40% from 40–59 years, ~75% from 60–79 years, and ~86% in those above the age of 80 years[32]. While the age range in this thesis encompasses a time at which the majority of cardiovascular disease events manifest, it is possible that in an age-stratified analysis, a differential association with CHD would have been observed.

Second, the observational results in this thesis were adjusted for variables that have been previously shown in existing literature to confound the association between TG and CHD. It is possible the observed associations may be explained by residual confounding due to factors not included in the analysis. Such factors include socioeconomic status and the influence of lipid lowering medication on TG and cholesterol concentrations. Socioeconomic status indicators including education, income, and occupation are associated with CHD risk factors. In most industrialised nations, individuals with less education, lower income, and 'blue collar' occupations have the highest CHD rates. Data on these variables were not available in the UCLEB cohort and is an important limitation to consider when interpreting the results, as inclusion of these variables may modify the association of TG and cholesterol associations with CHD. Given the age range included in this study, it is likely a substantial proportion of subjects will be taking lipid lowering medication. It is probable the association between TG and cholesterol content in the lipoprotein subfractions with CHD would have yielded smaller point estimates, were we able to account for lipid lowering medication. A further limitation is the lack of follow-up time. Had time-to CHD or stroke event data been available in UCLEB, Cox

regression analyses may have been used. In the absence of such data, it is appropriate to use logistic regression given the binary outcome of CHD and stroke.

Forth, blood samples for lipoprotein lipid NMR quantification were sourced from the UCLEB study population in a mixed fasting and non-fasting state. Historically, TG has been measured in the fasting state due to the lower biological variability of TG measurements when fasting, and the increase of TG concentrations in the post-prandial state. Most epidemiological studies have measured fasting TG to exclude the possibility of erroneous or overestimation of postprandial TG associations with CHD. It is postulated that due to varying food in-take patterns, the non-fasting state predominates the fasting state in 24-hour cycle as fasting for more than 8 hours normally only occurs before breakfast. Nordestgaard and colleagues report the maximal mean changes measured in random non-fasting versus fasting blood samples as +0.3 mmol/L TG, -0.2mmol/L total cholesterol, -0.2 mmol/L LDL-C and -0.2mmol/L non-HDL cholesterol, and do not translate to clinically significant differences, especially when evaluating associations with CVD[21]. A shift away from the longstanding tradition of using fasting to non-fasting lipid profiles is endorsed in multiple guidelines. This shift has been seen in countries including, Denmark, the United Kingdom, Europe, Canada and Brazil following the consensus view that non-fasting lipid profiles represent a simplified process for both clinicians and patients, without negative implications for prognostic or diagnostic options, for example in the case of CVD prevention[22,23]. Nonetheless, the impact of fasting status was assessed in Chapter 4 and 5. In a stratified analysis, there were similar associations of TG containing lipoprotein subfractions with CHD among fasted and non-fasted subjects from the SABRE study. Further supporting the possible shift away from using fasting measures for lipid profiling.

Triglycerides can be measured using direct and indirect methods in the clinical laboratory. Indirect estimations are calculated from the difference between serum concentrations of total fatty acids and concentration of cholesterol and phospholipid fatty acid esters[37]. Direct methods are relatively more precise and include fluorometric, colorimetric and enzymatic estimation. Similarly, methods of LDL-C measurement comprise non-direct methods including ultracentrifugation and electrophoresis, and direct methods such as chemical precipitation, immuno-separation, and homogenous assay methods[37,41]. The most common method for estimating LDL-C in clinical laboratories is using the Friedewald equation; LDL-cholesterol (mmol/L); total cholesterol minus HDL-cholesterol minus TG concentrations/2.2[42]. More recent reports since the 1990s dispute the estimation of LDL-C using the Friedewald equation as it may not be sufficiently accurate at high TG concentrations or non-fasting assay samples[43]. Moreover, the Friedewald estimation method is nonspecific to LDL-C and includes cholesterol carried in IDL and some VLDL particles[44]. In the UCLEB studies, clinical chemistry measures of TG have been measured using enzymatic estimates and LDL-C concentrations have been estimated used the Friedewald method. The variability and imprecise measurement of TG may contribute to erroneous associations with CHD.

The limitations of MR must also be appreciated. A potential cause for bias in MR is horizontal pleiotropy of genetic instruments used in analysis. Given the correlation between genetic instruments for the 14 lipoprotein subfractions for TG and cholesterol, it is likely the variants affect TG concentrations in the other subfractions, which in turn influence the association with CHD independently of the hypothesised exposure. This can result in biased MR estimates because of violation

of the exclusion restriction assumption. For example, if the genetic instrumental variable for TG concentrations in small VLDL robustly associates with TG concentrations in medium VLDL, then the MR estimate will be the combined effect of both the lipoprotein subfractions – not the effect of one lipoprotein subfraction alone and invalidates any causal inferences that may be made. Methods such as MR-Egger have been developed to explore and account for the impact of horizontal pleiotropy in MR studies. MR-Egger regression is a statistical approach that provides robust causal estimates in the presence of extreme horizontal pleiotropy. In Chapter 6, to overcome the limitations of horizontal pleiotropy and to reliably determine causal effects with greater certainty, a Rucker model selection framework was used to select between IVW and Egger methods.

7.4 Concluding comments

Elevated triglyceride concentrations have been associated with increased risk of cardiovascular disease events however, the causality of triglycerides and triglyceride-rich lipoproteins has been challenging due to associations with cardiovascular risk markers LDL-C and HDL-C. It is likely that TG are not causal and rather represent a marker for elevated cholesterol cargo of TRL, or remnant cholesterol, which are considered to be atherogenic. It is possible cholesterol in TRL accounts for the association of TG and CHD. Genetic evidence selecting variants from across the genome points to TG-mediated pathways as causal however, it can be difficult to isolate the TG only effect on disease, due to pleiotropy of genetic variants. Genetically reduced TG via variants in TG genes also have pleiotropic effects on other lipids with which they are highly correlated. Therefore, it is possible lowering of TG also results in lowering cholesterol in TRL and therefore contributes

to a reduction in CHD risk. Further work may help to provide a more thorough understanding of the genetic effects on different lipid subfractions and metabolites to fully understand the biological processes leading to atherosclerosis and CHD. Triglyceride-lowering therapies that alter TRL via the LPL pathway may prove to have efficacy in reduction of CHD.

7.5 Future work: Selection of therapeutic targets by mendelian randomisation

How do the findings presented in this thesis aid in the context of developing drugs that modify lipoprotein TG concentrations and predicting their effects on risk of CHD? Most prior MR analyses utilise multiple SNPs identified from GWAS used as instrumental variables leveraging the genetic association with TG and the genetic association with CHD, as was the case for results presented in Chapter 6 of this thesis. To answer the question of TG or TG-mediated pathway causality on CHD, SNPs are drawn from across the genome. Under certain assumptions, using this method helps to address the causal relevance of a biomarker (in this case TG) on disease. Whereas in order to ascertain whether modification of a specific gene product (i.e. a protein) will reduce CHD, an alternative method is proposed. For genetic traits, the instrumental variable must involve the TG-mediated pathway and not represent an aggregate measure of multiple genes regulating the concentration of TG that individually may have pleotropic associations with correlated lipids and CHD via non-TG mediated pathways, violating a key assumption of MR and sidestepping the exposure of interest[66,67]. In this perspective, pharmacogenomics enables the investigation of TG-specific genetic variants on the response of individuals to a TG-lowering drug. Variants in the target encoding gene (acting in *cis*) are used to evaluate the effects of modulating the same target pharmacologically

and is useful in drug target validation studies to address if modification of a protein encoded by a specific gene will result in reduction of disease outcome[66,68]. A genetic association demonstrating causality for a disease greatly increases the likelihood of success for a drug engaging the target encoded by the gene, whether a protein or proxy biomarker for protein concentration. a well studied example is that of that of 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR)[67,69]. Variants in the *HMGCR* gene are associated with lower LDL-C concentrations and reduced risk of CAD, confirming the effects of HMGCR inhibition by statins and reduced CVD in randomised trials. This approach, helps to define the mechanism based effects of intervening on the target as well as classify the on-target and off-target effects as illustrated by Plump and Davey-Smith using PCSK9 (proprotein convertase subtilisin/kexin type 9) and C-reactive protein (CRP)[67,70,71]. In the case of *PCSK9* gene, both loss and gain of function mutation has been identified to associate with circulating PCSK9 and LDL-C levels, which associate with CHD in observational epidemiological evidence. The identification of PCSK9 inhibitors contributed to the approval of two clinically verified therapies for cardio-protection, as was predicted by *PCSK9* MR studies[66,67,72]. Similarly, *CRP* genes associate with circulating CRP levels, however despite strong epidemiological evidence, CRP genetic variants do not relate to coronary artery disease risk[27]. Results from MR studies deem CRP as a predictive marker for risk but not causal in disease and therefore not a therapeutic target, as such CRP is now excluded from drug discovery efforts[73–75].

Schmidt et al[66] have developed a mathematical framework for drug-target MR and discuss the applicability of downstream biomarkers in cis-MR analyses as a valid test of a protein effect on disease. Cis-MR analysis requires the selection of

variants from within or near a druggable protein-coding gene[76,77]. This has been made possible via recent efforts to delineate the druggable genome and has culminated in the identification of 4,479 genes that comprise targets of existing therapeutics. This approach, referred to as 'drug-target MR', is progressively defined by reducing the identification of genetic instruments from across the whole genome of around 20,000 protein coding genes, to less than 5000 genes encoding druggable targets[78]. To prove utility, authors implement four loci; HMGCR, PCSK9, NPC1L1, and CETP that encode licenced or clinical phase drugs as positive controls to empirically evaluate instrument selection strategies to maximise study power and prevent erroneous significance.

The framework outlined by Schmidt and colleagues and the success of this approach in retrospectively confirming trial outcomes, make drug target MR a powerful method to validate putative TG lowering therapies. Approximately 82% of phase II and 50% of phase III trials exhibit high failure rates comprising 64% of total R&D budget[68,79]. Most drugs failing due to lack of efficacy, suggesting selection of invalid drug targets as a potential explanation. Given the failure or contentious results of TG-lowering drugs fibrates, niacin and omega-3-fatty acids, drug target MR offers an approach to confirm trial outcomes and validate therapeutic targets, prior to or alongside, the initiation of human clinical trials. This approach could be applied to test whether modifying TG via LPL-mediated emerging therapies ANGPLT3 inhibition and apoC-III translates into beneficial CHD outcomes. Moreover, the compelling association of cholesterol in TRL as the predominant causal lipid in CHD, and the evolving body of evidence implicating cholesterol in TRL and apoB containing lipoproteins in CHD represent important risk factors for

prioritisation for identification of gene products associated with TRL and apoC-III pathways as potential drug targets.

7.6 References

1. Ginsberg HN, Packard CJ, Chapman MJ, et al. Triglyceride-rich lipoproteins and their remnants: metabolic insights, role in atherosclerotic cardiovascular disease, and emerging therapeutic strategies—a consensus statement from the European Atherosclerosis Society. *Eur Heart J*. 2021;00:1-21. doi:10.1093/EURHEARTJ/EHAB551

2. Lifelong health belongs to everyone. Accessed April 14, 2020. https://nightingalehealth.com/

3. Handelsman Y, Shapiro MD. Triglycerides, atherosclerosis, and cardiovascular outcome studies: Focus on omega-3 fatty acids. *Endocr Pract*. 2017;23(1):100-112. doi:10.4158/EP161445.RA

4. Wang Q, Oliver-Williams C, Raitakari OT, et al. Metabolic profiling of angiopoietin-like protein 3 and 4 inhibition: a drug-target Mendelian randomization analysis. *Eur Heart J*. Published online December 22, 2020. doi:10.1093/eurheartj/ehaa972

5. Sacks FM, Alaupovic P, Moye LA, et al. VLDL, Apolipoproteins B, CIII, and E, and Risk of Recurrent Coronary Events in the Cholesterol and Recurrent Events (CARE) Trial. *Circulation*. 2000;102(16):1886-1892. doi:10.1161/01.CIR.102.16.1886

6. Wolska A, Lo L, Sviridov DO, et al. A dual apolipoprotein C-II mimetic-apolipoprotein C-III antagonist peptide lowers plasma triglycerides. *Sci Transl Med*. 2020;12(528):7905. doi:10.1126/scitranslmed.aaw7905

7. Talayero BG, Sacks FM. The role of triglycerides in atherosclerosis. *Curr Cardiol Rep*. 2011;13(6):544-552. doi:10.1007/s11886-011-0220-3

8. Nordestgaard BG, Varbo A. Triglycerides and cardiovascular disease. *Lancet*. 2014;384(9943):626-635. doi:10.1016/S0140-6736(14)61177-6

9. Zilversmit DB. Atherogenesis: a postprandial phenomenon. *Circulation*. 1979;60(3):473-485. doi:10.1161/01.cir.60.3.473

10. Goldstein JL, Brown MS. The LDL receptor. *Arterioscler Thromb Vasc Biol*. 2009;29(4):431-438. doi:10.1161/ATVBAHA.108.179564

11. Ballantyne CM, Olsson AG, Cook TJ, Mercuri MF, Pedersen TR, Kjekshus J.

Influence of low high-density lipoprotein cholesterol and elevated triglyceride on coronary heart disease events and response to simvastatin therapy in 4S. *Circulation*. 2001;104(25):3046-3051.

12. Group THC. Effects of Anacetrapib in Patients with Atherosclerotic Vascular Disease. *N Engl J Med*. 2017;377(13):1217-1227. doi:10.1056/NEJMoa1706444

13. Landray MJ, REVEAL Collaborative Group. Randomized Evaluation of the Effects of Anacetrapib through Lipid-modification (REVEAL)—A large-scale, randomized, placebo-controlled trial of the clinical effects of anacetrapib among people with established vascular disease: Trial design, recruitment, a. *Am Heart J*. 2017;187:182-190. doi:https://doi.org/10.1016/j.ahj.2017.02.021

14. März W, Kleber ME, Scharnagl H, et al. HDL cholesterol: reappraisal of its clinical relevance. *Clin Res Cardiol*. 2017;106(9):663-675. doi:10.1007/s00392-017-1106-1

15. Lewis GF, Rader DJ. New Insights Into the Regulation of HDL Metabolism and Reverse Cholesterol Transport. *Circ Res*. 2005;96(12):1221-1232. doi:10.1161/01.RES.0000170946.56981.5c

16. Feingold KR, Grunfeld C. *Introduction to Lipids and Lipoproteins*. MDText.com, Inc.; 2000. Accessed June 5, 2020. http://www.ncbi.nlm.nih.gov/pubmed/26247089

17. Kaess B, Fischer M, Baessler A, et al. The lipoprotein subfraction profile: heritability and identification of quantitative trait loci. *J Lipid Res*. 2008;49(4):715-723. doi:10.1194/jlr.M700338-JLR200

18. Pencina MJ, D'Agostino RB, Larson MG, Massaro JM, Vasan RS, Vasan RS. Predicting the 30-year risk of cardiovascular disease: the framingham heart study. *Circulation*. 2009;119(24):3078-3084. doi:10.1161/CIRCULATIONAHA.108.816694

19. Nordestgaard BG. Triglyceride-Rich Lipoproteins and Atherosclerotic Cardiovascular Disease: New Insights From Epidemiology, Genetics, and Biology. *Circ Res*. 2016;118(4):547-563. doi:10.1161/CIRCRESAHA.115.306249

20. Budoff M. Triglycerides and triglyceride-rich lipoproteins in the causal

pathway of cardiovascular disease. In: *American Journal of Cardiology*. Vol 118. Elsevier Inc.; 2016:138-145. doi:10.1016/j.amjcard.2016.04.004

21. Varbo A, Nordestgaard BG. Remnant Cholesterol and Triglyceride-Rich Lipoproteins in Atherosclerosis Progression and Cardiovascular Disease. *Arterioscler Thromb Vasc Biol*. 2016;36(11):2133-2135. doi:10.1161/ATVBAHA.116.308305

22. Kindel T, Lee DM, Tso P. The mechanism of the formation and secretion of chylomicrons. *Atheroscler Suppl*. 2010;11(1):11-16. doi:10.1016/j.atherosclerosissup.2010.03.003

23. Toth PP. Triglyceride-rich lipoproteins as a causal factor for cardiovascular disease. *Vasc Health Risk Manag*. 2016;12:171-183. doi:10.2147/VHRM.S104369

24. Varbo A, Nordestgaard BG. Remnant lipoproteins. *Curr Opin Lipidol*. 2017;28(4):300-307. doi:10.1097/MOL.0000000000000429

25. Timpson NJ, Lawlor DA, Harbord RM, et al. C-reactive protein and its role in metabolic syndrome: mendelian randomisation study. *Lancet*. 2005;366(9501):1954-1959.

26. Holmes M V, Ala-Korpela M, Smith GD. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. *Nat Rev Cardiol*. 2017;14(10):577-590.

27. Hingorani AD, Sofat R, Morris RW, et al. Is it important to measure or reduce C-reactive protein in people at risk of cardiovascular disease? doi:10.1093/eurheartj/ehs168

28. Do R, Willer CJ, Schmidt EM, et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nat Genet*. 2013;45(11):1345-1352. doi:10.1038/ng.2795

29. Holmes M V., Asselbergs FW, Palmer TM, et al. Mendelian randomization of blood lipids for coronary heart disease. *Eur Heart J*. 2015;36(9):539-550. doi:10.1093/eurheartj/eht571

30. Allara E, Morani G, Carter P, et al. Genetic Determinants of Lipids and Cardiovascular Disease Outcomes: A Wide-Angled Mendelian Randomization Investigation. *Circ Genomic Precis Med*. 2019;12(12):543-

551. doi:10.1161/CIRCGEN.119.002711

31. Richardson TG, Sanderson E, Palmerid TM, et al. Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: A multivariable Mendelian randomisation analysis. *PLoS Med*. 2020;17(3):e1003062. doi:10.1371/JOURNAL.PMED.1003062

32. Kathiresan S, Srivastava D. Genetics of Human Cardiovascular Disease. *Cell*. 2012;148(6):1242-1257. doi:10.1016/j.cell.2012.03.001

33. Triglyceride-mediated pathways and coronary disease: collaborative analysis of 101 studies. *Lancet*. 2010;375(9726):1634-1639. doi:10.1016/S0140-6736(10)60545-4

34. Nordestgaard B, Abildgaard S, Circulation HW-, 1997 undefined. Heterozygous lipoprotein lipase deficiency: frequency in the general population, effect on plasma lipid levels, and risk of ischemic heart disease. *Am Hear Assoc*. Accessed December 6, 2018.

35. Wittrup HH, Tybjærg-Hansen A, Abildgaard S, Steffensen R, Schnohr P, Nordestgaard BG. A common substitution (Asn291Ser) in lipoprotein lipase is associated with increased risk of ischemic heart disease. *J Clin Invest*. 1997;99(7):1606-1613.

36. Wittrup HH, Tybjærg-Hansen A, Nordestgaard BG. Lipoprotein lipase mutations, plasma lipids and lipoproteins, and risk of ischemic heart disease: a meta-analysis. *Circulation*. 1999;99(22):2901-2907.

37. Stitziel NO, Khera A V, Wang X, et al. ANGPTL3 Deficiency and Protection Against Coronary Artery Disease. *J Am Coll Cardiol*. 2017;69(16):2054-2063. doi:10.1016/j.jacc.2017.02.030

38. Dewey FE, Gusarova V, O'Dushlaine C, et al. Inactivating variants in ANGPTL4 and risk of coronary artery disease. *N Engl J Med*. 2016;374(12):1123-1133.

39. Jørgensen AB, Frikke-Schmidt R, Nordestgaard BG, Tybjærg-Hansen A. Loss-of-Function Mutations in *APOC3* and Risk of Ischemic Vascular Disease. *N Engl J Med*. 2014;371(1):32-41. doi:10.1056/NEJMoa1308027

40. Pollin TI, Damcott CM, Shen H, et al. A null mutation in human APOC3 confers a favorable plasma lipid profile and apparent cardioprotection.

Science (80-). 2008;322(5908):1702-1705.

41. Cohen JC, Stender S, Hobbs HH. APOC3, coronary disease, and complexities of Mendelian randomization. *Cell Metab*. 2014;20(3):387-389.

42. Keech A, Simes RJ, Barter P, et al. Effects of long-term fenofibrate therapy on cardiovascular events in 9795 people with type 2 diabetes mellitus (the FIELD study): randomised controlled trial. *Lancet*. 2005;366(9500):1849-1861. doi:10.1016/S0140-6736(05)67667-2

43. Abdelhamid AS, Brown TJ, Brainard JS, et al. Omega-3 fatty acids for the primary and secondary prevention of cardiovascular disease. *Cochrane Database Syst Rev*. 2020;2020(3). doi:10.1002/14651858.CD003177.pub5

44. Barter PJ, Rye K-A. Is there a role for fibrates in the management of dyslipidemia in the metabolic syndrome? *Arterioscler Thromb Vasc Biol*. 2008;28(1):39-46.

45. Jun M, Foote C, Lv J, et al. Effects of fibrates on cardiovascular outcomes: a systematic review and meta-analysis. *Lancet*. 2010;375(9729):1875-1884. doi:10.1016/S0140-6736(10)60656-3

46. Jakob T, Nordmann AJ, Schandelmaier S, Ferreira-González I, Briel M. Fibrates for primary prevention of cardiovascular disease events. *Cochrane Database Syst Rev*. 2016;11:CD009753. doi:10.1002/14651858.CD009753.pub2

47. Pradhan AD, Paynter NP, Everett BM, et al. Rationale and design of the Pemafibrate to Reduce Cardiovascular Outcomes by Reducing Triglycerides in Patients with Diabetes (PROMINENT) study. *Am Heart J*. 2018;206:80-93. doi:https://doi.org/10.1016/j.ahj.2018.09.011

48. Fruchart J-C. Pemafibrate (K-877), a novel selective peroxisome proliferator-activated receptor alpha modulator for management of atherogenic dyslipidaemia. *Cardiovasc Diabetol*. 2017;16(1):1-12.

49. Dallinga-Thie GM, Kroon J, Borén J, Chapman MJ. Triglyceride-Rich Lipoproteins and Remnants: Targets for Therapy? *Curr Cardiol Rep*. 2016;18(7). doi:10.1007/s11886-016-0745-6

50. Abdelhamid AS, Brown TJ, Brainard JS, et al. Omega-3 fatty acids for the primary and secondary prevention of cardiovascular disease. *Cochrane*

Database Syst Rev. 2018;(11). doi:10.1002/14651858.CD003177.pub4

51. Boden WE, Bhatt DL, Toth PP, Ray KK, Chapman MJ, Lüscher TF. Profound reductions in first and total cardiovascular events with icosapent ethyl in the REDUCE-IT trial: why these results usher in a new era in dyslipidaemia therapeutics. *Eur Heart J*. Published online December 23, 2019. doi:10.1093/eurheartj/ehz778

52. Linton MF, Yancey PG, Davies SS, Jerome WG (Jay), Linton EF, Vickers KC. *The Role of Lipids and Lipoproteins in Atherosclerosis*. MDText.com, Inc.; 2000. Accessed October 22, 2018. http://www.ncbi.nlm.nih.gov/pubmed/26844337

53. Kastelein JJP, Stroes ESG. FISHing for the Miracle of Eicosapentaenoic Acid. *N Engl J Med*. 2019;380(1):89-90. doi:10.1056/NEJMe1814004

54. Nurmohamed NS, Dallinga – Thie GM, Stroes ESG. Targeting apoC-III and ANGPTL3 in the treatment of hypertriglyceridemia. *Expert Rev Cardiovasc Ther*. 2020;18(6):355-361. doi:10.1080/14779072.2020.1768848

55. Nicholls SJ, Lincoff AM, Garcia M, et al. Effect of High-Dose Omega-3 Fatty Acids vs Corn Oil on Major Adverse Cardiovascular Events in Patients at High Cardiovascular Risk: The STRENGTH Randomized Clinical Trial. *JAMA*. 2020;324(22):2268-2280. doi:10.1001/jama.2020.22258

56. Graham MJ, Lee RG, Brandt TA, et al. Cardiovascular and metabolic effects of ANGPTL3 antisense oligonucleotides. *N Engl J Med*. 2017;377:222-232.

57. Dewey FE, Gusarova V, Dunbar RL, et al. Genetic and pharmacologic inactivation of ANGPTL3 and cardiovascular disease. *N Engl J Med*. 2017;377(3):211-221.

58. Richardson T, Sanderson E, Palmer T, et al. Apolipoprotein B underlies the causal relationship of circulating blood lipids with coronary heart disease. *Apolipoprotein B underlies causal Relatsh Circ blood lipids with Coron Hear Dis*. Published online August 29, 2019:19004895. doi:10.1101/19004895

59. Gaudet D, Gipe DA, Pordy R, et al. ANGPTL3 inhibition in homozygous familial hypercholesterolemia. *N Engl J Med*. 2017;377(3):296.

60. Ference BA, Kastelein JJP, Ray KK, et al. Association of Triglyceride-Lowering LPL Variants and LDL-C-Lowering LDLR Variants with Risk of

Coronary Heart Disease. *JAMA - J Am Med Assoc*. 2019;321(4):364-373. doi:10.1001/jama.2018.20045

61. Borén J, Packard CJ, Taskinen MR. The Roles of ApoC-III on the Metabolism of Triglyceride-Rich Lipoproteins in Humans. *Front Endocrinol (Lausanne)*. 2020;11:474. doi:10.3389/fendo.2020.00474

62. Taskinen MR, Packard CJ, Borén J. Emerging Evidence that ApoC-III Inhibitors Provide Novel Options to Reduce the Residual CVD. *Curr Atheroscler Rep*. 2019;21(8):1-10. doi:10.1007/s11883-019-0791-9

63. Gaudet D, Brisson D, Tremblay K, et al. Targeting APOC3 in the familial chylomicronemia syndrome. *N Engl J Med*. 2014;371(23):2200-2206.

64. Gouni-Berthold I. The role of antisense oligonucleotide therapy against apolipoprotein-CIII in hypertriglyceridemia. *Atheroscler Suppl*. 2017;30:19-27.

65. Kettunen J, Tukiainen T, Sarin A-P, et al. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat Genet*. 2012;44(3):269-276. doi:10.1038/ng.1073

66. Schmidt AF, Finan C, Gordillo-Marañón M, et al. Genetic drug target validation using Mendelian randomisation. *Nat Commun*. 2020;11(1):3255. doi:10.1038/s41467-020-16969-0

67. Plump A, Davey Smith G. Identifying and Validating New Drug Targets for Stroke and Beyond: Can Mendelian Randomization Help? *Circulation*. 2019;140(10):831-835. doi:10.1161/CIRCULATIONAHA.119.042005

68. Hingorani AD, Kuan V, Finan C, et al. Improving the odds of drug development success through human genomics: modelling study. *Sci Rep*. 2019;9(1):18911. doi:10.1038/s41598-019-54849-w

69. Bovijn J, Censin JC, Lindgren CM, Holmes M V. Commentary: Using human genetics to guide the repurposing of medicines. *Int J Epidemiol*. 2020;49(4):1140-1146. doi:10.1093/ije/dyaa015

70. Polymorphisms G. Separating the Mechanism-Based and Off-Target Actions of Cholesteryl Ester. Published online 2009.

71. Rosenson RS, Koenig W. Mendelian Randomization Analyses for Selection

of Therapeutic Targets for Cardiovascular Disease Prevention: a Note of Circumspection. *Cardiovasc Drugs Ther*. 2016;30(1):65-74. doi:10.1007/s10557-016-6642-9

72. Ference BA, Robinson JG, Brook RD, et al. Variation in PCSK9 and HMGCR and risk of cardiovascular disease and diabetes. *N Engl J Med*. 2016;375(22):2144-2153.

73. Casas JP, Shah T, Hingorani AD, Danesh J, Pepys MB. C-reactive protein and coronary heart disease: a critical review. *J Intern Med*. 2008;264(4):295-314.

74. Hingorani AD, Shah T, Casas JP, Humphries SE, Talmud PJ. C-reactive protein and coronary heart disease: predictive test or therapeutic target? *Clin Chem*. 2009;55(2):239-255.

75. Wensley F, Gao P, Burgess S. Collaboration CRPCHDG. Association between C reactive protein and coronary heart disease: Mendelian randomisation analysis based on individual participant data. *BMJ*. 2011;342:d548.

76. Hopkins AL, Groom CR. The druggable genome. *Nat Rev Drug Discov*. 2002;1(9):727-730.

77. Russ AP, Lampel S. The druggable genome: an update. *Drug Discov Today*. 2005;10(23-24):1607.

78. Finan C, Gaulton A, Kruger FA, et al. The druggable genome and support for target identification and validation in drug development. *Sci Transl Med*. 2017;9(383):eaag1166.

79. Mokry LE, Ahmad O, Forgetta V, Thanassoulis G, Richards JB. Mendelian randomisation applied to drug development in cardiovascular disease: a review. *J Med Genet*. 2015;52(2):71-79.