Metabolic Syndrome (MetS)-Associated Lipoprotein Lipase Gene (*LPL*) Single Nucleotide Polymorphisms (SNPs) and Diagnosis, Prognosis and Management of Type 2 Diabetes and Cardiovascular Conditions in Adults:

Data from 12,872 Participants from the UK Biobank

Esphie Grace Fojas

A Thesis Submitted in Partial Fulfilment of the Requirements for the

Degree

Professional Doctorate in Healthcare Science

School of Health, Education, Policing and Sciences
University of Staffordshire

August 2025

ACKNOWLEDGMENTS

I am grateful for the unwavering support, guidance and patience of my primary supervisor, Prof. Roozbeh Naemi, throughout the process of completing this thesis – your advice on research methods and statistics, remarkable proficiency and sincere commitment to share your knowledge as well as your consistent drive to encourage students do not go unnoticed. I am thankful for the notable contribution of Dr. Samina Naseeb whose expertise in genetics has been particularly effective for this study. Sincere thanks to Dr. Ahmad Haidery – your knowledge as a clinician has provided significant insights for the completion of the project. My genuine gratitude as well to the UK Biobank and DNANexus support teams for their invaluable guidance with the access and use of the data.

Thank you to all the teachers, administrators, and staff at University of Staffordshire - you have always been remarkable providers of knowledge and assistance, which made this journey much more endurable and enjoyable to see it through to completion.

To my husband, Syd, and my daughter, Prayse – you will always be the biggest inspirations for all my personal aspirations. Thank you for the love, encouragement and forbearance, particularly during moments of “near-surrender” on this endeavor. To our family and friends, specially my brother Van and Uncle Dad, Papa Ting - your prayers and support are well-appreciated.

And to my Lord Jesus Christ who Has placed this dream in my heart, thank You for Your steadfast faithfulness for helping me achieve this goal, as You always do - all glory is, and will always be, Yours.

ABSTRACT

Metabolic syndrome (MetS) is the clustering of risk factors for type 2 diabetes (T2D) and cardiovascular conditions (CVC). MetS features include abdominal obesity, hyperglycemia, and dyslipidemia. The lipoprotein lipase (LPL) gene *(LPL)* provides instructions for making the enzyme LPL, hence is crucial for lipid regulation. *LPL* polymorphisms, including single nucleotide polymorphisms (SNPs), which have been previously associated with MetS, may be of particular significance to the progression of CVC and T2D. This study aimed to investigate the molecular pathogenesis of T2D and CVC amongst individuals with polymorphisms of *LPL*. SNPs rs268, rs11542065, rs116403115, rs118204057, rs118204061, rs144466625, and rs547644955, in particular, were investigated using data from the UK Biobank. Specifically, the confirmed diagnoses of T2D and CVC in the cohort with these SNPs were assessed. In addition, this study also aimed to predict the confirmed diagnosis of T2D and CVC in the cohort. Variables associated with MetS, T2D and CVC were selected from the dataset and were analysed using SPSS. The total number of participants analysed in the cohort was 12,872 (mean age 56 years $\pm$8.1; 90.0% were of British ethnicity; 53.9% were females). Significant ($p < 0.05$) associations between all the SNPs and diagnosis of both T2D and CVC were found. Statistically significant differences in weight, BMI, diastolic BP, total lipids in lipoprotein, HbA1c, WC, HDL, and LDL were found between SNPs. BMI and WC were significantly higher in individuals who were diagnosed with both T2D and CVC; when sexes were compared, men with T2D and CVC had slightly increased BMI and WC than women. Prediction models using clinical parameters showed good AUC for predicting the diagnosis of T2D and CVC in ROC analysis (AUC = .959 for T2D, AUC = .772 for CVC). The addition of Polygenic Risk Scores (PRSs) showed diagnosis prediction improvement for both (AUC = .961 for T2D, AUC = .790 for CVC), and further

addition of SNPs showed more AUC increase (AUC = .965 for T2D, AUC = .837 for CVC). This study shows that the investigated *LPL* SNPs are associated with the diagnosis of T2D and CVC. In addition, this study demonstrates that T2D and CVC diagnoses may be predicted by clinically available factors, which may be further enhanced by incorporating associated PRSs and SNPs, including the reported *LPL* SNPs. These results can have particular implications for T2D and CVC prevention and treatment with the utilisation of stratified and personalised medicine. In this light, pharmacogenetic investigations of T2D and CVC related to these *LPL* SNPs combined with current pharmacogenomics knowledge may pave the way for improved preventive and therapeutic clinical guidelines.

Table of Contents

ABBREVIATIONS

ANOVA -	Analysis of Variance
AUC -	Area Under the Curve
BIRF -	Bioresource Research Impact Factor
BMI -	Body mass index
CAD -	Coronary artery disease
CHD -	Coronary heart disease
CIOMS-	Council for International Organizations of Medical Sciences
ClinGen-	Clinical Genome Resource
CVC -	Cardiovascular Conditions
CVD -	Cardiovascular Disease
DNA -	Deoxyribonucleic acid
DSE -	Data and sample exchange
GDPR -	General Data Protection Regulation
GIFT -	Global Initiative for the Ethical Use of Human Specimens
GWAS-	Genome-wide Association Study(-ies)
HbA1c -	Glycated haemoglobin
HDL -	High-density lipoprotein
HLPP1 -	Hyperlipoproteinemia
HTA -	Human Tissue Authority
IC -	Informed Consent
ICN -	International Clinical Trial Center Network
ID -	identification
IR -	Insulin Resistance
IRB -	Institutional Review Board
ISBER-	International Society for Biological and Environmental Repositories
ISO -	International Organization for Standardization
LDL -	Low-density lipoprotein
LIMS -	Laboratory Information Management System(s)
LPL -	Lipoprotein Lipase
LPL -	Lipoprotein lipase gene (italicised LPL)
MetS -	Metabolic Syndrome
NIH -	National Institutes of Health
NRES -	National Research Ethics Service
OECD -	Organization for Economic Cooperation and Development
PM -	Personalised Medicine
PRS(s)-	Polygenic Risk Score(s)
REC -	Research Ethics Committee
REG -	Research Ethics and Governance
ROC -	Receiver Operating Characteristic
RNA -	Ribonucleic Acid
SM -	Stratified Medicine
SNP(s)-	Single Nucleotide Polymorphism(s)
SPSS -	Statistical Package for the Social Sciences
T2D -	Type 2 diabetes
UDBHR-	Universal Declaration of Bioethics and Human Rights
UK -	United Kingdom
UKB -	UK Biobank
UNESCO-	United Nations Educational, Scientific and Cultural Organization
USA -	United States of America
WC -	Waist circumference
WMA -	World Medical Association

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Metabolic syndrome (MetS) is a major global health concern which has seen unprecedented rise in the recent decades. It is a non-communicable disease associated with the rapidly increasing obesity prevalence. MetS refers to the clustering of cardiovascular disease (CVD) and type 2 diabetes (T2D) risk factors including insulin resistance (IR), abdominal obesity, dyslipidemia, hypertension, and hyperglycemia. In the United States, it is estimated that around one third of the adult population has MetS (Saklayen, 2018). Globally, the prevalence is approximately 20-30% in adults corresponding to over a billion people affected by MetS (Grundy, 2008, Saklayen, 2018).

Although extensive investigations on individual components of MetS have been widely reported, studies on MetS as an entity are notably scarce in general, and genetic studies are limited in particular (Monda et al., 2010). The heritability of each component individually has been found to range from 16 to 60% with lipids/glucose, obesity at 44% and blood pressure at 20%. In contrast, a study conducted in Italy found the heritability of MetS to be 27% (O'Neill and O'Driscoll, 2015). Nevertheless, many studies are in agreement that IR is at the core in the pathogenesis of MetS. This is related to the role of insulin as a peptide hormone secreted by pancreatic beta cells in response to increased blood glucose levels to maintain normal metabolic state. Through its anabolic functions, insulin stimulates transport of glucose in the liver, muscles, and adipocytes (Fahed et al., 2022, Saltiel, 2021). When IR develops, the metabolic consequences include hyperglycaemia, hypertension, and dyslipidemia (Freeman et al., 2024).

Additionally, when IR develops in adipose tissues, thos impairment leads to increased levels of circulating free fatty acids (FFAs) (Boden and Shulman, 2002, Griffin et al., 1999). These FFAs promote lipogenesis and gluconeogenesis upon

acting on the liver. Increasing the negative effects, FFAs also have lipotoxic effects on pancreatic beta cells and increases triglyceride synthesis which subsequently increases the production of very low-density lipoproteins (VLDLs) (Boden and Shulman, 2002, Griffin et al., 1999, Murakami et al., 1995, Unger and Zhou, 2001). The resulting disruption of lipoprotein homeostasis highlights the potential role of lipoproteins in the development of obesity, MetS and related chronic diseases (Fahed et al., 2022).

There are two main pathways in lipid metabolism: (1) the exogenous pathway, which derives lipids from dietary sources; and (2) the endogenous pathway, which involves lipids synthesized by the liver (Lent-Schochet and Jialal, 2024). Several transfer proteins and enzymes are involved in lipid regulation, including hepatic lipase, endothelial lipase, lecithin: cholesterol acyltransferase (LCAT), and lipoprotein lipase (LPL) (Feingold, 2000). LPL is involved in both the endogenous and exogenous pathways. It breaks down triglycerides (TG), and the fat molecules are used by the body as energy or stored in fatty tissue, which may harden over time potentially leading to CVD (Kumari et al., 2021, Pirahanchi et al., 2023). Additionally, when T2D is untreated and LPL activity is subnormal, this may result in increased serum triglycerides and decreased HDL level, further contributing to the development of CVD (Kumari et al., 2021).

The lipoprotein lipase gene (*LPL*) provides instructions for making the rate-limiting enzyme LPL(Pirahanchi et al., 2023). *LPL* is located on 8p22, spans ~30 kB (kilobase), and contains 10 exons, which are the coding sections of an RNA transcript (Xie et al., 2010). Several reports indicate that *LPL* variations may cause IR changes therefore potentially resulting to obesity, MetS, and T2D (Goodarzi et al., 2004, Huang, 2009, Kumari et al., 2021, Pirahanchi et al., 2023). In *The Lancet,* Hopkins (1997) reported in his article entitled “LPL gene may shape diabetic future”

that a polymorphism of the *LPL* gene has been associated with shorter time between diagnosis of non-insulin-dependent diabetes and the development of cardiovascular diseases. Other studies which have also reported evidence that variations in the LPL gene associates with IR changes, which have repercussions on obesity and MetS but these studies were population-specific. For instance, the study by Goodarzi published in 2004 (Goodarzi et al., 2004) focused on Mexican Americans, while Huang's study in 2011 was conducted among Chinese Han youths (Huang et al., 2011).

Therefore, *LPL* polymorphisms may be of particular significance in the progression of T2D and CVD, especially in the context of MetS. These include single nucleotide polymorphisms (SNPs), which are single nucleotide substitutions at a specific genomic locus. SNPs may help elucidate the susceptibility of certain individuals to different diseases, including MetS, primarily via disease gene mapping. This process involves relational assessments between variants and disease phenotypes (Bell, 2002). Furthermore, SNPs in *LPL* have been reported to potentially have diagnostic applications for MetS (Kang et al., 2023). In the, National Institutes of Health (NIH) National Library Medicine, an online accessible source of information for genetic studies, several SNPs of *LPL* are shown to have conflicting interpretations of pathogenicity (i.e., some reports specify pathogenic for MetS or MetS-associated diseases, some are contradictory, and others report no significance or association). The SNPs in question include rs268, rs11542065, rs116403115, rs118204057, rs118204061, rs144466625, and rs547644955 (seven SNPs).

These SNPs with conflicting interpretations of pathogenicity may be of particular interest for MetS and other healthcare research as additional studies could clarify inconsistent findings in different populations. Specifically, further investigations have the potential value of enhancing variant classification systems in the long run-

for example, via the understanding of sources of disagreements (e.g. reports from commercial versus academic laboratories, different methods of interpretations used) and identification of limitations (e.g. different classification frameworks/guidelines used such as ACMPG/AMP (The American College of Medical Genetics and Genomics/Association for Molecular Pathology), and ClinGen (Clinical Genome Resource) (Nussbaum et al., 2017, Zukin et al., 2023). This could lead to suggestions of improved models for variant classifications and/or better evidence for reclassification may be provided, (Lazareva et al., 2024) further taking into consideration factors such as out-dated studies, population-specific reports, or limited evidence. This may result in improved public genomic resources in terms of consistency, accuracy, robustness, and inclusivity amongst others, which may subsequently lead to improved diagnostic accuracy and reliability of clinical decisions, particularly for patients who undergo genetic testing or being considered for personalised/precision medicine. Furthermore, studies on these SNPs in larger populations may reduce conflicting disparities and provide important additional information to help elucidate their role in the development of disease (Walsh et al., 2021). Overall, studies on SNPs with conflicting interpretations of pathogenicity are important contributors to improved public healthcare outcomes.

Using UK Biobank (UKB) data, this study aims to investigate the development of T2D and CVD amongst individuals with the MetS-associated *LPL* SNPs specifically rs268, rs11542065, rs116403115, rs118204057, rs118204061, rs144466625, and rs547644955. Additionally, the study aims to predict the incidence of T2D or CVC, as confirmed by definitive diagnosis amongst individuals in the specified cohort using logistic regression analysis.

The objectives of this study are as follows:

(1) To identify *LPL* SNPs with conflicting interpretations of pathogenicity in the online resource National Institutes of Health (NIH) National Library Medicine

(2) To detect individuals in the UKB with *LPL* SNPs rs268, rs11542065, rs116403115, rs118204057, rs118204061, rs144466625, and rs547644955

(3) To select and download variables associated with MetS, T2D, and CVC from the dataset that included: sex, age, weight, BMI, waist circumference (WC), hip circumference (HC), smoking and alcohol drinking status, physical activity, and diet variation, blood pressure (systolic and diastolic), cholesterol levels, HbA1c and glucose levels, T2D and CVC were also included, standard polygenic risk scores (PRS) for the relevant parameters in this study (PRSs for T2D, cardiovascular disease (CVD), body mass index (BMI), glycated haemoglobin, coronary artery disease (CAD), atrial fibrillation (AF), high-density lipoprotein (HDL), low-density lipoprotein (LDL), and hypertension)

(4) To analyse the data using SPSS version 29, focusing on the primary outcomes as follows:

4.A) Descriptive characteristics of the cohort

4.B) Comparison of groups in relation to T2D and CVC diagnosis

4.B.I) SNPs vs diagnosis of T2D and CVC, using chi-square test for independence

4.B.II) Zygosity vs diagnosis of T2D and CVC, using chi-square test for independence

4.B.III) SNPs vs clinical parameters, using one-way ANOVA and Post-hoc Tukey analyses

4.B.IV) SNPs vs Polygenic Risk Scores, using one-way ANOVA

4.C) Association of variables across different groups

4.C.I) Correlation between BMI and waist circumference with T2D and CVC diagnosis in individuals with these SNPS

4.C.II) Correlation between BMI and waist circumference with T2D and CVC diagnosis in individuals with the SNPS: Males and Females compared

4.C.III) Partial Correlation between waist circumference and LDL levels while controlling for age

4.D) Prediction of confirmed T2D diagnosis by logistic regression

4.D.1) Prediction of confirmed T2D diagnosis by clinical parameters

4.D.II) Prediction of confirmed T2D diagnosis by clinical parameters and T2D-associated Polygenic Risk Scores (PRS): Model D2

4.D.III) Prediction of confirmed T2D diagnosis by clinical parameters, T2D-associated PRS and SNPs: Full model for T2D

4.D.IV) Prediction of confirmed T2D diagnosis by SNPs using Model D2

4.E) Prediction of confirmed CVC diagnosis by logistic regression

4.E.1) Prediction of confirmed CVC diagnosis by clinical parameters

4.E.II) Prediction of confirmed CVC diagnosis by clinical parameters and T2D-associated Polygenic Risk Scores (PRS): Model E2

4.E.III) Prediction of confirmed CVC diagnosis by clinical parameters, CVC-associated PRS and SNPs: Full model for CVC

4.E.IV) Prediction of confirmed CVC diagnosis by SNPs using Model E2

Publication of the outputs of this research project is aimed for two publications in obesity, endocrinology, lipidology, cardiovascular, or other relevant journals. The guidelines of UKB regarding publication with the use of their data will be followed.

2. BACKGROUND
2.A. Contextual Information
2.A.I. MetS Definition and Criteria

The definition of MetS was first described by Gerald Reaven in 1988, which he first coined as Syndrome X. This was presented during his Banting lecture hosted by the American Diabetes Association and the findings were susbsequently published in the journal *Diabetes* (Reaven, 1988). Later in 2001, he published a short history of Syndrome X, wherein an explanation of 'X' was mentioned to stem from the fact that the importance of IR as a coronary heart disease (CHD) risk was relatively unknown at the time (Reaven, 2001). Subsequent articles and numerous reports have agreed that for the simplest measure, IR appears to be the key driver of metabolic disturbances in MetS (Fahed et al., 2022, Lemieux and Després, 2020, Roberts et al., 2013).

The definition of MetS varies from then based on several criteria from various health authorities, including World Health Organization, which first formalized MetS definition (WHO; 1998), European Group of Insulin Resistance (EGIR; 1999), National Cholesterol Education Program Adult Treatment Panel III (NCEP:ATPIII; 2001), American Association of Clinical Endocrinologists (AACE; 2003), International Diabetes Federation (IDF; 2005), and American Heart Association/National Heart, Lung, and Blood Institute (AHA/NHLBI; 2004) . A consensus definition was later developed, incorporating elements from the AHA/NHLBI and IDF definitions, which was established in 2009 (Huang, 2009, Kassi et al., 2011). Although the different criteria are all linked and in many aspects similar, evidently, there is a lack of universal MetS definition. The criteria for the different definitions of MetS in adults,

with comprehensive details and summary is shown in Appendix 1 (A. textual criteria, B. tabular comparison).

2.A.II. T2D and CVC: The Big Picture and Pathogenesis

Diabetes mellitus (DM) affects 1 in 11 adults worldwide and is the ninth major cause of death (Zheng et al., 2018). DM is a metabolic disease which involves atypically increased blood glucose levels due to poor glycaemic control (Sapra and Bhandari, 2024). WHO recognized DM as an epidemic, but has also attained the status of a truly global pandemic (Chobot et al., 2018, Unnikrishnan et al., 2017). There are several forms of DM such as Type 1, Type 2, maturity onset diabetes of the young (MODY), gestational diabetes, and neonatal diabetes; the most common of which is T2D accounting for 90% of adults with DM (Sapra and Bhandari, 2024). The economic burden associated with medical expenditures is equally alarming, ranging between U$140 to U$2990 per person with T2D per year (Ramzan et al., 2019). In addition, increasing prevalence of T2D is also predicted, with 415 million patients in 2015 to an estimated rise to 642 million in 2040 (Aghaei Meybodi et al., 2017).

CVD, on the other hand, is the leading cause of deaths and disability-adjusted life years (DALY) globally, accounting for one-third of all deaths worldwide (Joseph et al., 2017). Ischemic heart disease (IHD) and stroke are the leading contributors to the global CVD burden (Roth et al., 2017). In 2004, the report from the article “Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study” published in the Lancet stated that dyslipidemia was the most significant risk factor for the first occurrence of myocardial infarction (MI), more known as heart attack (Yusuf et al., 2004). However, the contributions of the primary risk factors for heart failure vary substantially across

different regions (Damasceno et al., 2012, He et al., 2001), which may signify the role of genetics and epigenetics in the pathogenesis of CVD.

Obesity is very strongly correlated with T2D and CVC as it is the most important culprit of insulin resistance, along with other closely linked comorbid conditions (Iglay et al., 2016). Microvascular and macrovascular complications of T2D (eg. diabetic retinopathy, nephropathy, neuropathy; and coronary artery disease, peripheral arterial disease, and stroke; respectively), with their accompanying financial burden, are also known to cause psychological and physical distress (Chatterjee et al., 2017, Gregg, 2017, Khanam et al., 2017). As such, T2D and CVC are undoubtedly two of the most investigated diseases in many high and middle-income countries across the globe as per report of the International Diabetes Federation (IDF) in the 2023 statement “Diabetes and cardiovascular disease.” These studies aim to elucidate its pathogenesis, risk factors, and numerous other features with the overall goal of T2D and CVC prevention and treatment.

The dysregulation of glucose metabolism in T2D is primarily due to IR and impaired insulin secretion, ultimately manifesting as uncontrolled elevations in blood glucose levels (DeFronzo et al., 2015, Lawlor et al., 2017). At the core of its development is pancreatic β-cell dysfunction which leads to glucose intolerance affecting various tissues (Kaneto, 2015). Several mechanisms have been proposed regarding the gradual decline of β-cell function, such as increased non-esterified fatty acids, inflammatory cytokines, adipokines, mitochondrial dysfunction for insulin resistance, glucotoxicity, lipotoxicity, and amyloid formation (Stumvoll et al., 2005). It is now well-established that T2D results from the interaction of the environment with the subject’s genetic makeup. Studies which support the multifactorial aetiology and nature of the disease include the thrifty phenotype hypothesis, the role of air pollution

and noise, and the effects of endocrine-disrupting chemicals also termed as obesogens or diabetogens (Kahn, 2003).

T2D, as a polygenic disorder, develops due to multiplex interplay between numerous genes and environmental factors. The precise mechanisms of how these genes interact with each other and with the environment is not completely understood. The genetic component of T2D risk appears to be the outcome of interaction of many genes across the genome, and is therefore not concentrated in one region (Ali, 2013). It is possible that the genetic component of T2D is because of multiple rare genetic variants, or a few rare genetic variants of large effect (Gibson, 2012, Queitsch et al., 2012). As with other diseases, diabetes risk genes identification is paramount to understanding the genetic components of T2D – this includes linkage studies, candidate gene studies, and genome-wide association studies (GWAS); the technology and processes involved in these are further discussed in succeeding sections. Many of the diabetes risk genes identified from these studies are active in beta cells (e.g. risk alleles at *MTNR1B*,*SLC30A8*, *THADA*, *TCF7L2*, *KCNQ1*, *CAMK1D*, *CDKAL1*, *IGF2BP2*, *HNF 1B* and *CENTD2* loci were associated with reduced beta-cell function*)* or involved in insulin secretion (e.g. risk alleles at *PPARG, FTO,* and *KLF14* loci were associated with reduced insulin sensitivity) (Voight et al., 2010). Hence, this supports the idea that beta-cell dysfunction and IR are crucial final steps in the development of T2D (Florez, 2008). Other more recent studies show that polymorphisms in the *PON1, LCAT, APOE, FTO*, and *TCF7L2* are significantly associated with T2D by an increase in free fatty acid (FFA) (Himanshu et al., 2020). Although, as mentioned, due to being a multifactorial chronic disease, the gene-environment interactions in addition to an individual's genetic component is extremely complex, and it is for this particular reason that T2D heritability and pathogenesis can not be completely

explained even at this time. Nevertheless, studies on diabetes risk genes and their contribution to the development of obesity and MetS are significant to the continuous unearthing of the underlying molecular pathogenesis of these diseases.

On the other hand, at the core of CVD pathogenesis is atherosclerosis, (Frostegård, 2013, Walden and Tomlinson, 2011), the hardening or thickening of the arteries due to plaque build up from deposits of fatty substances in the artery lining. This results in coronary artery disease (CAD), cerebrovascular disease, and peripheral vascular disease which subsequently leads to heart failure or cardiac arrhythmias (Walden and Tomlinson, 2011). Elevated low-density lipoprotein (LDL) cholesterol and triglyceride levels are established as major predictor of CAD (Talayero and Sacks, 2011). The contribution and role of lipids and lipoprotein particles have been identified in the pathology of CVD, with numerous studies supporting the need for more in-depth investigations (Soppert et al., 2020). The major risk factors for CVD have also been identified, including T2D or even slight glucose abnormalities (Chahwala and Arora, 2009).

As with T2D, linkage studies, candidate gene studies, and GWAS are used to discover genes for CVD. Family and twin studies have demonstrated the heritability of CVD, with genetic variants which predispose to CVD spanning from rare mutations to common polymorphisms (Abbate et al., 2008). Association studies have identified polymorphisms with *APOE, APOA5, and MC4R* as determinants of plasma cholesterol levels, plasma triglycerides, and body weight, respectively, among others (Vrablik et al., 2021). Another CVD-asscociated gene detected by GWAS which has been deemed as one of the most interesting genes, is *FTO*, having been confirmed previously as associated with BMI and T2D. In addition, *FTO* has also been associated with other diseases such as Alzheimer's, diabetes complications, and even as a determinant of total mortality (Vrablik et al., 2021). Considering plasma

triglycerides which has been discussed as an established predictor for CVD, a study by Johansen et al. in 2010 showed that the genetic architecture for triglycerides in the population studied comprised of large-effect variants rare in frequency, small effect variants common in frequency, and environmental factors (Johansen et al., 2010, Kathiresan and Srivastava, 2012). This was the conclusion they derived when both common and rare genetic variants explained 42% of total variation in the diagnosis of hypertriglyceridemia: clinical variables explained 20%, common genetic variants explained 21%, and rare genetic variants explained 1% (Johansen et al., 2010) . This result which appears as a mosaic, hence termed the mosaic model (Kathiresan and Srivastava, 2012), clearly show here once more the complexity and interplay of numerous influences contributing to CVD pathogenesis similar to T2D.

Various lifestyle characteristics were described to confer T2D and CVC risks, most known of which are increased caloric intake low in fibre, and little physical activity (Kolb and Martin, 2017). Greater risks were also reported with increased levels of noise and air pollution due to the activation of the hypothalamic-pituitary-adrenal (HPA) axis and the autonomic nervous and immune systems which are linked to depression (Dendup et al., 2018). Although a recognized polygenic disease, the incorporation of human genetics along with lifestyle and environmental factors have been gaining popularity in the recent decade, particularly in light of personalised medicine for T2D management (Gloyn and Drucker, 2018). The addition of this genotype approach to the currently practiced phenotype method (i.e. highly dependent on patient's clinical characteristics such as demography, comorbidities, and biological characteristics) of T2D management, is the backbone of personalised medicine (PM) for T2D diagnosis and subsequent treatment (Scheen, 2016). This is posited to complement current T2D therapeutic measures (e.g. use of anti-diabetes medications and metabolic surgery) for better, more

efficient, and cheaper T2D management in general and in the longer-term. All these aspects may also be applicable to the diagnosis, prevention and treatment for various CVC.

2.A.III. The Role of LPL and *LPL*

Lipoproteins are lipid particles containing different components with a central core made of triglycerides and cholesterol esters that transport plasma lipids (Genest et al., 1992, Lent-Schochet and Jialal, 2024). Hence, because lipids are not soluble in water, lipoproteins are needed in the circulation. There are seven classes of lipoproteins: (1) chylomicrons, (2) chylomicron remnants, (3) very low-density lipoproteins (VLDL), (4) intermediate density lipoproteins (IDL), (5) low-density lipoproteins (LDL), (6) high-density lipoproteins (HDL), and (7) lipoprotein (a) (Lp-a) (Feingold, 2000). There are two lipoprotein pathways: (1) endogenous, and (2) exogenous. The schematic on the roles of components including enzymes for both pathways are shown in Appendix 2, and the processes are summarized below, as very concisely but clearly presented by Feingold in January 2024 (Feingold, 2000):

For the exogenous lipoprotein pathway, the start is the incorporation of dietary lipids into chylomicrons in the intestine. Then in the circulation, the triglycerides carried in chylomicrons are metabolized in muscle and adipose tissue by lipoprotein lipase releasing free fatty acids, which are subsequently metabolized by muscle and adipose tissue, and chylomicron remnants are formed. Chylomicron remnants are then taken up by the liver.

For the endogenous lipoprotein pathway, it begins in the liver with the formation of VLDL. The triglycerides carried in VLDL are metabolized in muscle and adipose tissue by lipoprotein lipase releasing free fatty acids and IDL are formed. The IDL are further metabolized to LDL, which are taken up by the LDL receptor in

numerous tissues including the liver, the predominant site of uptake. Reverse cholesterol transport begins with the formation of nascent HDL by the liver and intestine. These small HDL particles can then acquire cholesterol and phospholipids that are effluxed from cells, a process mediated by ABCA1 resulting in the formation of mature HDL. Mature HDL can acquire addition cholesterol from cells via ABCG1, SR-B1, or passive diffusion. The HDL then transports the cholesterol to the liver either directly by interacting with hepatic SR-B1 or indirectly by transferring the cholesterol to VLDL or LDL, a process facilitated by CETP. Cholesterol efflux from macrophages to HDL plays an important role in protecting from the development of atherosclerosis.

Lipoprotein lipase is one of the four enzymes (other three are hepatic lipase, endothelial lipase, and lecithin: cholesterol acyltransferace (LCAT)) involved in lipoprotein metabolism (Feingold, 2000, Olivecrona, 2016). Its systematic name is triacylglycerol acylhydrolase, and it is mainly distributed in adipose, heart, and skeletal muscle tissue (Appendix 2) (Wang et al., 1992). The active site of the LPL is a Ser/Asp/His triad which is in a hydrophobic groove blocked from solvent by the lid (Mead et al., 2002). On a single lipoprotein, estimate is that up to forty LPL dimers may act at the same time, and the release of the product in the circulation is believed to be a rate-limiting step (Mead et al., 2002, Wang et al., 1992).

Some lipoproteins are risk factors for CVD and other metabolic disorders, and an impairment in lipid metabolism may cause drastic outcomes in a person's health (Lent-Schochet and Jialal, 2024). In a mice study, it has been reported that LPL caused insulin resistance and promoted obesity (Delezie et al., 2012). In humans, a study reported that a high adipose tissue LPL response to a high-carbohydrate diet may predispose toward fat gain (Ferland et al., 2012). Clearance of triacylglycerol-rich lipoproteins, a crucial step to release fatty acids for usage or storage, is believed

to be slowed down in metabolic disease and has been observed both in mice and humans (Olivecrona, 2016). In relation to glycaemic control and diabetes, LPL activity in adipose tissue and skeletal muscle is insulin-dependent, therefore LPL activity varies depending on insulin level and insulin sensitivity (Taskinen, 1987). LPL regulation also has important role in atherosclerosis, and is known to more directly impact CVD, because impaired LPL activity leads to accumulation of chylomicrons and VLDL in plasma, which results in hypertriglyceridemia (Kumari et al., 2021).

The gene that encodes LPL is the LPL gene (*LPL*), which is located 8p22, spans ~30 kB, and contains 10 exons as specified (Xie et al., 2010). It was in 1960 when LPL deficiency was discovered by Havel and Gordon (Havel and Gordon, 1960), and from then, several mutations have been detected in *LPL (Henderson et al., 1991)*. Nevertheless, anomalies or changes in *LPL,* including SNPs as discussed, may directly or indirectly impact LPL and its activity with subsequent repercussions on lipid metabolism. A schematic on the relationship among *LPL* and dyslipidemia, T2D, and CHD is presented in Appendix 3.

2.A.IV. The *LPL* SNPs in this study

The summary of information for the seven *LPL* SNPs in this study (rs268, rs11542065, rs116403115, rs118204057, rs118204061, rs144466625, and rs547644955) identified, with information copied, from the NIH Library of Medicine online for reference is presented in Appendix 4.

The details include variant type (i.e. single nucleotide variant (SNV), the specie (*homo sapiens)*, alleles (e.g. A>G specify that the reference allele is A (i.e. adenine) and the variant allele is G (i.e. guanine), chromosome, functional consequence (e.g. missense variant), and clinical significance (i.e. conflicting

interpretations of pathogenicity). Further discussion on the LPL SNPs investigated in this study are presented in the next section.

2.A.V. Advanced Molecular Techniques: SNPs and Clinical Research

Some background has already been discussed regarding SNPs. This section will expound further and provide additional details, discuss SNP types, SNP frequency, current SNP detection methods, applications and importance including that for clinical research, and databases.

Going back to what SNPs are, a note worth mentioning and discussing here is the difference between SNPs and mutations. Variants in the sequence of human DNA and proteins are identified as either a mutation or polymorphism. If the frequency of the variation in the DNA sequence in a population is 1% or higher, it is called a polymorphism (i.e. SNP, if variation is in a single nucleotide, the most common polymorphism); otherwise, it is termed a mutation (Karki et al., 2015). The latter may be more known globally to the public given adverts on mutants particularly in scientific fiction stories in printed materials (e.g. comic books), the television and movies. However, in healthcare research, both mutations and SNPs are widely investigated. Polymorphisms being more common in the population suggest that it is naturally occurring, and their effects are investigated in predisposition studies to certain diseases such as what has been explored in this research.

Going back further to the basics of molecular biology, almost every cell in the human body is nucleated, thereby containing deoxyribonucleic acid or DNA, the genetic information-carrier molecule. The DNA is composed of two chains that coil as a double helix, which are hydrogen-bonded via base-pairing rules (i.e. adenine (A) with thymine (T), and cytosine (C) with guanine (G)). Alleles, or also called allelomorphs, are variants of a nucleotide sequence at a locus in a DNA molecule.

SNVs, or also called SNPs when found in at least 1% of the population as described, differ at a single position – for instance, the seven SNPs investigated in this study. These variations have functional consequences, pertaining to the effect of the variation. The functional consequence of a missense variant alteration pertains to production of an amino acid that is different from the usual amino acid produced in the same position. The intron variant is a non-coding section. SNPs in the LPL gene may impact LPL enzyme activity with varying degree of effect from none/negligible/minimal to drastic changes in the reduction or increase of enzyme activity, subsequently affecting individuals' phenotypes (Perera et al., 2025). The functional and biochemical consequences of the *LPL* SNPs investigated in this study are presented in the following paragraphs (succeeding seven paragraphs discussing each SNP). The SNPs in investigation and their functional consequences as well as their clinical significance are shown as a summary for each SNP in Appendix 4; further information on the biochemical consequences of the SNPs can also be found in the Universal Protein Resource (UniProt), an online resource, with the accession number P06858 for LPL.

The rs268 allele is A>G (meaning adenine is replaced with guanine), with the functional consequence of a missense variant. The LPL SNP rs268 (UniProt VAR_004239), also known as p.Asn291Ser or Asn291Ser, as well as LPL N291S or N318S, is a loss-of-function variant that reduces LPL enzyme activity. This leads to a significant increase in plasma triglyceride levels due to the slower clearance of triglyceride-rich lipoproteins, and also disrupts the normal metabolism of HDL, causing a decrease in plasma HDL cholesterol levels. These changes relate to evidence, which demonstrates that the rs268 SNP is a key genetic determinant of an unfavorable lipid profile (Reymer et al., 1995a). The rs268 has been identified in UniProt to be involved in hyperlipoproteinemia (HLPP1) and as a risk factor for

familial combined hyperlipidemia-3 (FCHL3) (de Bruin et al., 1996, Morabia et al., 2003).

The rs11542065 alleles are C>G,T (comma (,) specifying either of the two bases), with the functional consequence of a missense variant. However, rs11542065 is a relatively unstudied variant compared to other LPL polymorphisms like rs268, which may be attributed to the wide discrepancy in frequency of these SNPs in various populations (i.e. rs268 being the most common among the investigated SNPs). Although research on *LPL* SNPs as rs11542065 is ongoing, they are often integrated into broader polygenic risk score models to assess overall risk for metabolic and cardiovascular conditions (Dron et al., 2019).

The rs116403115 alleles are T>C,G (comma (,) specifying either of the two bases), with the functional consequence of a missense variant. The rs118204061 allele is T>C, with the functional consequence of a missense variant. The rs144466625 allele is G>A, with the functional consequence of a missense variant. As with rs11542065, the rs116403115, rs118204061, and rs144466625 have limited reports and therefore would benefit from further investigations such as this research for additional information on their effects and clinical significance across varied populations.

The rs118204057 (UniProt VAR_004225) alleles are G>A,C (comma (,) specifying either of the two bases), with the functional consequence of a missense variant. This variant is described to result in loss of enzyme activity and its involvement in disease has been described to be in HLPP1; there are 14 publications referenced in UniProt to date, including publications from the UK and Europe (Kavazarakis et al., 2004, Mailly et al., 1997), and mixed ancestry populations (Monsalve et al., 1990).

The variant type for rs547644955 is DELINS (or Deletion and Insertion) with alleles T>-,TT ('-' specifying deletion). The functional consequence of the SNP rs547644955 is specified as an intron variant. There are very limited reports and no publications (National Library of Medicine) on this SNP, and an additional factor contributing to the need for subsequent significant study is it being an intron variant. Introns, being non-coding sections of the DNA or RNA, present unique challenges in determining pathogenicity primarily due to their location and regulatory complexity (Li et al., 2017). Unpredictable effects on splicing (eg. pre-mRNA splicing by modifying splice donor/acceptor sites that are also difficult to predict computationally) (Anna and Monika, 2018) and lack of functional assays (eg. lacking or non-standardized) may also be considered as contributory factors (Cooper, 2010). In addition, many intronic variants are under-represented in clinical databases or available population frequency may be insufficient, therefore many such variants remain classified with uncertain significance (Landrum et al., 2018). As discussed, the collective clinical significance is conflicting interpretations of pathogenicity for all these SNPs. In the online resource (this appears similar to how it appears in Appendix 4), clicking on the specific SNP would lead to further information. The clinical significance is presented in the "Clinical Significance" tab, which shows the ClinVar Accession, Disease Names, and Clinical Significance. When the Clinical Significance varies or is conflicting in nature (e.g. for rs268: pathogenic, risk factor, uncertain significance, conflicting interpretations of pathogenicity, and benign are listed), its clinical significance is listed as "conflicting interpretations of pathogenicity" as a whole. This is also related to the results of the publications (listed in the "Publications" tab) on the specific SNPs.

For rs268, for example, the first publication listed in the National Library of Medicine was in 1995, which concluded that a defective LPL is at least one of the

contributing factors to the familial combined hyperlipidemia phenotype (Reymer et al., 1995b). There are around 30 publications on this SNP to date on this list up to 2022. One of the largest earlier studies in 2008, the HuGE association and meta-analysis, which had a total of over 70,000 CHD cases and controls, reported that there was only a modestly adverse lipid profiles for carriers of rs268; but was also in agreement on the need for further investigation of this genotype on CHD risk (Sagoo et al., 2008). In 2010, a study by Ariza et al in Spain, investigated the additive effect of LPL variant with other genetic variants which have effects in triglyceride (TG) metabolism such as APOA5 and APOE, and found that rs268 has a significant independent additive effect on TG levels (Ariza et al., 2010).

In Section 2.A.III, the endogenous (synthesis within the body, i.e, the liver) and exogenous (from food or dietary sources) lipoprotein pathways have been discussed. Figure A shows a flow chart when there is an irregularity with LPL function; this may be attributed to an aberration of LPL gene (eg. SNP or mutation), causing dysregulation of lipid metabolism and homeostasis which may lead to dyslipidemia, affecting cascades of imbalance in bodily mechanisms as fatty acids play a major role in heart, muscle, and adipose tissue metabolism.

The association of LPL and dyslipidemia have long been established and has been found to be regardless of ethnicity (Havel and Gordon, 1960, Henderson et al., 1991, Liu et al., 2004). As an enzyme, when LPL hydrolyses triglycerides, non-festered fatty acids (NEFA) and 2-monoacylglycerols are provided for many tissues- in the adipose tissue, NEFA is stored as triacylglycerol (TAG) via re-esterification; while in the muscles, NEFA is the major energy source, suggesting that LPL gene is the candidate gene for dyslipidemia (Mead and Ramji, 2002, Merkel et al., 2002, Seip and Semenkovich, 1998). Further, some studies show that LPL deficiency leads to hypertriglyceridemia (Feoli-Fonseca et al., 1998, Liu et al., 2004), and several

other studies in different populations support that *LPL* variants associate with varying lipid levels such as the study in French subjects - Jemaa et al. (Jemaa et al., 1995), in Dutch population - Groenemeijer et al. (Groenemeijer et al., 1997), and in Japanese individuals - King et al. (King et al., 1998) and Yamada et al. (Yamada et al., 2007), and amongst Chinese - Yang et al. (Yang et al., 2005). In addition target therapy studies in mice by Ross et al. (Ross et al., 2004) and anti-LPL autoantibody investigations by Kodera et al. (Kodera et al., 2005) support the role of LPL gene aberrations in the pathogenesis of dyslipidemia.

Dyslipidemia may then lead to insulin resistance and/or pancreatic Beta cell apoptosis, which may result to T2D. Hypertriglyceridemia, being characteristic of dyslipidemia, prioritizes TAG utilization inhibiting the intake and oxidation of glucose (Ferreira et al., 2001). Moreover, fatty acid metabolites in the cell interfere with the cascade of insulin signalling (Pulawa and Eckel, 2002), and Beta-cell function may be impaired which can lead to apoptosis when more free fatty acids are delivered to pancreatic Beta cells (Cruz et al., 2001). All these are suggested contributing factors to the direct association of LPL aberration to the development of T2D, which may be considered for all the SNPs included in the study, given that all have shown association to the development of T2D. Other studies also report that some *LPL* SNPs have further effects in diabetes complications, such as the study of Wu et al., which suggested that investigated *LPL* SNPs conferred susceptibility to diabetic kidney disease and rapid loss of renal function (Wu et al., 2023). It is important to note here and again, however, that the outcome is dependent on the *LPL* impairment, considering type or degree. For instance, a meta-analysis study by Liu et al. in 2020 has showed that certain alleles in the LPL gene were associated with lower risk of T2D, although this was still attributed to change in lipid levels (Liu et al., 2021).

Essential hypertension (EH), as with the other related chronic diseases mentioned in this study, has genetic, environmental, and epigenetic causes. And as with T2D, abnormalities in lipid metabolism and insulin resistance are suggested as major causes in EH development, particularly hemodynamics due to increased TG levels, sodium reabsorption, retention, and vascular hypertrophy (Jemaa et al., 1995); as well as the role of LPL in arterial stiffness regulation (Yang et al., 2004). This is supported by research including linkage analysis and disequilibrium studies in Chinese populations showing LPL gene markers being associated with systolic blood pressure (SBP) and diastolic blood pressure (DBP) (Tu et al., 2005, Yang et al., 2004, Yang et al., 2003). Although earlier studies did not see similar results in Caucasians (Hunt et al., 1999), a more recent meta-analysis generalized association between the LPL gene S447X and hypertension; however, the association was found to be stronger in Asians (Wang et al., 2017), supporting that *LPL* may need race-specific investigations in terms of the development of hypertension.

There are several risk factors already discussed in this chapter that have established association with the pathogenesis of coronary heart disease (CHD) and other cardiovascular diseases, such as insulin resistance, hypertension, and even T2D. Atherosclerosis, as also previously discussed, as well as thromboembolism, is a known effect of dyslipidaemia. Thromboembolism arises when a localized blood clot (thrombus) breaks off from a site, travels through the bloodstream and obstructs blood vessels eventually causing ischemia (lack of oxygen) and organ damage (Gollamudi et al., 2022). Atherosclerosis and thromboembolism can both cause CHD / CVD. But in the grounds of *LPL* aberration alone, several studies have reported direct linkage between *LPL* SNPs and aetiology of CVD (Bos et al., 2004, Lamarche et al., 1997, Rip et al., 2006, Zee et al., 2006). As to the underlying complex

mechanism, some causes and cascades include lipid-saturated macrophages penetrating the endothelium cells in the vascular wall middle layer given an overexpression of LPL (Babaev et al., 2000), and proliferation of smooth muscle cells due to LPL-triggered pro-pathogenic events (Mamputu et al., 2000). In a mice study by Wilson, et al. (Wilson et al., 2001), LPL deficiency in the macrophage showed a decrease in diet-induced atherosclerosis, while this was accelerated in apoE-deficient mice via expression of human LPL in the macrophage specifically, suggesting tissue-specific effects of *LPL* variations.

Some definitions of SNPs require that the subsitution of a single nucleotide at a specific position in the genome should be present in a large population (e.g. at least 1%) (Sherry et al., 1999); howevever, many publications do not apply this cut-off (Auton et al., 2015, Lander et al., 2001, Sherry et al., 1999). SNPs may be within coding, non-coding, or intergenic (i.e. between genes) regions. SNPs in coding regions may be synonymous substitutions (i.e. do not result in amino acid change), or non-synonymous substitutions; the latter may be missense (single change in the base results to amino acid change, resulting in disease), or nonsense (results in premature stop codon) (Auton et al., 2015, Cordovado et al., 2012). In the global population, according to MedlinePlus, over 600 million SNPs have been identified.

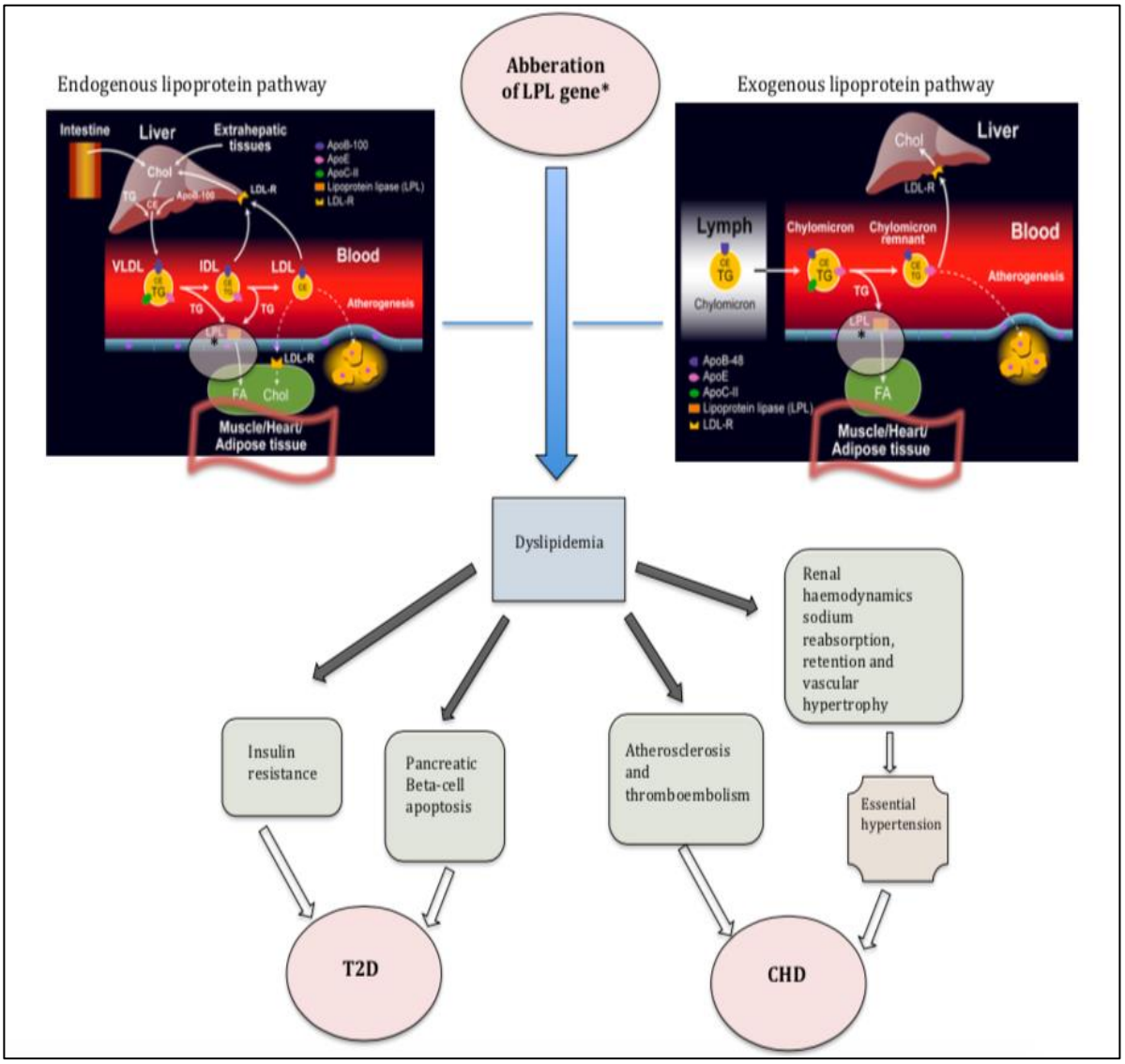


Figure A. The effects of aberration of LPL gene resulting in T2D and CVC (Feingold, 2000, Xie et al., 2010): Diagram showing the effects of aberration of LPL gene in the lipoprotein pathways (endogenous and exogenous) resulting in dyslipidemia and subsequently T2D and CHD, the most common CVD which may be taken as a surrogate marker for all CVD/CVC. *LDL-(R) = low-density lipoprotein (receptor), Chol = cholesterol, VLDL = very low-density lipoprotein, IDL = intermediate density lipoproteins, TG = triglycerides, CE = cholesterol esters, FA = fatty acids, T2D = type 2 diabetes, CHD = coronary heart disease, CVD = cardiovascular disease, CVC = cardiovascular conditions.*

There are several SNP analysis or detection techniques available, including DNA sequencing, mass spectrometry, capillary electrophoresis, denaturing HPLC and gel electrophoresis, hybridization analysis, and polymerase chain reaction (known as PCR) followed by gel electrophoresis (Ye et al., 2001). DNA amplication via PCR, in fact, is often used for many of these techniques (Tu et al., 2018); however, this is often costly and time-consuming. Because of this, alternative

techniques are being explored such as the method developed by Xia, et al in 2021 which involved the use of biosensor and fluorescence (Xia et al., 2021).

In terms of applications, the utility of SNPs for GWAS and candidate gene association studies have been briefly discussed among others. GWAS require genome-wide genetic data maybe generated by several technology, such as whole genome sequencing (WGS) which provides the most comprehensive genetic data in an organism at a single time, and whole exome sequencing (WES) which sequences the protein-coding regions of genes in a genome (Schwarze et al., 2018). A new technology also increasing in popularity is next-generation sequencing (NGS) which allows for massive parallel sequencing of DNA or RNA sequences or whole genome in a short period of time; NGS involves several steps as DNA/RNA fragmentation, library preparation, sequencing, bioinformatics analysis and data interpretation (Qin, 2019). Another important application of SNPs is in pharmacogenetics, which is discussed in detail in the next sections, particularly for T2D and CVC medications.

Bioinformatics databases, which are mostly online and easily accessible, provide valuable resources for studying SNPs. Notable example include *dbSNP* (used for this study in identifying the SNPs with conflicting interpretations of pathogenicity for *LPL*), as well as the OMIM database, Kaviar, dbSAP, SNPedia, International HapMap Project, and GWAS Central.

2.A.VI. T2D and CVC: Current Treatment Strategies

The first line of prevention for T2D and CVC are lifestyle modifications. Because it is generally believed that energy-dense diet combined with sedentary lifestyle are the primary cause of T2D, it follows that modifying these factors may reverse T2D (Kolb and Martin, 2017). One of the current hard evidence supporting the intensive weight management approach through caloric restriction for diabetes

remission is the Diabetes Remission Clinical Trial (DiRECT). The intervention comprised withdrawal of antidiabetic and antihypertensive drugs, total diet replacement (825–853 kcal/day formula diet for 3–5 months), followed by stepped food reintroduction (2–8 weeks), and structured support for long-term weight loss maintenance. Results showed that almost half of the participants achieved remission to a non-diabetic state and off antidiabetic drugs at 12 months (Lean et al., 2018). Structured exercise intervention programs have also been shown to be effective for IR. Controlled trials using continuous glucose measurements suggest that exercise has beneficial effects on insulin sensitivity and glycaemic control (Kolb and Martin, 2017, Sampath Kumar et al., 2019).

Conventionally, when lifestyle modification measures are deemed unsuccessful, pharmacotherapy is used. Pharmacotherapy, also known as pharmacological therapy or drug treatment, is universally known as the use of one or more pharmaceutical drugs to improve symptoms, treat conditions or prevent diseases. MetS treatment and therapies, on the other hand, are primarily targeted on only one metabolic trait (for example hyperglycaemia, hyperlipedimia, or hypertension). As such, MetS medication categories include antidiabetics (metformin, thiazolidinediones, SGLT2 inhibitors, glucagon-like peptide-1 agonist), lipid-lowering agents (statins and non-statins), ACE inhibitors, ARBs, and antiplatelet agents (Nguyen et al., 2017). These therapeutic options have their distinct benefits and disadvantages, mechanisms of action, and predicted outcomes. Suitability of the patients are evaluated by physicians, particularly endocrinologists, diabetologists, or cardiologists depending on the patient's clinical characteristics, comorbidities, contraindications, and several other related factors (Aghaei Meybodi et al., 2017, Elk and Iwuchukwu, 2017, Gloyn and Drucker, 2018, Srinivasan et al., 2018). There are also combinations of these drugs, primarily antihypertensives and lipid modifying

drugs, known as the polypill which is gaining increasing recognition and use (Rosolová, 2017). The combination of drugs in the form of the polypill may be a suitable solution for preventing both T2D and CVD particularly for patients with MetS (Rosolová, 2017).

When lifestyle modification and/or pharmacotherapy fail, bariatric surgery (BS) is often the next treatment option offered to patients. BS is recognized to be the most effective treatment for obesity, with effects that go beyond weight loss as a high percentage of cases achieve remission of comorbidities, hence also known as metabolic surgery (Benaiges et al., 2015, Buchwald and Buchwald, 2019). Contemporary bariatric operations which are now deemed safe include Roux-en-Y gastric bypass, sleeve gastrectomy, adjustable gastric band and the duodenal switch; the vast majority of which are currently performed using laparoscopic technique offering rapid recovery (Nguyen and Varela, 2017). In addition, large, long-term observational studies demonstrate that bariatric/metabolic surgery is associated with reductions in all cardiovascular risk factors, actual cardiovascular events, cancer and death (Cummings and Rubino, 2018). More investigations are necessary to explore the mechanisms of glycaemic control post-bariatric surgery and the optimal surgical procedure for the treatment of obese patients with T2D and CVC (Maleckas et al., 2015).

Other unconventional, alternative approaches have also been applied for T2D and CVC management. Intake of polyunsaturated fatty acids has been shown to improve glycaemic control, particularly in Asian subjects (Coelho et al., 2017). Dietary supplements, including chromium, as well as nutritional anti-inflammatories have also been implicated as potential candidates for T2D management, particularly as adjunct remedies (Huang et al., 2018, Merone and McDermott, 2017). The effect of psychological practices such as meditation and yoga have also been evaluated,

with positive outcomes ranging from lowering of inflammatory gene expression to potential reduction in T2D complications, thereby improving quality of life and overall well-being (Lee et al., 2019, Thind et al., 2018, Varghese et al., 2018).

2.A.VII Personalised Medicine in T2D and CVC: The Role of Genetics and Genomics in Clinical Enhancing Management

Personalised Medicine (PM) is the individualization of therapies based on patient-specific attributes for better management or treatment of the disease, in contrast to the currently practised “one size fits all” or phenotype-based approach to clinical management (Aghaei Meybodi et al., 2017, Estampador and Franks, 2018). Hence, PM in T2D and CVC, or any other disease for that matter, would entail incorporating the patient’s genetic architecture to an individual’s clinical characteristics for tailored medical management. Stratified medicine (SM), on the other hand, appears intermediary between current clinical approach and personalised medicine; SM entails grouping of individuals based on disease risk or therapy response (Bell, 2014). Another term, which is often interchangeably used more for PM rather than SM is precision medicine. However, there are arguments surrounding the use of this term including the editorial published by Siest (2014) stating that medicine is not at all or could not be precise (Siest, 2014).

The application of PM in diabetes began over just a decade ago with the advent of molecular biology techniques, which have increasingly been made more available and affordable by various manufacturers (Aghaei Meybodi et al., 2017, Florez, 2017). The primary and effective first uses were on monogenic forms of diabetes (i.e. maturity-onset diabetes of the young (MODY) and neonatal diabetes) due to single gene target – in MODY, target genes included *HNF4A, GCK, HNF1A, PDX1, HNF1B,* and *NeuroD1*, while in neonatal diabetes target genes included

KCNJ1, ABCC8, IDDM2, PTF1A, and *FOXP3* (Malandrino and Smith, 2011). For T2D management, however, PM poses several challenges mainly due to the polygenic nature of T2D. No single gene region has been found for T2D underlining the complexity of laboratory and clinical investigations including pathway mechanisms and interplay with other biochemical processes, environmental contributions, and epigenetic changes (Aghaei Meybodi et al., 2017) . This complexity also applies to CVD as no single gene can explain its cause (Ho et al., 2020). Genetic testing in the context of PM for designing T2D treatment regimen is mainly categorized in three purposes or factors, namely: (1) to identify risk for disease development; (2) to identify genetic variation(s) of T2D; and (3) to predict drug response (Elk and Iwuchukwu, 2017), which may also be applicable to CVC.

Pharmacogenetics is the discipline that examines genetic variations and investigates how they affect therapeutic outcomes and incidence of adverse effects (Mannino et al., 2019). Pharmacogenomics, on the other hand, is the study of the simultaneous impact of multiple mutations in the genome which may be determinants of drug response and effects (Dere and Suto, 2009). Both pharmacogenetics and pharmacogenomics, therefore, are at the core of PM for T2D and CVC. Linkage studies, T2D candidate gene investigations, and genome-wide association studies (GWAS) have identified several T2D susceptible genes, and many investigations have established the effects of these gene variations on drug efficacy and toxicity (Aghaei Meybodi et al., 2017, Elk and Iwuchukwu, 2017). This is further explained with examples cited in Section 5.C.1. This is also true for CVD; further explanations and examples can be found in Section 5.C.II. Undoubtedly, pharmacogenetics and pharmacogenomics are highly relevant in developing pharmacoeconomically viable and relevant treatment strategies for complex multifactorial diseases such as T2D and CVC (Elk and Iwuchukwu, 2017).

Nevertheless, the use of genetics and genomics in the study of multifactorial diseases is underscored, from investigations on disease treatment with pharmacogenetics and pharmacogenomics at the core as discussed in this section, as well as studies on disease pathogenesis, prevention, and epigenetics.

Epigenetics is the study of processes involved in the alteration of gene activity without changing the DNA sequence; these modifications can be passed on to daughter cells which may or may not be reversible (Weinhold, 2006). Two of the most studied epigenetic processes are DNA methylation and chromatin modification, and current evidence suggests that SNPs are associated with these two processes (Leung et al., 2012, Weinhold, 2006). For instance, Bell et al characterized a differentially methylated region having T2D-associated SNPs near *FTO* (Bell et al., 2010). Overall, several publications report that genetic variants can modify epigenetic features, and it appears true vice versa – epigenetic variations may also mediate genetic variations (Leung et al., 2012). Genetics and genomics studies are therefore clearly significant contributors for the study of multifactorial diseases such as T2D and CVC.

2.A.VIII. The Data Accessed: UK Biobank (UKB)

In the advent of big data and data sharing, biobanks have gained increasing popularity over the last two decades, particularly for healthcare and clinical research applications (Bernasconi et al., 2020, Henderson et al., 2019, Virani and Longstaff, 2015). The main aim of biobanks is the provision of bio-samples and related data for future use in biomedical research (Langhof et al., 2018). Biobanks are often large scale resources linked to medical or public health data and are distinguished from biorepositories in that the latter are merely stored biological samples from clinical investigations (Thompson and McNamee, 2017). The global-wide interest birthed

from this revolutionary idea with an apparent advantage of huge data set formation, which may address various scientific hypotheses along with growing breakthroughs in genetics, led to considerable ventures among many governments and private industries (Caulfield and Murdoch, 2017, Thompson and McNamee, 2017). Given that biobanks are huge entities with organizational structures, systems, and business plans in place, several stakeholders are involved including prominent research bodies such as ethical and biomedical experts, with several structural and operational differences compared to traditional research (Henderson et al., 2019, Verlinden et al., 2016).

Informed consent (IC), a universally recognized requirement in most, if not all, forms of research, is perhaps the most controversial in biobanks mainly due to samples and data re-use, raising questions on the need for re-consent thereby re-contact of participants (Goisauf et al., 2019). The gravity of its impact in biobanks with regard to probable lawsuit has been exemplified in an occurrence in Texas, USA when over 5 million blood samples from newborn babies have been destroyed when five parents sued the company for failure to obtain consent (Caulfield and Murdoch, 2017). Other issues surface around age-related concerns, such as involvement of pediatric or adolescent populations (McGregor and Ott, 2019). Evidence also suggests that the trust of the participants to the inviting organization is fundamental for increased participation (Broekstra et al., 2019). Critical assessment of biobanks and the UKB framework is presented in Section III.7.

The contents of this next paragraph are derived from the UK Biobank Ethics and Governance Framework Version 3.0 (October 2007). The UK Biobank is a global healthcare research resource resulting from longitudinal personal, medical, and biological data collected from approximately 500,000 UK residents, aged 40-69 years old. Its governing and funding bodies include UK Biobank Limited, Board of

Directors, the University of Manchester as the Coordinating Centre, six Regional Collaborating Centres composing the Steering Committee involved in scientific design of the resource, and an independent Ethics and Governance Council.

The framework is subdivided into four primary parts: (1) Relationship with participants (describes recruitment, understandings and consents, and confidentiality; (2) relationship with research users (stewardship of data and samples, and research access to data and samples); (3) relationship with society (management and accountability, external governance, benefit sharing, transfer of assets or closure), and (4) adoption, implementation and revision.

2.B. Systematic Review of MetS Association Studies
2.B.I. MetS Genome-wide Association Studies

Several MetS GWASs, which are easily accessible online via GWAS Central, as well as many other MetS association studies particularly candidate gene association studies, have been published. The identification and critical analysis of reports on all SNPs from all MetS association studies worldwide may be important for further research, such as selecting participants at higher risk for development of MetS complications. In addition, systematic suggestions may be offered, which may further shed light to the study of MetS in general, and MetS genetics in particular.

In GWAS Central, reports within the decade preceding the Covid pandemic (i.e. 2010-2020; 2020 included given approximately 14 months from research project development to publication (Tumin et al., 2022)), twelve studies (Kong and Cho, 2019, Kraja et al., 2011, Kristiansson et al., 2012, Lee et al., 2018, Lin et al., 2017, Lind, 2019, Moon et al., 2018, Oh et al., 2020, Shim et al., 2014, Willems et al., 2020, Zabaneh and Balding, 2010, Zhu et al., 2017) were found on record for MetS GWAS with MetS as at least a binary trait (i.e. at least two abnormal measurements or levels of glucose, BMI or waist circumference, triglycerides, and blood pressure)

(Kong and Cho, 2019, Kraja et al., 2011, Kristiansson et al., 2012, Lee et al., 2018, Lin et al., 2017, Lind, 2019, Moon et al., 2018, Oh et al., 2020, Shim et al., 2014, Willems et al., 2020, Zabaneh and Balding, 2010, Zhu et al., 2017). These are presented in Appendix 5.

Korea had the most number of articles included with five publications, followed by the United States (US) with two reports, and one publication each from the United Kingdom (UK), Finland, Taiwan, China, and Sweden. The publications from Korea were produced from different institutions, groups, or authors, which may indicate that Korea is making significant progress in genetics studies on MetS. Between zero and eighty novel SNPs were reported to be associated with MetS as at least a binary trait. The publication from the UK that utilised the UK Biobank data reported 80 novel SNPs, which was the highest number reported (Lind, 2019). No novel SNP for compound MetS phenotype was reported in the publication from the UK in 2010 (Zabaneh and Balding, 2010). Notably, all twelve reports were open access. The majority (five reports; 42%) of the articles were published in genetics journals.

2.B.II. MetS Candidate Gene Association Studies

For candidate gene association studies, a systematic search in PubMed was performed in February 2023-April 2023. Key concepts were: (1) Metabolic syndrome (MetS) and (2) Single nucleotide polymorphism (SNP).

Controlled vocabulary terms or subject terms were: Metabolic Syndrome, Cardiometabolic Syndrome, Dysmetabolic Syndrome X, Insulin Resistance Syndrome X, Metabolic Cardiovascular Syndrome, Metabolic Syndrome X, Metabolic X Syndrome, and Reaven Syndrome X.

The protocol used for literature search on Metabolic Syndrome (MetS) and associated Single Nucleotide Polymorphisms (SNPs) was as follows:

(“Metabolic Syndrome”[Title/Abstract] OR “MetS”[Title/Abstract] OR “Metabolic Syndrome”[MeSH] OR “Cardiometabolic Syndrome”[MeSH] OR “Dysmetabolic Syndrome X”[MeSH] OR “Insulin Resistance Syndrome X” [MeSH] OR “Metabolic Cardiovascular Syndrome”[MeSH] OR “Metabolic Syndrome X”[Mesh]) OR “Metabolic X Syndrome” [MeSH] OR “Reaven Syndrome X”[MeSH] AND (“Single nucleotide polymorphism”[Title/Abstract] OR “SNP”[Title/Abstract] OR “SNPs”[Title/Abstract] OR “single nucleotide polymorphism”[Mesh] OR “SNPs”[Mesh]).

The result yielded a total of 886 publications. Excluding irrelevant articles (e.g. not related to MetS, for other diseases), duplications, and irretrievable articles, the total number of publications assessed was 135 studies.

Of these, numerous SNPs from various genes were identified as being associated with MetS. The oldest included study was from 2005, with rs718049 *PTPN1* SNP found to be associated with MetS (Spencer-Jones et al., 2005). Several of the most recent studies, published in 2022, reported associations with MetS including rs7895833 (*Sirt1) (Tao et al., 2022)*, rs266729 and rs3774261 (*ADIPOQ)* (Truong et al., 2022), and rs1169288, rs2464196, and rs735396 (*HNF1A)* (Dallali et al., 2022). *FTO* was one of the most studied and reported genes, with various SNPs identified as being associated with MetS (Kawajiri et al., 2012, Molina-Luque et al., 2021, Nagrani et al., 2020, Velazquez-Roman et al., 2021).

In the article of Chuluun-Erdene in 2020 (Chuluun-Erdene et al., 2020), aside from *ADIPOQ, PGC1,* and *FTO, rs285,* a SNP of *LPL* was included to be reported as associated with MetS amongst Mongolian subjects. However, *LPL* is not widely reported compared to other genes in terms of association with MetS, T2D, or CVC, although evidence (Goodarzi et al., 2004, Huang et al., 2011) and more recent studies from various populations (Alinaghian et al., 2019, Bozina et al., 2013,

Castellano-Castillo et al., 2018, Czyzewska et al., 2010, Hsu et al., 2021, Mus et al., 2019, Vishram et al., 2016) suggest that it may be of particular significance to the pathogenesis of these related disorders. The selection of this particular gene, which may subsequently be of importance for preventive and therapeutic studies to curb these highly prevalent chronic diseases, is therefore strongly supported by current publications. Focusing on this LPL gene, particularly those with conflicting interpretations of pathogenicity as mentioned, will facilitate a better understanding of its contribution, while also providing a manageable yet significant data size for analyses. Moreover, finding associations and comparisons on T2D and CVC diagnoses in individuals with these *LPL* SNPs, as well as predicting confirmed diagnoses of both conditions in this cohort, will add valuable information to the currently limited knowledge on these SNPs. The advent increasing use, and popularity of genetic and genomic studies, added with modern technology, will undoubtedly be useful tools in the deepening the understanding of MetS, T2D, and CVC, including the role of *LPL*.

3. MATERIAL AND METHODS

The UKB data was used for this study (UKB reference for Research Ethics Committee (REC) approval 16/NW/0274). The UKB Research Analysis Platform (RAP), an online platform managed by DNANexus, was accessed from October 2023-December 2023. The data of interest from the cohort was filtered and downloaded. The inclusion criteria were participants in the UKB (age: 40-69 years old) with the seven SNPs of the LPL gene reported to have conflicting interpretations of pathogenicity: rs268, rs11542065, rs116403115, rs118204057, rs118204061, rs144466625, and rs547644955; exclusion criteria included presence of cancer and other serious illness from recruitment baseline. Variables associated with MetS, T2D,

and CVC were selected from the data set that included: sex, age, weight, BMI, waist circumference (WC), hip circumference (HC), smoking and alcohol drinking status, physical activity, and diet variation (Meigs et al., 2006) were accessed for each participant. Blood pressure (systolic and diastolic), cholesterol levels, HbA1c and glucose levels, as well as diagnosis of T2D and CVC were also included. In addition, the standard polygenic risk scores (PRS) for the relevant parameters in this study (PRSs for T2D, cardiovascular disease (CVD), body mass index (BMI), glycated haemoglobin, coronary artery disease (CAD), atrial fibrillation (AF), high-density lipoprotein (HDL), low-density lipoprotein (LDL), and hypertension) were added.

CVC is defined for this study as any or combination of heart attack, angina, stroke, and high blood pressure; in the UKB data, this was collectively presented as presence of heart or vascular problems (HVP) as diagnosed by a doctor. CVC is used as a surrogate marked of cardiovascular disease (CVD) in this study. T2D was identified in participants with ICD-10 (International Classification of Diseases 10th Revision) diagnosis code E11 (code for Type 2 diabetes mellitus). Normal distribution was found on test of data normality for all continuous variables (Appendix 2). Average and standard deviation were calculated for each group for the selected parameters.

Data were analysed using SPSS ver. 29. Primary outcomes were incidence of T2D and CVC, and prediction of T2D or CVC diagnosis among individuals with the *LPL* SNPs of interest. Direct logistic regression was performed to assess the impact of a set of predictor variables on the odds that the participants have been diagnosed with T2D or CVC at the time of recruitment. A total of four models have been assessed for both T2D and CVC diagnoses as outlined in the subheadings of the following sections.

3.A. Descriptive characteristics

Test of data normality for the variables was conducted and were all found to have normal distribution (Appendix 6). Average and standard deviation were calculated for each group for the selected parameters. The total number of individuals in the UK Biobank for the seven specified SNPs was 17,386 when filtered individually, and 17,364 when filtered together, wherein the difference is attributed to participants having at least two SNPs. No participant was excluded in the study due to participation withdrawal as per UKB's notice to UKB researchers. Participants with cancer (N=1,355) and with other serious non-cancer medical condition or disability at baseline (N=3,137) were excluded in the study. The total number of participants included in the study was 12,872.

3.B. Comparison of groups in relation to T2D and CVC diagnosis

Data analyses of group comparisons on T2D and CVC incidence based on type of SNPs (i.e. seven SNP groups (i.e. one group corresponds to all participants with the same SNP) plus another, 8^{th}, group (i.e. participants with 2 heterozygous SNPs)) and zygosity (i.e. heterozygous versus homozygous versus combination of 2 hetrozygous SNPs) were performed using chi-square independent test for bivariate association analyses.

One-way Analysis of Variance (ANOVA) was used to compare differences amongst groups with the different metric-value parameters such as BMI, waist circumference, lipids, HbA1c, and blood glucose levels as well as PRS scores. Post-hoc analyses were performed.

3.C. Comparison of groups in relation to variables

Correlation and partial correlation analyses were performed on BMI and waist circumference versus T2D and CVC diagnosis, and was further differentiated to

compare results between males and females. Partial correlation was performed between LDL levels and waist circumference while controlling for age. Simple boxplots were generated using IBM SPSS Statistics version 29. All data points, including outliers, were retained in Figure 2 to provide a complete representation of data variability and to allow visual identification of participants with extreme anthropometric values, which may have clinical and epidemiological significance; outliers were excluded in Figure 3 to improve visual clarity and facilitate comparison of median and interquartile ranges between groups. This approach enhances the interpretability of the plots by allowing the central distribution of the data to be more easily visualized, which aligns with recommendations for descriptive graphics on such comparisons (Frigge et al., 1989, McGill et al., 1978).

3.D. Prediction of Confirmed T2D diagnosis

3.D.I. Prediction of confirmed T2D diagnosis by clinical parameters

The model contained 15 independent variables normally accessible in clinical settings (age, sex, weight, height, BMI, waist circumference (WC), hip circumference HC), systolic BP, diastolic BP, number of days per week of moderate physical activity, diet variation, smoking status, alcohol drinking status, random blood glucose, HbA1c).

3.D.II. Prediction of confirmed T2D diagnosis by clinical parameters and T2D-associated Polygenic Risk Scores (PRS): Model D2

The model contained a total of 18 independent T2D-relevant variables including the above-stated clinical parameters (i.e. 15 variables) plus three T2D-relevant PRSs (PRS for T2D, BMI, and glycated haemoglobin).

3.D.III. Prediction of confirmed T2D diagnosis by clinical parameters, T2D-associated PRS and SNPs: Full model for T2D

Direct logistic regression was performed with the addition of the *LPL* SNP groups as a parameter to the above model (i.e. total 19 variables).

3.D.IV. Prediction of confirmed T2D diagnosis by SNPs using Model D2

LPL SNP groups were filtered or selected as separate cases (i.e. Model D2 was used for individuals with rs268 only, and so on using other SNPs) and direct logistic was performed.

3.E. Prediction of confirmed CVC diagnosis

3.E.I. Prediction of confirmed CVC diagnosis by clinical parameters

The model contained 15 independent variables normally accessible in clinical settings, as with D.1.

3.E.II. Prediction of confirmed CVC diagnosis by clinical parameters and CVC-associated PRS: Model E2

The model contained a total of 22 independent CVC-relevant variables including the above-stated clinical parameters (i.e. 15 variables) plus seven CVC-relevant PRSs (PRS for BMI, CVD, atrial fibrillation, coronary artery disease (CAD), hypertension, LDL, HDL).

3.E.III. Prediction of confirmed CVC diagnosis by clinical parameters, CVC-associated PRS and SNPs: Full model for CVC

Direct logistic regression was performed with the addition of the *LPL* SNP groups as a parameter to the above model (i.e. total 23 variables).

3.E.IV. Prediction of confirmed CVC diagnosis by SNPs using Model B2

LPL SNP groups were filtered or selected as separate cases and direct logistic was performed using Model E2.

4. RESULTS

4.A. Descriptive characteristics

The total number of subjects analysed in the cohort was 12,872 (mean age 56 years ±8.1. 90.0% were of British ethnicity, and 53.9% were females. All 18 participants with multiple SNPs had 2 SNPs each, which were heterozygous for both SNPs (**Table 1**). A total of 111 (0.86%) participants had homozygous variations involving rs116403115, rs115426065, and rs268 (1, 1, and 109 individuals, respectively).

The mean baseline weight and BMI were 77.8 ±15.5 kg and 27.3 ±4.6 kg/m^2, respectively. There were 474 (3.7%) participants who had Type 2 diabetes (T2D), and 3,651 (28.4%) had heart or vascular problems (HVP; interchangeably called cardiovascular conditions or CVC in this study) as diagnosed by a doctor. Presence of T2D was identified with diagnosis of ICD-10 code E11 (non-insulin dependent diabetes mellitus) amongst participants, and CVC was defined by the presence of one or more of the following: high blood pressure, angina, heart attack, or stroke. Baseline characteristics are presented in **Table 2**.

Most participants (93.3%) were alcohol drinkers, and a majority (59.0%) had history of smoking on study enrolment. Most also reported engaging in moderate exercise for at least 10 minutes on 5.4 ±1.9 days per week (n=12,681). Dietary variation was reported as ‘sometimes’ by 56.9% of participants and ‘never/rarely’ by 34.8%.

Variant ID	Zygosity	n	Total
rs118204061	Heterozygous	10	10
rs144466625	Heterozygous	19	19
rs116403115	Homozygous Heterozygous	1 68	69
rs115426065	Homozygous Heterozygous	1 275	276
rs118204057	Heterozygous	311	311
rs547644955	Heterozygous	224	224
rs268	Homozygous Heterozygous	109 11,836	11,945
rs144466625; rs547644955	Hetero; hetero	1	18
rs115426065; rs547644955	Hetero; hetero	6	
rs115426065; rs268	Hetero; hetero	1	
rs118204057; rs268	Hetero; hetero	8	
rs118204057; rs547644955	Hetero; hetero	1	
rs547644955; rs268	Hetero; hetero	1	
TOTAL			12,872

Table 1. Frequency distribution per variant ID shows zygosity, number of participants, totals and overall total. Notably, there are pronounced differences in sample sizes, particularly when compared with variant ID rs268.
Hetero = heterozygous n = number of participants

	N	Mean	Std Dev
Weight (kg)	12,693	77.8	15.5
BMI (kg/m2)	12,843	27.3	4.6
Waist circumference (cm)	12,856	89.9	13.2
Systolic BP_mmHg	12,051	139.6	19.5
Diastolic BP_mmHg	12,052	82.3	10.7
Glucose(mmol/L)	11,199	5.1	1.2
Glycated haemoglobin (HbA1c) (mmol/mol)	12,224	36.0	6.7
Total Lipids in HDL (mmol/L)	3,043	2.9	0.6
Total Lipids in LDL (mmol/L)	3,043	2.5	0.6
Age diabetes diagnosed	612	51.6	11.9

Table 2. Baseline characteristics of participants, including key MetS-associated clinical parameters, N, mean, and standard deviation, providing comparative context for normative population data
Std Dev =standard deviation; N = total number of participants;
BMI = body mass index; BP = blood pressure; HDL = high-density lipoprotein; LDL= low-density lipoprotein

4.B. Comparison of groups in relation to T2D and CVC diagnosis

4.B.I. SNPs vs diagnosis of T2D and CVC

A chi-square test for independence indicated a significant association between SNPs and T2D (χ2 (7) = 64.09, p<.001), as well as between SNPs and CVC (χ2 (7) = 36.68, p <. 001). (See **Figure 1**).

4.B.II. Zygosity vs diagnosis of T2D and CVC

A chi-square test for independence indicated no significant association between zygosity and T2D (Fisher Exact Test p = .176chi-square assumptions not met); nor between zygosity and CVC (χ2 (2) = .546, p = .761).

4.B.III. SNPs vs clinical parameters

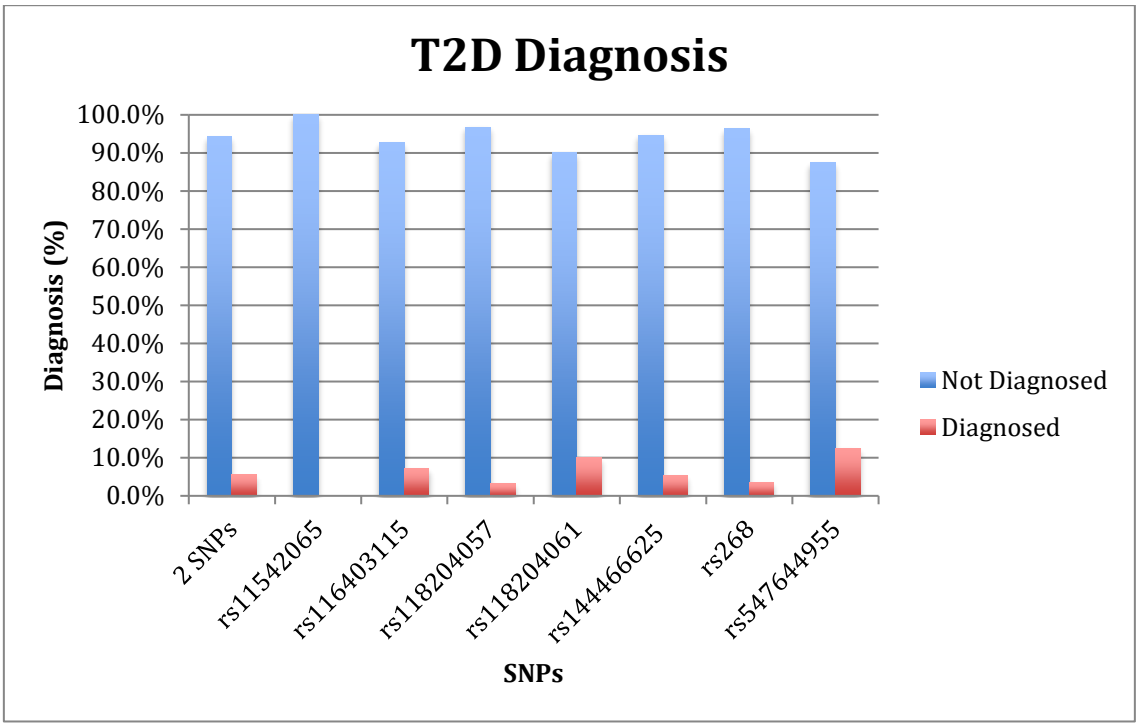
Statistically significant differences between groups were observed for weight, BMI, diastolic BP, total lipids in lipoprotein between groups, and HbA1c as determined by one-way ANOVA (p < .001). Significant differences were also found for waist circumference, HDL, and LDL (p < .05). (See **Table 3**).

Post-hoc Tukey analyses revealed significant differences between rs11542065 and rs268 as well as between rs118204057 and rs54764995, rs54764995 and rs268 (p < .001). Additional significant differences were found between rs118204057 and rs115426065, and between rs268 and rs54764995 (p < .05).

4.B.IV. SNPs vs Polygenic Risk Scores (PRS)

Significant differences were observed in all parameters tested: PRS for BMI and HDL (p <. 001); PRS for T2D, CVD, AF, CAD, HbA1c, HT, and LDL (p < .05).

A.



B.

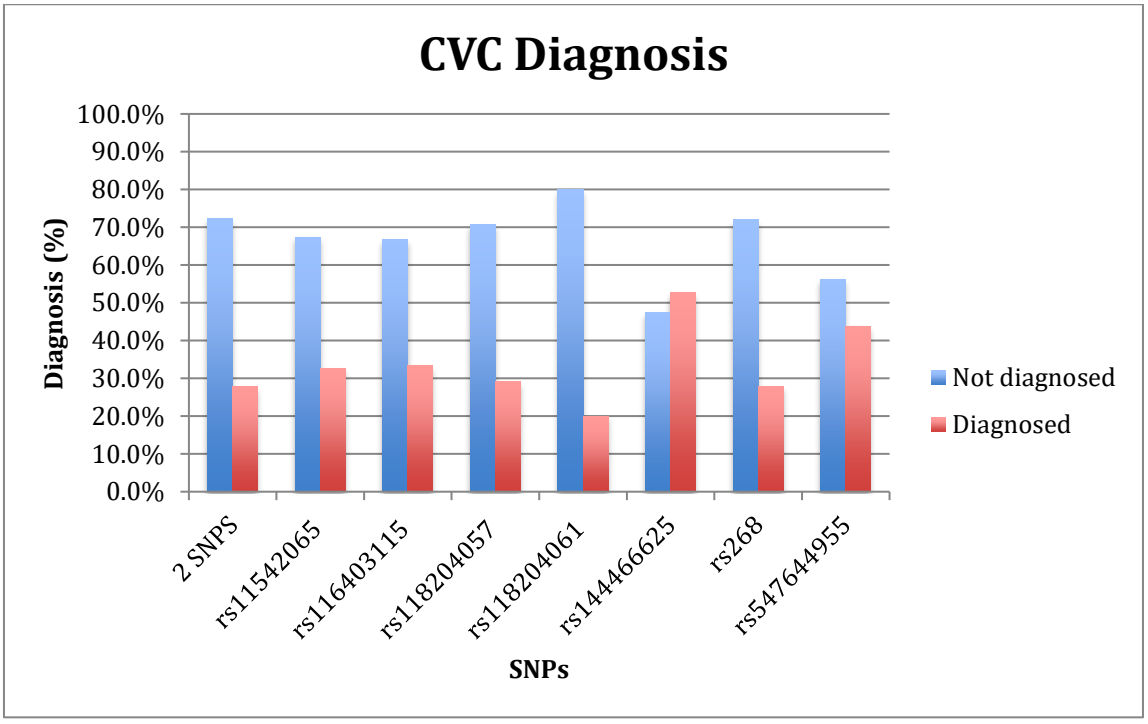


Figure 1. Diagnosis of outcomes amongst individuals with the SNPs for: A. T2D, B. CVC, showing the percentage of diagnosed versus non-diagnosed individuals. Significant associations were found between SNPs and both T2D and CVC.
T2D = Type 2 diabetes, CVC = cardiovascular conditions, 2 SNPs = presence of 2 SNPs as defined in Table 1

	SNPs	N	Mean	Std. Deviation
Weight (kg)	1 rs118204061	10	79.1	21.2
	2 rs144466625	19	80.0	18.9
	3 rs116403115	69	76.4	16.8
	4 rs115426065	271	82.1	15.8
	5 rs118204057	308	77.8	16.2
	6 rs54764995	221	81.7	16.7
	7 rs268	11,777	77.6	15.5
	8 2_SNPs	18	72.2	14.9
	Total	12,693	77.8	15.5
		ANOVA	**df** **F** **p**	**7** **5.7** **< .001**
BMI(kg/m^2)	1 rs118204061	10	26.2	5.7
	2 rs144466625	19	28.3	5.5
	3 rs116403115	69	27.6	4.6
	4 rs115426065	273	28.9	5.0
	5 rs118204057	311	27.4	5.0
	6 rs54764995	221	29.3	5.1
	7 rs268	11,922	27.2	4.5

	SNPs	N	Mean	Std. Deviation
Waist circumference (cm)	1 rs118204061	10	91.0	17.2
	2 rs144466625	19	90.9	13.6
	3 rs116403115	69	90.3	14.9
	4 rs115426065	275	92.1	12.8
	5 rs118204057	311	90.4	13.6
	6 rs54764995	223	92.6	13.0
	7 rs268	11,931	89.8	13.2
	8 2_SNPs	18	86.2	13.6
	Total	12,856	89.9	13.2
		ANOVA	**df** **F** **p**	**7** **2.8** **0.006**
Total Lipids in HDL (mmol/L)	1 rs118204061	4	2.7	0.6
	2 rs144466625	4	2.6	0.5
	3 rs116403115	14	2.9	0.7
	4 rs115426065	75	3.0	0.7
	5 rs118204057	70	2.7	0.5
	6 rs54764995	62	2.8	0.5
	7 rs268	2,810	2.9	0.6

	8 2_SNPs	18	26.9	4.6
	Total	12,843	27.3	4.6
		ANOVA	**df**	**7**
			F	**12.1**
			p	**< .001**
Diastolic BP (mmHg)	1 rs118204061	9	87.1	8.9
	2 rs144466625	18	88.2	10.3
	3 rs116403115	64	80.9	9.6
	4 rs115426065	269	84.6	10.9
	5 rs118204057	297	82.1	11.0
	6 rs54764995	219	84.6	12.4
	7 rs268	11,158	82.2	10.6
	8 2_SNPs	18	79.8	13.9
	Total	12,052	82.3	10.7
		ANOVA	**df**	**7**
			F	**4.7**
			p	**< .001**
Total Lipids in Lipoprotein Particles (mmol/L)	1 rs118204061	4	8.4	1.2
	2 rs144466625	4	7.5	1.2
	3 rs116403115	14	8.0	1.8
	4 rs115426065	75	7.9	1.6
	5 rs118204057	70	8.9	1.7

B.

	8 2_SNPs	4	2.5	0.4
	Total	3043	2.9	0.6
		ANOVA	**df**	**7**
			F	**2.1**
			p	**0.043**
Total Lipids in LDL (mmol/L)	1 rs118204061	4	2.3	0.7
	2 rs144466625	4	2.3	0.6
	3 rs116403115	14	2.3	0.7
	4 rs115426065	75	2.2	0.6
	5 rs118204057	70	2.5	0.7
	6 rs54764995	62	2.3	0.5
	7 rs268	2,810	2.5	0.6
	8 2_SNPs	4	2.4	0.9
	Total	3,043	2.5	0.6
		ANOVA	**df**	**7**
			F	**2.4**
			p	**0.019**
Systolic BP (mmHg)	1 rs118204061	9	142.2	19.8
	2 rs144466625	18	149.8	25.0
	3 rs116403115	64	135.7	18.1
	4 rs115426065	269	139.8	19.0
	5 rs118204057	297	139.3	19.9

	6 rs54764995	62	7.8	1.3
	7 rs268	2,810	8.8	1.7
	8 2_SNPs	4	8.0	2.3
	Total	**3,043**	**8.8**	**1.7**
			df	**7**
		ANOVA	**F**	**6.8**
			p	**< .001**
Glycated haemoglobin (HbA1c)	1 rs118204061	9	39.3	14.0
	2 rs144466625	18	36.3	6.4
	3 rs116403115	67	37.0	6.8
	4 rs115426065	235	39.2	11.2
	5 rs118204057	298	36.3	7.4
	6 rs54764995	176	39.9	9.6
	7 rs268	11,404	35.9	6.4
	8 2_SNPs	17	35.3	3.4
	Total	12,224	36.0	6.7
			df	**7**
		ANOVA	**F**	**17.5**
A.			**p**	**< .001**

	6 rs54764995	219	141.3	22.0
	7 rs268	11,157	139.5	19.5
	8 2_SNPs	18	136.8	24.8
	Total	12,051	139.6	19.5
			df	**7**
		ANOVA	**F**	**1.4**
			p	**0.194**
Glucose(mmol/L)	1 rs118204061	9	5.7	2.4
	2 rs144466625	17	5.1	0.5
	3 rs116403115	61	5.1	1.3
	4 rs115426065	232	5.3	1.6
	5 rs118204057	272	5.1	1.3
	6 rs54764995	199	5.1	1.2
	7 rs268	10,394	5.1	1.2
	8 2_SNPs	15	5.1	0.8
	Total	11,199	5.1	1.2
			df	**7**
		ANOVA	**F**	**0.9**
C.			**p**	**0.510**

Table 3. One-way Analysis of Variance (ANOVA) between groups showing significant difference at (A) p < .001 (weight, BMI, diastolic BP, total lipids in lipoprotein between groups, and HbA1c as determined by one-way ANOVA); and (B) at p < .05) (waist circumference, HDL, and LDL) ; C showing no significant difference amongst SNPs (numbered 1 to 8) and assessed parameters (as per clinical parameter indicated). *SNP = single nucleotide polymorphism, N = number of participants, Std. = standard, 2 SNPs = presence of 2 SNPs as per Table 1, df = degrees of freedom, F = ANOVA F-statistic indicating magnitude of difference between group means, BMI = body mass index, BP = blood pressure, HDL = high-density lipoprotein, LDL = low-density lipoprotein*

4.C. Association of variables across different groups

4.C.I. Correlation between BMI and waist circumference with T2D and CVC diagnosis in individuals with these SNPs

Amongst individuals with SNPs investigated in this study, there was a positive correlation between T2D diagnosis and both BMI (r = .180) and waist circumference (r = .201), p < .001. A positive r value indicated that higher BMI or waist circumference was associated with a greater likelihood of T2D diagnosis. The correlation coefficients fall within the “small” range according to Cohen’s criteria (0.10–0.29), suggesting that, while the associations were not strong, they were consistent and unlikely to be due to chance given the large sample size.

Similarly, a positive correlation was observed between CVC diagnosis and both BMI (r = .248) and waist circumference (r = .261), p < .001. These coefficients were slightly higher than those observed for T2D, indicating a marginally stronger association between increased adiposity measures and CVC diagnosis in this cohort. Thus, individuals with higher BMI and larger waist circumference were more likely to have a CVC diagnosis, with waist circumference showing a slightly stronger relationship than BMI for both conditions.

Overall, the results suggest that, in this SNP-defined population, central adiposity (as measured by waist circumference) may be a marginally better predictor of both T2D and CVC diagnoses than BMI, with these associations being more pronounced for CVC than for T2D.

Simple boxplots of BMI and waist circumference by diagnosis of T2D and CVC are shown in **Figure 2.**

4.C.II. Correlation between BMI and waist circumference with T2D and CVC diagnosis in individuals with the SNPs: Males and Females compared

When comparing males and females, the correlation between BMI and waist circumference was consistently positive in both sexes, indicating that higher values for these anthropometric measures were associated with a greater likelihood of diagnosis.

For T2D, the correlation coefficients for females were $r = 0.164$ (BMI) and $r = 0.178$ (waist circumference), while for males they were slightly higher at $r = 0.199$ (BMI) and $r = 0.209$ (waist circumference). These positive r values, all statistically significant, fall within the “small” range according to Cohen’s conventions (0.10–

0.29), but the slightly higher values in males suggest a marginally stronger relationship between adiposity and T2D diagnosis in men compared to women.

For CVC, the pattern was similar. Among females, the correlation coefficients were r = 0.238 (BMI) and r = 0.241 (waist circumference), while among males they were slightly higher at r = 0.252 (BMI) and r = 0.258 (waist circumference). These values approach the upper end of the “small” range and were stronger than those observed for T2D, indicating that both BMI and waist circumference have, to some extent, greater association with CVC diagnosis than with T2D in this population. Again, the higher r values in men suggest that the association between measures of adiposity and CVC is modestly stronger in males compared to females.

In summary, the consistently positive and statistically significant correlation coefficients indicate that in both sexes, higher BMI and waist circumference were related to increased likelihood of T2D and CVC diagnoses, with waist circumference showing slightly stronger associations than BMI, and these relationships being marginally more pronounced in males.

Boxplots of BMI and waist circumference by diagnosis of T2D and CVC, stratified by sex are shown in **Figure 3.**

4.C.III. Partial Correlation between waist circumference and LDL levels while controlling for age

There was no significant partial correlation between waist circumference and LDL levels while controlling for age, r = -.006, p > .05. The r value was extremely close to zero, indicating an almost complete absence of a linear relationship between waist circumference and LDL levels in this sample when the effect of age was accounted for. An inspection of the corresponding zero-order correlation coefficient (r = −0.007) showed an essentially identical value, suggesting that controlling for age did not meaningfully alter the association. The minimal difference (Δr = 0.001) indicates that age was not a confounding factor in the relationship between waist circumference and LDL levels in this dataset.

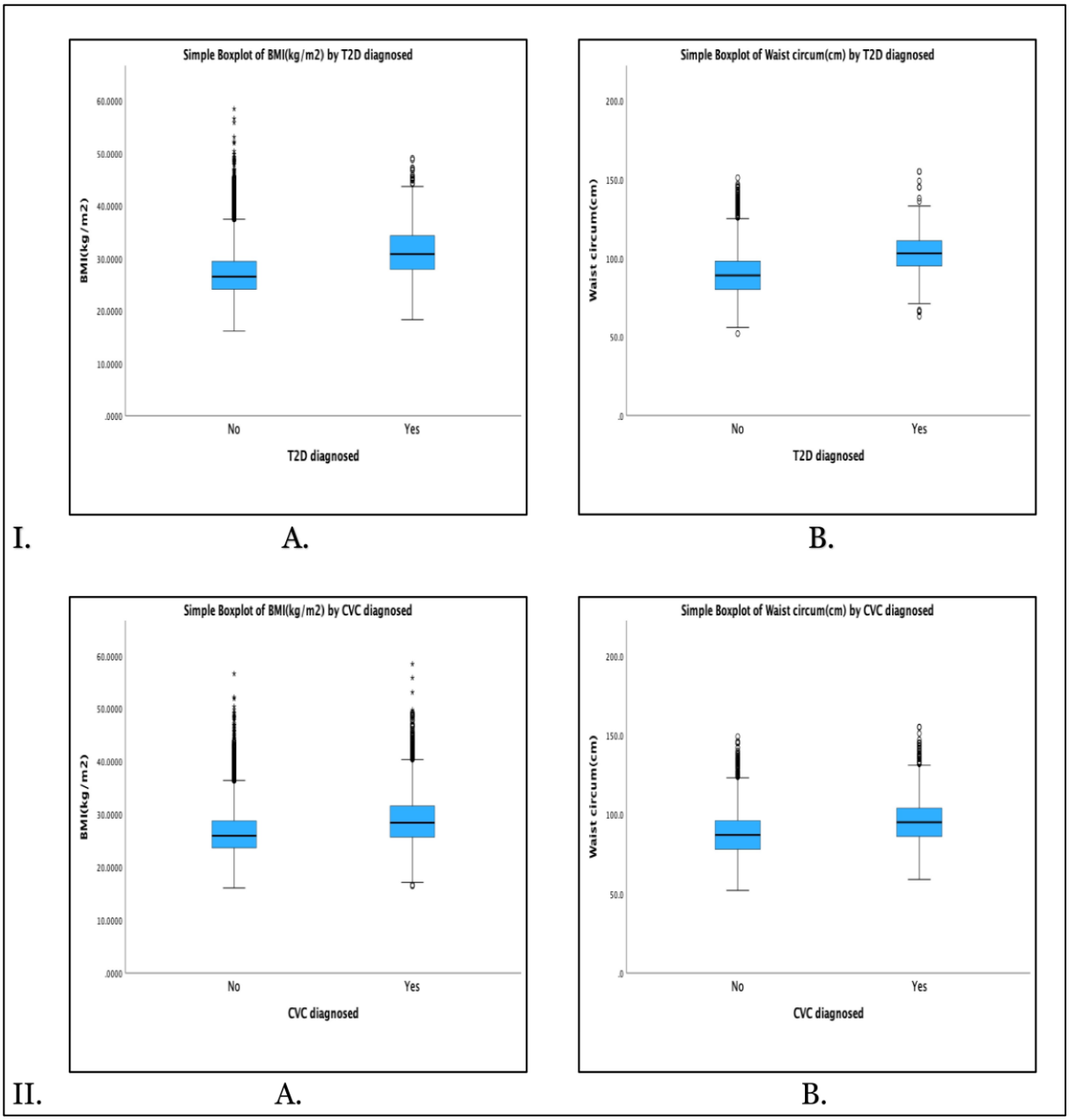


Figure 2. Simple boxplots of body mass index (BMI, kg/m²) and waist circumference (circum; cm) by diagnosis of type 2 diabetes (T2D) and cardiovascular conditions (CVC). Panels I.A and I.B show the distribution of BMI and waist circumference, respectively, for participants with and without T2D. Panels II.A and II.B show the same for participants with and without CVC. In each plot, the central line inside the box indicates the median, the lower and upper edges of the box represent the 25th and 75th percentiles (interquartile range, IQR), and the whiskers extend to the most extreme values within 1.5 × IQR from the box. Data points beyond the whiskers are plotted individually as circles (mild outliers; >1.5 × IQR but ≤3 × IQR from the quartiles) or asterisks (extreme outliers; >3 × IQR from the quartiles). A large number of outliers were present in all groups, representing individuals with particularly high BMI or waist circumference, which may be clinically relevant for assessing metabolic and cardiovascular risk. Median BMI and waist circumference were higher in participants with T2D than those without, whereas marginal differences were observed between CVC and non-CVC groups.

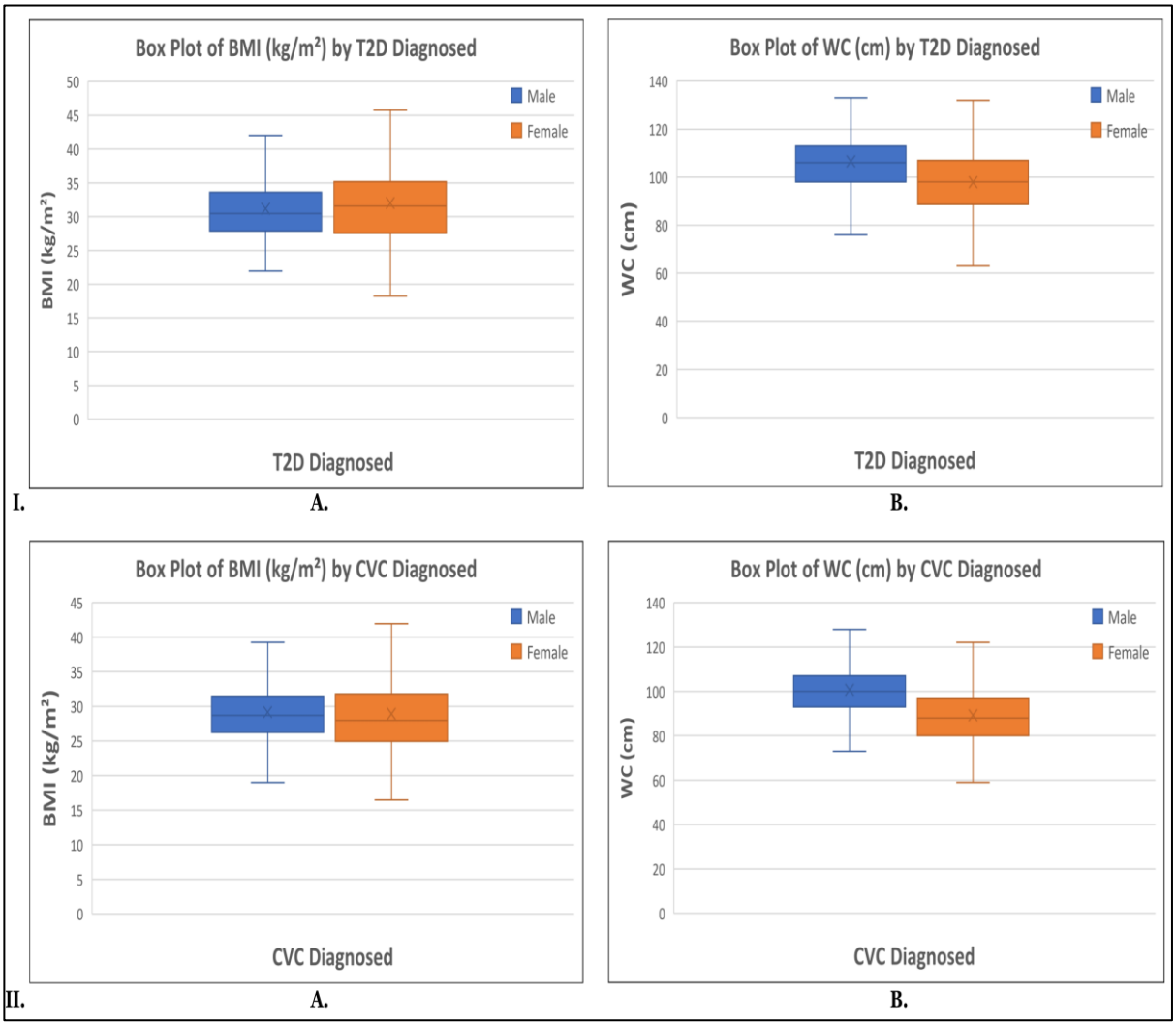


Figure 3. Boxplots of body mass index (BMI, kg/m²) and waist circumference (WC, cm) by diagnosis of type 2 diabetes (T2D) and cardiovascular conditions (CVC), stratified by sex. Panels I.A and I.B display BMI and WC distributions, respectively, for participants with T2D. Panels II.A and II.B show the same for participants with CVC. Blue boxes represent males and orange boxes represent females. In each plot, the central line inside the box denotes the median, X represents the mean, the lower and upper edges represent the 25th and 75th percentiles (interquartile range, IQR), and whiskers extend to the most extreme values within 1.5 × IQR from the quartiles. There were marginal differences in BMI for males and females with T2D or CVC, while males with T2D or CVC had higher waist circumference than females with T2D or CVC.

4.D. Prediction of confirmed T2D diagnosis

4.D.I. Prediction of confirmed T2D diagnosis by clinical parameters

The model containing all predictors (a total of 15 variables) was statistically significant, χ^2 (20, N = 9,668) = 1399.3, p < .001. The accuracy, specificity, and

sensitivity for the model were 97.1%, 99.4%, 38.5%, respectively. The AUC in ROC analysis was .959 ($p < .001$).

4.D.II. Prediction of confirmed T2D diagnosis by clinical parameters and T2D-associated Polygenic Risk Scores (PRS): Model D2

The model containing all predictors (a total of 18 variables) was statistically significant, χ^2 (23, N = 9,623) = 1,427.2, $p < .001$. The model as a whole correctly classified 97.1% of the cases; specificity was 99.4%, and sensitivity was 38.0%. The AUC in ROC analysis was .961 ($p < .001$).

4.D.III. Prediction of confirmed T2D diagnosis by clinical parameters, T2D-associated PRS and SNPs: Full model for T2D

The addition of the SNPs as predictor (a total of 19 variables; χ^2 (30, N = 9,623) = 1516.0, $p < .001$) correctly classified 97.3% of the cases; specificity was the same at 99.4%, and sensitivity increased to 42.5%, representing a 4.5% increase.

Six independent variables made a unique statistically significant contribution to the model, namely age, BMI, HP, diastolic BP, and standard PRS for T2D ($p < .05$; **Table 4**). The strongest predictor of T2D diagnosis in participants with the *LPL* SNPs investigated in this study was the Standard PRS for T2D. This indicated that the odds of participants being diagnosed with T2D were 1.6 times greater for each unit increase in PRS score, controlling for other factors in the model.

A ROC curve based on the full prediction model is presented in **Figure 4**; the AUC was .965 ($p < .001$).

4.D.IV. Prediction of confirmed T2D diagnosis by SNPs using Model D2

Four SNPs had a sufficient number of participants for data analysis: rs116403115 (n = 69, 7.2% with T2D), rs118204057 n = 311, 3.2% with T2D), rs547644955 (n = 224, 12.5% with T2D), and rs268 (n = 11,945, 3.6% with T2D). On investigation, three SNPs– rs116403115, rs118204057, and rs547644955- demonstrated 100.0% sensitivity and specificity. The sensitivity for rs268, which had the highest number of N, was 42.5%.

For the three SNPs with 100.0% specificity, sensitivity, and accuracy (i.e., rs116403115, rs118204057, and rs54764495), the AUCs were 1.0 as expected ($p < .001$). The AUC for rs268 was .963 ($p < .001$).

Variables	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
							Lower	Upper
Age at Recruitment	0.064	0.012	27.239	1	**< .001**	1.066	1.041	1.092
Sex(1)	-0.055	0.265	0.044	1	0.834	0.946	0.563	1.591
Weight (kg)	-0.068	0.044	2.392	1	0.122	0.934	0.856	1.018
Height Standing (cm)	0.088	0.048	3.352	1	0.067	1.092	0.994	1.200
BMI (kg/m^2)	0.355	0.13	7.491	1	**0.006**	1.426	1.106	1.84
Waist circumference (cm)	0.034	0.013	6.389	1	0.011	1.035	1.008	1.062
Hip circumference (cm)	-0.061	0.017	12.790	1	**< .001**	0.941	0.910	0.973
Ever Smoked			0.573	2	0.751			
Ever Smoked(1)	1.055	1.494	0.498	1	0.48	2.871	0.154	53.695
Ever Smoked(2)	1.004	1.492	0.453	1	0.501	2.730	0.147	50.798
Alcohol Drinker status			1.787	3	0.618			
Alcohol Drinker status(1)	-0.399	0.352	1.283	1	0.257	0.671	0.336	1.338
Alcohol Drinker status(2)	-17.238	17612.815	0.000	1	0.999	0.000	0.000	
Alcohol Drinker status(3)	0.250	0.382	0.430	1	0.512	1.284	0.608	2.714
Systolic BP (mmHg)	0	0.005	0.007	1	0.932	1	0.989	1.010
Diastolic BP (mmHg)	-0.045	0.010	20.983	1	**< .001**	0.956	0.938	0.975
Physical Activity	-0.036	0.036	0.958	1	0.328	0.965	0.899	1.036
Variation in Diet			2.461	3	0.482			
Variation in Diet(1)	-0.832	0.880	0.894	1	0.344	0.435	0.078	2.441
Variation in Diet(2)	-0.488	0.896	0.297	1	0.586	0.614	0.106	3.551
Variation in Diet(3)	-0.737	0.874	0.711	1	0.399	0.479	0.086	2.655
Glucose (mmol/L)	-0.003	0.045	0.005	1	0.941	0.997	0.912	1.089
HbA1c (mmol/mol)	0.196	0.011	303.998	1	**< .001**	1.216	1.190	1.243
Variant Group			8.581	7	0.284			
Variant Group(1)	1.438	3.692	0.152	1	0.697	4.211	0.003	5848.791
Variant Group (2)	0.745	3.595	0.043	1	0.836	2.106	0.002	2416.766
Variant Group (3)	-27.199	3198.500	0.000	1	0.993	0	0	.
Variant Group (4)	-0.734	3.582	0.042	1	0.838	0.480	0	536.881
Variant Group (5)	1.109	3.551	0.097	1	0.755	3.030	0.003	3194.619
Variant Group (6)	0.353	3.527	0.010	1	0.920	1.424	0.001	1431.903

Variant Group (7)	1.569	3.723	0.178	1	0.673	4.802	0.003	7087.110
Standard PRS for T2D	0.491	0.084	33.959	1	**< .001**	1.634	1.385	1.927
Standard PRS for BMI	0.042	0.081	0.270	1	0.603	1.043	0.890	1.222
Standard PRS for HbA1c	-0.042	0.073	0.333	1	0.564	0.959	0.831	1.106
Constant	-28.068	8.919	9.902	1	0.002	0.000		

Table 4. Logistic regression outcome for T2D diagnosis using the full prediction model. Table shows that six independent variables (age, BMI, HP, diastolic BP, and standard PRS for T2D) made a unique statistically significant contribution to the model ($p < .05$), with the Standard PRS for T2D being the strongest predictor of being diagnosed with T2D, indicating that the odds were 1.6 times greater that the participants were diagnosed with T2D with per unit increase of PRS score, controlling for other factors in the model. B = B coefficient (representing change in the log-odds of the outcome for a unit change in variable), S.E. = standard error, Wald = (B/S.E.)^2, T2D = type 2 diabetes, BMI = body mass index, BP = blood pressure, HbA1c = glycated haemoglobin, PRS = polygenic risk score, df = degrees of freedom, Sig. = significance ($p < .05$ bolded to emphasize significance), Exp(B) = exponential value of B coefficient (e^B), 95% C.I. (confidence interval) for EXP(B) = range within which the true odds ratio is likely to fall with 95% confidence (lower and upper values shown defining the boundaries of the 95% C.I.). Physical activity is defined as number of days/week of moderate physical activity 10+ minutes. *SPSS coding used* (reference- used as baseline for comparison): Sex = 0 for female, 1 for male; Ever smoked = 0 for No, 1 for Yes; Alcohol drinking status = 0 for never, 1 for previous, 2 for current (reference); Variation in diet = 0 for never/rarely, 1 for sometimes, 2 for often; Variant Groups: 1 = rs118204061 (reference), 2 = rs1444466625, 3 = rs116403115, 4 = rs11542065, 5 = rs118204057, 6 = rs547644955, 7 = rs268, 8 = 2 SNPs.

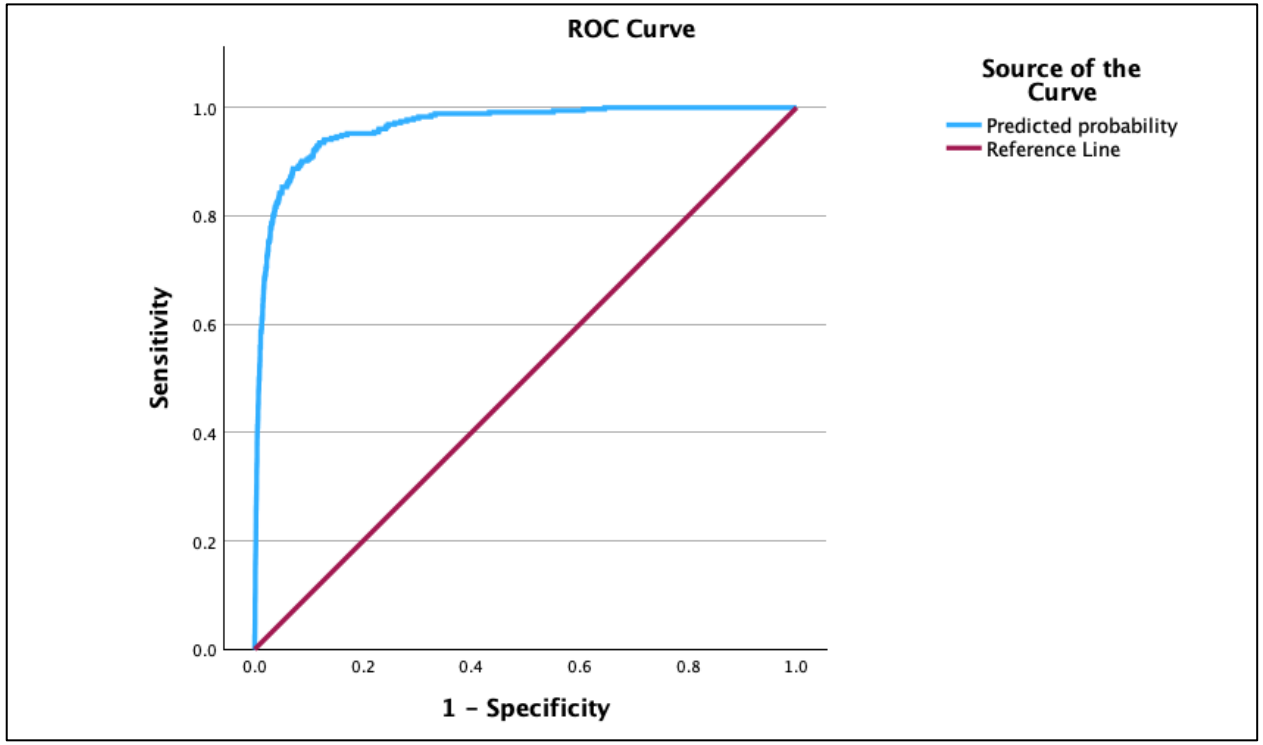


Figure 4. Receiver Operating Characteristic (ROC) curve of the full model for prediction of T2D diagnosis, showing predicted probability (blue curve)) in relation to reference line (red line), indicating high classification accuracy (AUC = .965, $p < .05$) for the studied cohort (N = 12,872). T2D = type 2 diabetes, AUC = area under the curve

4.E. Prediction of confirmed CVC diagnosis

4.E.I. Prediction of confirmed CVC diagnosis by clinical parameters

The model containing all predictors (a total of 15 variables) was statistically significant, χ^2 (20, N = 9668) = 1,852.2, p < .001. The accuracy, specificity, and sensitivity for the model were 74.8%, 91.4%, 32.1%, respectively. The AUC in ROC analysis was .772 (p < .001).

4.E.II. Prediction of confirmed CVC diagnosis by clinical parameters and CVC-associated PRS: Model E2

The model containing all predictors (a total of 22 variables) was statistically significant, χ^2 (27, N = 9623) = 2,132.9, p < .001. The model as a whole correctly classified 75.8% of the cases, specificity was 90.9%, and sensitivity was 37.1%. The AUC in ROC analysis was .790 (p < .001).

4.E.III. Prediction of confirmed CVC diagnosis by clinical parameters, CVC-associated PRS and SNPs: Full model for CVC

The addition of the SNPs as predictor (total 23 variables; χ^2 (34, N = 9,623) = 2158.6, p < .001) correctly classified 75.9% of the cases, specificity was the same at 90.9%, and sensitivity increased to 37.5%, a small increase of 0.4%.

Twelve independent variables made a unique statistically significant contribution to the model: age, WC, HC, alcohol drinker status, systolic BP, diastolic BP, variation in diet, HbA1c, variant group, and standard PRS for CVD, BMI, and hypertension (p < .05; **Table 5**); In reference to other SNPs, the odds of being diagnosed with CVC differs. Being a current alcohol drinker as well as the PRS for hypertension show higher odds ration of 1.7 and 1.5, respectively, controlling for other factors in the model.

A ROC curve based on the full prediction model for CVC (AUC = .837, $p < .001$) is presented in **Figure 5**. The AUC for the full model was .837 ($p < .001$).

4.E.IV. Prediction of confirmed CVC diagnosis by SNPs using Model E2

Upon investigation of SNP groups with sufficient number of participants for data analysis (four SNPs: rs11542065 (n=276, 32.6% with CVC), rs118204057 (n=311, 29.3% with CVC), rs547644955 (n=224, 43.8% with CVC), and rs268 (n=11,945, 27.9% with CVC)), rs547644955 had the highest sensitivity at 75.9%, specificity 83.1%, and accuracy 80.9% (ROC curve showing AUC = .910, $p < .001$, is shown in **Figure 5**). The other 3 SNPs (rs11542065, rs118204057, and rs268) had lower sensitivity (50.8, 33.8, and 36.9, respectively).

Variables	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
							Lower	Upper
Age at Recruitment	0.0700	0.004	311.141	1	**< .001**	1.072	1.064	1.081
Sex(1)	-0.103	0.094	1.198	1	0.274	0.902	0.750	1.085
Weight (kg)	0.021	0.018	1.375	1	0.241	1.021	0.986	1.058
Height Standing (cm)	-0.020	0.018	1.268	1	0.260	0.980	0.946	1.015
BMI(kg/m2)	0.035	0.053	0.434	1	0.510	1.035	0.934	1.148
Waist circumference (cm)	0.030	0.005	36.574	1	**< .001**	1.030	1.020	1.040
Hip circumference (cm)	-0.038	0.007	32.137	1	**< .001**	0.962	0.950	0.975
Ever Smoked			5.076	2	0.079			
Ever Smoked(1)	0.247	0.468	0.278	1	0.598	1.280	0.512	3.201
Ever Smoked(2)	0.363	0.467	0.604	1	0.437	1.438	0.576	3.590
Alcohol Drinker status			12.934	3	0.005			
Alcohol Drinker status(1)	0.064	0.135	0.225	1	0.635	1.066	0.818	1.390
Alcohol Drinker status(2)	0.762	0.952	0.642	1	0.423	2.143	0.332	13.847
Alcohol Drinker status(3)	0.518	0.148	12.226	1	**< .001**	1.678	1.256	2.244
Systolic BP (mmHg)	0.016	0.002	69.259	1	**< .001**	1.016	1.012	1.020
Diastolic BP (mmHg)	0.021	0.003	37.719	1	**< .001**	1.021	1.014	1.028
Physical activity	-0.008	0.014	0.319	1	0.572	0.992	0.966	1.019
Variation in Diet			22.046	3	**< .001**			
Variation in Diet(1)	-0.015	0.388	0.001	1	0.969	0.985	0.461	2.108
Variation in Diet(2)	0.138	0.396	0.122	1	0.727	1.148	0.529	2.493
Variation in Diet(3)	0.250	0.387	0.418	1	0.518	1.284	0.602	2.740
Glucose (mmol/L)	0.006	0.026	0.050	1	0.823	1.006	0.957	1.057
HbA1c (mmol/mol)	0.013	0.005	6.400	1	**0.011**	1.013	1.003	1.024
Variant Group			25.564	7	**< .001**			
Variant Group(1)	2.924	1.268	5.314	1	0.021	18.617	1.549	223.696
Variant Group(2)	2.238	1.169	3.663	1	0.056	9.377	0.948	92.786
Variant Group(3)	1.966	1.134	3.005	1	0.083	7.143	0.773	65.966
Variant Group(4)	1.518	1.131	1.803	1	0.179	4.565	0.498	41.865
Variant Group(5)	2.276	1.138	3.998	1	0.046	9.734	1.046	90.597
Variant Group(6)	1.586	1.120	2.006	1	0.157	4.884	0.544	43.853
Variant Group(7)	1.353	1.331	1.034	1	0.309	3.871	0.285	52.593

Standard PRS for CVD	0.103	0.039	7.076	1	**0.008**	1.108	1.027	1.195
Standard PRS for BMI	-0.059	0.027	4.552	1	**0.033**	0.943	0.894	0.995
Standard PRS for AF	-0.007	0.029	0.066	1	0.797	0.993	0.938	1.050
Standard PRS for CAD	0.053	0.039	1.807	1	0.179	1.054	0.976	1.138
Standard PRS for hypertension	0.411	0.030	189.265	1	**< .001**	1.508	1.422	1.599
Standard PRS for HDL cholesterol	-0.023	0.026	0.782	1	0.376	0.977	0.929	1.028
Standard PRS for LDL cholesterol	0.010	0.025	0.162	1	0.688	1.010	0.961	1.062
Constant	-9.443	3.238	8.507	1	0.004	0		

Table 5. Logistic regression outcome for CVC diagnosis using the full prediction model. Table shows that twelve independent variables (age, WC, HC, alcohol drinker status, systolic BP, diastolic BP, variation in diet, HbA1c, variant group, and standard PRS for CVD, BMI, and hypertension) made a unique statistically significant ($p < .05$) contribution to the model. Being a current alcohol drinker as well as the PRS for hypertension show high odds ration of 1.7 and 1.5, respectively, controlling for other factors in the model. B = B coefficient (representing change in the log-odds of the outcome for a unit change in variable), S.E. = standard error, Wald = (B/S.E.)^2, df = degrees of freedom, Sig. = significance ($p < .05$ bolded to emphasize significance), Exp(B) = exponential value of B coefficient (e^B), 95% C.I. (confidence interval) for EXP(B) = range within which the true odds ratio is likely to fall with 95% confidence (lower and upper values shown defining the boundaries of the 95% C.I.), CVC = cardiovascular conditions, BMI = body mass index, BP = blood pressure, HbA1c = glycated haemoglobin, PRS = polygenic risk score, CVD =cardiovascular disease, AF = atrial fibrillation, CAD = coronary artery disease, HDL = high-density lipoprotein, LDL = low-density lipoprotein. Physical activity is defined as number of days/week of moderate physical activity 10+ minutes. *SPSS coding used* (reference- used as baseline for comparison): Sex = 0 for female, 1 for male; Ever smoked = 0 for No, 1 for Yes; Alcohol drinking status = 0 for never, 1 for previous, 2 for current (reference); Variation in diet = 0 for never/rarely, 1 for sometimes, 2 for often; Variant Groups: 1 = rs118204061 (reference), 2 = rs1444466625, 3 = rs116403115, 4 = rs11542065, 5 = rs118204057, 6 = rs547644955, 7 = rs268, 8 = 2 SNPs

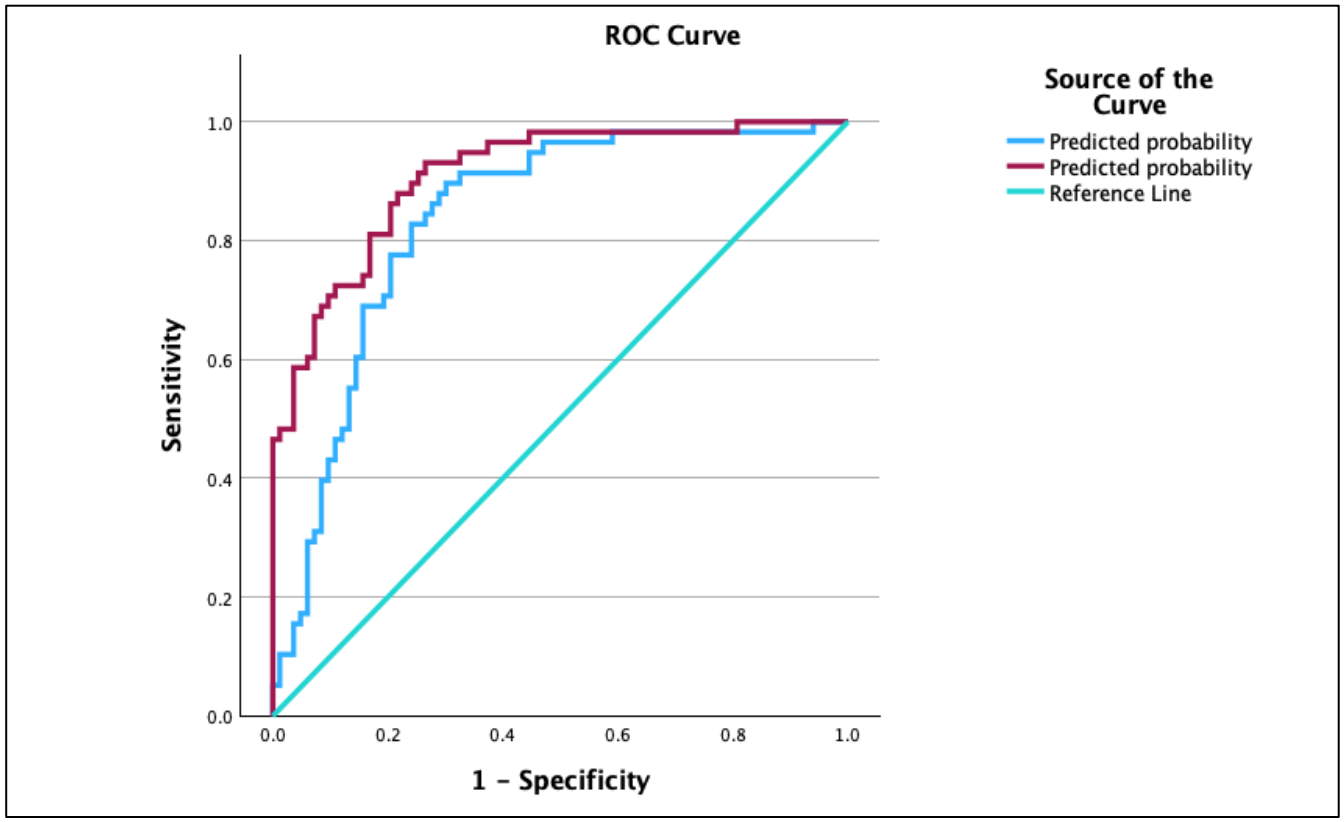


Figure 5. Receiver Operating Characteristic (ROC) curve of the model for prediction of CVC diagnosis, showing full model (blue curve) and in individuals with rs547644955 (red curve) in relation to reference line (green line), indicating higher classification accuracy for rs547644955 (AUC = .910, $p < .001$; N = 224) than the full model (AUC = .837, $p < .001$; N = 12,872. CVC = cardiovascular conditions, AUC = area under the curve

5. DISCUSSION

5.A. Comparison amongst SNPs and association studies

A review by Brown and Walker in 2016 on MetS GWAS reported that the major genomic risk loci for MetS were in, or close to, lipid-regulating genes including *LPL* (Brown and Walker, 2016). In 2019, a GWAS on MetS in the UK Biobank was published, as previously discussed, reporting 80 novel independent loci using the harmonized National Cholesterol Education Program (NCEP; Appendix) criteria for MetS (Lind, 2019). According to Heart UK, MetS is very common amongst the British, with an estimated prevalence of 25%, a rate which is consistent with the worldwide MetS prevalence (Noubiap et al., 2022). Expanding the body of knowledge on the genetic patterns of MetS may provide causal links, preventative insights, and potential therapeutic propositions with significant global healthcare impact.

In this study, SNPs previously associated with MetS were evaluated in relation to the development of T2D and CVC. Significant associations between the SNPs rs268, rs11542065, rs116403115, rs118204057, rs118204061, rs144466625, and rs54764495 and diagnosis of both T2D and CVC (4.B.1, Figure 1) were found. Previous reports also identified associations of *LPL SNPs* with MetS: rs320, rs328, rs1801177, rs268 (Ariza et al., 2010) and rs328 (Cahua-Pablo et al., 2015). In this cohort, rs268 was found to be the most common, and previous large studies have reported significant findings in relation to MetS (Franceschini et al., 2011, Nejati et al., 2018, Sagoo et al., 2008). For example, the Human Genome Epidemiology (HuGE) Review found consistent lipid effects on coronary heart disease (CHD) risk in white participants. Here, however, a MetS feature was the focus,

that is CHD, rather than MetS as defined (Sagoo et al., 2008). The Population Architecture using Genomics and Epidemiology (PAGE) Study was another large study evaluating rs268 and also CHD, but did not find significant associations (Franceschini et al., 2011). Its participants were primarily Hispanics and Asians, which may signify the crucial role of ethnicity. Notably, in this study - comprising 90% British participants - significant differencee were found between the seven SNPs and CVC. Therefore, the difference of ethicity in the PAGE study (Hispanics and Asians) may be significant to that in this study (mostly British). A computational analysis performed on a recent study has shown that rs268 (as well as rs328) may affect the protein structure, while a meta-analysis done in the same study indicated that stroke risk was decreased in other *LPL* SNPs (i.e. rs320 and rs285) (Nejati et al., 2018).

The differences amongst groups with the assessed clinical parameters with significance such as weight, BMI, diastolic BP, and waist circumference may be of particular interest for the further study of these SNPs. HbA1c was of significant difference, for instance, but not random blood glucose (Table 3A and Table 3C). This is in contrast to the two previous studies which evaluated an *LPL* SNP, rs285, which reported the relation of this SNP to both BMI and fasting blood glucose (Bozina et al., 2013, Chuluun-Erdene et al., 2020). Although this SNP was not included in this study, differences as this warrant further scrutiny. In addition, combinations of frequencies in selected genes may exacerbate obesity as with the result from a study, which evaluated relationships between *LPL* m107 (rs1800590) and APOA5 S19W (rs3135506) and lipid and anthropometric measures (Smith et al., 2010). Combinations of

frequencies were not investigated in this study as only the seven *LPL* SNPs were evaluated; however, the combination of these *LPL* SNPs with other genes may be explored in future studies. The utility of PRS for risk prediction of both T2D and CVC, amongst many diseases, is gaining interest (Arnold and Koenig, 2021, Hahn et al., 2022); however, the non-significant association of PRS with the SNPs in this study (4.B.IV) is expected due the interplay of numerous factors including environmental influences and epigenetic regulation.

Numerous publications indicate the association of BMI with the pathogenesis of both T2D and CVC (Kolb and Martin, 2017, Larsson and Burgess, 2021, Meigs et al., 2006). Abdominal obesity, often presented as weight circumference as an anthropometric measure, has also been widely investigated and reported to be risk factor for both diseases (Franek et al., 2023, Qiao et al., 2022). A large genetic study in the United Arab Emirates has studied genetic associations between T2D and coronary artery disease (CAD) and their associations with several cardiometabolic features; this study found that the strongest association with CAD was detected with SNP rs264 in *LPL* (Osman et al., 2020). In the present study, a stronger association of the *LPL* SNPs is evident, compared to their association with T2D, which can be attributed to the role of LPL in fatty acid metabolism (**Figure 1**). Regarding the marginal higher BMI and WC amongst men compared to women with T2D which was found in this study (4.C.2), a review article on the sex differences on features of diabetes reported a somewhat different, though not entirely contradictory; their results indicated that BMI is a better predictor of T2D in men, while it is WC for women (Kautzky-Willer et al., 2016). For CVC, the

outcome in the study was similar, with slightly higher BMI and WC observed amongst men compared to women. However, more studies are required to establish distinct features between sexes in CVC (Regitz-Zagrosek and Gebhard, 2023). Nevertheless, these differences, including effects of other confounding variables may be further investigated with *LPL* SNPs to clarify their role in the pathogenesis of T2D and CVC.

Genome-wide association studies (GWAS) for MetS have been reported for various ethnic populations and from multiethnic backgrounds (Moon et al., 2018, Oh et al., 2020, Tekola-Ayele et al., 2015). In 2011, a systematic review on the genetic variants associated with MetS has outlined the most studied SNPs linked with MetS (rs9939609 (*FTO*), rs7903146 (*TCF7L2*), C56G (*APOA5*), T1131C (*APOA5*), C482T (*APOC3*), C455T (*APOC3*) and 174G>C (*IL6*)); *LPL* was not included in this report (Povel et al., 2011). A recent (2019) publication on MetS GWAS which has used the UK Biobank data has been published, which reported 80 novel independent loci; *LPL* SNP rs3844510 was included although not as a novel finding (Lind, 2019). The use of larger data sets such as in the latter has been argued to be of significance particularly for linkage and candidate gene studies including MetS (Monda et al., 2010). Nevertheless, current evidence suggests that the genetic risk factors for MetS are strongly connected with the components of MetS, including hyperglycemia and dyslipidemia (Taylor et al., 2013). This study provides a significant amount of additional information, expanding knowledge of *LPL* SNPs which are not widely investigated for MetS. Moreover, the use of the UKB data as a credible source of a large data set is a notable strength, adding to the research's value.

5.B. Prediction of Confirmed T2D or CVC

Prediction models have been trialled using various parameters, including risk factors, to estimate the probability of T2D and/or CVD development in multiple studies using logistic regression and machine learning approaches (Dinh et al., 2019, Edlitz and Segal, 2022, Joshi and Dhakal, 2021). These models may aid in formulating preventive measures for those who may be deemed at risk for develeping the disease. In this study, the logistic regression model for T2D and CVC had high accuracy, specificity, and AUC in ROC analysis. Sensitivity was considerably low, except for rs547644955, and rs116403115 and rs118204057 (Sections 4.D.1-III, 4.E.I-III). However, the ROC AUC is regarded as a superior assessment tool for medical diagnostic evaluation due to the arbitrary nature of specificity, sensitivity, and accuracy, which is deemed problematic (Hajian-Tilaki, 2013, Swets, 1988). Therefore, the ROC AUC better distinguishes between healthy versus diseased population (Metz, 1978), and the models assessed in this study may be of value (Figure 4 and Figure 5), including the addition of SNPs particularly for individuals diagnosed with T2D (Sections 4.D.III and 4.E.III).

The variant that showed major significance for both T2D and CVC was rs547644955 (Sections 4.D.IV and 4.E.IV). The other two variants with significance for T2D were rs116403115 and rs118204057 (Section 4.D.IV). There appear to be no substantial publications for rs547644955 and rs116403115; therefore these findings may pave the way for a greater understanding on these SNPs and clinical applications (e.g. diagnosis, prevention, treatment) for individuals identified with these SNPs. There were a

few reports directly associated with the variant rs118204057, including heritability in ethnic groups (Gagné et al., 1989, Henderson et al., 1992, Paulweber et al., 1991). The variant rs268 was the most common in the cohort studied, and previous publications reported MetS-specific resuts (Franceschini et al., 2011, Nejati et al., 2018, Sagoo et al., 2008). Nonetheless, as the full model for both T2D and CVC diagnosis had good predictability based on the ROC analysis, these may be have research value in addition to mentioned clinical applications, as well as to obesity studies overall.

Upon exclusion of the genetic parameters on both full models, the decrease in AUC was marginal (.959 versus .965 for the full model) for T2D (Results section B, D), while that for CVC may be considerable (.772 versus .837 for the full model; Results section C, D). Although the clinical relevance may need further investigation, the availability of the fifteen variables used in the model (i.e. age, sex, weight, height, BMI, WC, HC, systolic BP, diastolic BP, number of days per week of moderate physical activity, diet variation, smoking status, alcohol drinking status, random blood glucose, HbA1c) are readily accessible or easily obtainable in routine healthcare settings. T2D, although related to CVD through several similar risk factors, is itself a known risk factor for the development of CVD, but the reverse does not seem to be true (Dal Canto et al., 2019, Einarson et al., 2018, Kelsey et al., 2022). This consideration primarily influenced the selection of PRSs for the *full* models in the T2D and CVC diagnosis prediction models. The *full* model for CVC diagnosis was trialled with the addition of PRSs for T2D and glycated haemoglobin, however the result of the AUC in ROC analysis did not differ

(result not presented). The American Heart Association (AHA) has recently published (2022) a scientific statement regarding PRS for CVD as well as other related conditions such as T2D (O'Sullivan et al., 2022). PRS is normally derived from single nucleotide variant effect sizes from GWAS then adjusted for linkage disequilibrium (Choi et al., 2020), and large biorepositories such as the UKB provides these data as what has been used in this study. As per summary of the AHA statement, the utility of PRS for CVD and associated disorders appears somehow different based on specific disease states as evidenced by various research. In CVD, CAD is the most studied form in terms of PRS research and its use is mainly geared towards pharmacological management (Damask et al., 2020, Khera et al., 2016, Said et al., 2018). In T2D, earlier studies point to similar utility of PRS with clinical factors, while more recent evidence suggests that PRS may be additive to the latter (Mars et al., 2020, Meigs et al., 2008, Talmud et al., 2010). Yet other studies suggest unclear significance of T2D high-risk identification (Hivert et al., 2011, Said et al., 2018). These findings are relatively in accordance with the results of this research in terms of ambiguous usability of PRS addition to the prediction model. While some studies suggest PRS for T2D may be useful for assessing response to sulfonylureas (Li et al., 2021) and for glucose management (Shah et al., 2016), the clinical applications of PRS may be worth pursuing in this era of advanced genomic investigations.

5.C. Further Discussion Points and Results Implications

In the online resource dbSNP from the NIH National Library of Medicine, as described, several details regarding SNPs are available such as

frequency (total frequency and frequency from various ethnicities (e.g. European, African, Asian, etc.), variant details (such as genomic placements including sequence name and change; genomic regions, transcripts, and products), Clinical significance (disease name and clinical significance), human genome variant society (HGVS) standard, submissions (details of submitter, submission ID, and date), history, publications involving the SNPs, and flanks (adjacent nucleotide sequences). It is an easily accessible reference detailing this information for SNPs, which are used in further discussion in this section, with a focus on significant results; list for reference can be found in Appendix 4.

The *LPL* SNP rs268 (alleles: A>G) has the highest total frequency (global total frequency = 235,134 as of 02 May 2024) amongst the cohort studied, which was also observed in this study (i.e. highest n is for rs268, which was 11,945). There are 29 citations/publications listed which involve this SNP (also the highest among the SNPs of interest in this study). The clinical significance was mostly familial hyperlipidemia (pathogenic or risk factor) and hyperlipoproteinemia type 1 (uncertain significance), along with benign results and conflicting interpretations of pathogenicity. In this study, rs268 was significantly associated with confirmed diagnosis of both T2D and CVC. For prediction of confirmed T2D, however, the sensitivity for rs268 was low (42.5%), particularly in comparison to the three other SNPs (rs116403115, rs118204057, and rs547644955), which had sufficient number of participants for analyses – these three other SNPs, in fact, had a sensitivity of 100.0%. For prediction of confirmed CVC, the sensitivity for rs268 was low (36.9%), although this time, two of the three (rs11542065, rs118204057,

rs547644955,) other SNPs also had low sensitivity (50.8% and 33.8% for rs11542065, rs118204057, respectively) while rs547644955 had a good sensitivity of 75.9%. It is important to note here, however, that AUC is a more reliable measure, as discussed, and for both the full models for T2D and CVC, AUC were acceptable (AUC = .965 and .867, for T2D and CVC, respectively).

The *LPL* SNP rs268 is one of the *LPL* SNPs, which may be of further interest for research primarily because of its highest global frequency in the cohort studied compared to the other SNPs in this study. It is a coding sequence variant, with missense functional consequence, and located at chr8:19956018 (GRCh38.p14).

5.D. Current Conventions and Future Directions

As discussed, pharmacogenetics and pharmacogenomics are two of the most direct benefactors of genetic and genomic studies in different diseases including T2D and CVC for preventive and therapeutic benefits while minimizing adverse drug reactions. The primary focus of the study of pharmacogenetics is on single genes, while the effects of numerous genes in the genome are investigated in pharmacogenomics. There are several references online that provide comprehensive data on drug details including pharmacological, molecular, and chemical information such as DrugBank (https://go.drugbank.com/), PharmGKB (https://www.pharmgkb.org), ChEMBL (https://www.ebi.ac.uk/chembl), PubChem (https://pubchem.ncbi.nlm.nih.gov), KEGG DRUG (https://www.genome.jp/kegg/drug), and Therapeutic Target Database (TTD; http://db.idrblab.net/ttd/). These platforms have their own focus (eg. drug and drug target information for DrugBank, pharmacogenomics

for PharmGKB, chemical molecules and their activities in biological assays for PubChem), and are readily accessible for biomedical scientists worldwide.

Although pharmaceutical profiling was beyond the scope of this research particularly given its retrospective nature, it is valuable to provide a brief background on the application of these disciplines in T2D and CVD to aid in understanding on what we currently know on these diseases. In addition, this will help solidify the impact and contribution that this particular study provides in this field, being one of the springboards for further investigations on the evaluated *LPL* SNPs, given that only a considerably small number of reports have been published. As such, this section outlines the different conventional drug classes, drug examples, associated genes, mechanism of action, and other drug type-specific parameters for T2D (5.D.1) and CVC (5.D.II).

5.D.I. Pharmacogenetics in T2D

The following texts are all derived and/or summarized from at least one or combination of the presented references on this sentence (2016b, 2016a, Dawed et al., 2016, Garber et al., 2015, Gentilella et al., 2019, Gloyn and Drucker, 2018, Hieronymus and Griffin, 2015, Mannino et al., 2019, Rodbard, 2018, Srinivasan et al., 2018, Zhou et al., 2016). Accordingly, these texts are italicised without specific referencing for each paragraph or section. Genetic, genomic, or SNP associations, linkages, or connections to drug types/classes presented on this section refer to any or combination of literature on the effect of genes, genotype, genetic variants or SNPs to drug interactions, metabolism, safety, efficacy, response, use and other drug-related factors

encompassing clinical pharmacology (e.g. pharmacokinetics, pharmacodynamics, toxicology).

Biguanides

Next to lifestyle modification measures, which primarily come in the form of diet and physical activity, the use of metformin is usually the first-line medication for T2D in the absence of contraindications such as severe renal or hepatic insufficiency. The liver is the major site of action for Metformin, which belongs to the drug class of biguanides. It has been known to be associated with several genes including SLC22A1, SLC22A2, SLC22A3, SLC47A1, and SLC47A2, ATM, and IRS1. Numerous studies since year 2007 have investigated metformin's pharmacokinetics describing diverse phenotypes and specifying gene variants involved. The main advantages of metformin over other diabetes therapies include good safety profile, cheap cost, and effectivity. Evidence also suggests that it may aid weight loss and may be used for T2D prevention and polycystic ovary syndrome (PCOS) treatment. One of its known main disadvantages, however, is its insufficiency as a monotherapy to meet glycaemic control.

Sulfonylureas

Sulfonylureas are previously considered as the first choice for T2D management, but are now used as the second-line treatment in combination with metformin. Its primary site of action is the pancreas where it directly stimulates insulin secretin from the beta cells. Associated genes include KCNJ11, ABCC8, CYP2C9, and TCF7L2. Reported benefits of sulfonylureas include intensive glucose control for several years and reduction in

cardiovascular events as well as decreased all-cause mortality during longer-term follow-up. Its primary weakness, however, is increased rates of hypoglycaemia and weight gain.

Dipeptidyl peptidase (DPP)-4 inhibitors

Also known as gliptins, DPP4 inhibitors are classified as incretin mimetics because they inhibit a key enzyme in the incretin signalling pathway. They mainly act in the intestines and are linked with CYP3A4, CYP2C8, and TCF7L2 genes. The increasing use of DPP4 is essentially attributed to its good safety profile. As they are relatively new in the market, very few pharmacogenetics studies have been conducted. The liver is not important for the elimination of gliptins as their main mode of clearance is by renal excretion.

Thiazolidinediones (TZDs)

TZDs are a group of drugs that act on adipose tissue to increase glucose utilization and decrease glucose production. Primary associated gene is PPAR-γ; others include ADIPOQ1, CYP2C8, CYP2C9, and CYP3A4. Inconsistent outcomes have been reported which may be attributed to differences on type of TZDs used, treatment duration, inclusion criteria, baseline metabolic conditions, and ethnicity. Major cause of the limited use of TZDs has been attributed to severe adverse events such as heart failure, myocardial infarction, and bladder cancer.

Sodium-glucose transporters (SGLT)-2 inhibitors

Also called gliflozins, SGLT-2 inhibitors reduce hyperglycemia through glucose elimination via urine, thereby acting in the kidneys. So far, there have been no definitive pharmacogenetics studies, however, which directly relates genetic variants and SNPs in response to SGLT2 inhibitors. Hence, more genetic studies including long-term outcomes of the use of SGLT-2 inhibitors for the management of patients with T2D are called for.

α-glucosidase inhibitors

The α-glucosidase inhibitors are primarily represented by the drug acarbose, which inhibits the upper gastrointestinal enzymes (alpha-glucosidases) that convert polysaccharide carbohydrates into monosaccharides thereby decreasing glucose absorption in the intestines. Associated genes for drug efficacy/toxicity include PPAR-γ, HNF4A, and LIPC.

Glucagon-like Peptide-1 (GLP-1) receptor agonists (RAs)

GLP-1 RAs have well-established safety and efficacy profiles in patients with T2D. They enhance prandial insulin secretion in the pancreas and can either be long-acting or short-acting. GLP-1 RAs are known to have excellent potency in reducing HbA_{1c} and mean glucose, improving fasting plasma glucose, and inducing weight loss. They are likewise recognized to have cardioprotective effects. Although gastrointestinal side-effects were observed, this tends to diminish over time. Associated genes with GLP-1 RAs have not yet been identified.

Meglitinides

Meglitinides are short-acting glucose-lowering drugs which exert their effects via pancreatic beta cell receptors. Similar to sulfonylureas, they act by increasing insulin secretion, although they are distinct in structure. Implicated SNPs associated with T2D include KCNJ11, KCNQ1, UCP2, NAMPT, MDRI, PAX4, NEUROD1, and SLCO1B1.

Amylin mimetics

Most known in this group of anti-diabetic drugs is pramlintide, a synthetic amylin analog acting on pancreatic beta cells. It works by suppressing glucagon release in response to caloric intake, delaying the rate of gastric emptying, and stimulating the satiety center in the brain to limit caloric intake. Although it offers favorable effects to body weight, the risk of hypoglycemia is increased along with other adverse effects. Not much is known on the pharmacogenetics of this drug class.

5.D.II. Pharmacogenetics in CVC

Similarly, pharmacogenetic applications for CVC are also gaining popularity. A systematic review on pharmacogenetics in CVD published in 2012 outlined the readiness for clinical use (Verschuren et al., 2012), while a newer report in 2023 supports the additive role of pharmacogenomics in improving patient care and treatment outcomes (Saleh et al., 2023). Yet another two recent publications in 2024 (Ingelman-Sundberg and Pirmohamed, 2024) and 2023 (Padmanabhan et al., 2023) provided perspective and evaluations on the current role and applications of

pharmacogenetics and pharmacogenomics in CVD therapeutics in particular, and of precision medicine in general.

As with how the drug classes have been presented for T2D in the preceding section, the following texts summarize the most common CVD medications used and the underlining genetics involved - these are all derived from at least one or combination of the presented references on this sentence (Ingelman-Sundberg and Pirmohamed, 2024, Padmanabhan et al., 2023, Saleh et al., 2023, Verschuren et al., 2012). As such, these texts are italicised without specific referencing for each paragraph or section. Also as with the preceding section for T2D (5.D.I), genetic / genomic / SNP associations, linkages, or connections to drug types / classes presented on this section for CVC refer to any or combination of literature on the effect of genes, genotype, genetic variants or SNPs to drug interactions, metabolism, safety, efficacy, response, use and other drug-related factors encompassing clinical pharmacology (eg. pharmacokinetics, pharmacodynamics, toxicology).

Warfarin

*Preventing or treating thromboembolism is the main role of warfarin, a coumarin derivative, which is one of the main stays in CVD therapy. Significant pharmacogenetic implications of warfarin include interpatient dosing variability, wherein genetic variation accounts for 55-60%, while non-genetic factors (e.g. age, BMI) comprise a lower approximate of 20%. VKORC1, for example, accounts for 25% dosing variability, CYP2CP at approximately 15%, and CYP4F2*3 at approximately 1-7%. In addition,*

several CYP2C9 alleles require decreased dose due to reduced clearance of S-warfarin.

Clopidogrel

Clopidogrel is the most known medication in its class, which is primarily used as an antiplatelet therapy. Effectiveness of this medication is partly attributed to genetic variation. CYP2C19 alleles have different metaboliser phenotypes with clopidogrel (e.g. normal metaboliser, intermediate metaboliser, or poor metaboliser), while the antiplatelet drugs pasugrel and ticagrelol are not affected by this genotype. A number of real-world research have investigated the pharmacogenetic effects of these drugs, including relative risk of MACE (major adverse cardiovascular events), bleeding risk, stroke prevention, and effects when used when added to aspirin dosage.

Direct-acting oral anti-cogaulants (DOACs)

Although studies on pharmacokinetics due to genetic variation have been done, there are no clinical outcomes reported as yet. However, genotype sensitivity to dosing and bleeding risks were indicated, including a report that, compared to warfarin, dabigratan had reduced bleeding risk in CES1 rs2244613 minor allele.

Statins

Statins is another group of medications, which is a cornerstone of CVD prevention and therapy with its lipid-lowering capability. On pharmagenetics, it has been noted that risk of myopathy increases for more than 1.5 times per

*copy of SLCOB1B1*5 in patients on a doubled simvastatin dosage. On another note, there were inconsistent reports with CYP3A4, ABCB1, COQ2, and GATM.*

Beta-blockers

Beta-blockers are used for the treatment of heart failure, hypertension, and secondary prevention of myocardial infarction. Although currently, there is weak evidence for the pharmacogenetics of Beta-blockers, central to its study is the variations in CYP2D6, which is responsible for the biotransformation of up to 80% of metropolol oral dose. Other genes associated with pharmacodynamics rather than pharmacokinetics include ADRB1, ADRB2, and GRK5.

Hydralazine

*Another medication for hypertension is hydralazine. Some NAT2 alleles differ in phenotypic characteristics – e.g. homozygous NAT2*5, 6, and 7 exhibit slow acetylator phenotype, while heterozygous NAT2*4 and *5 are intermediate acetylators. A study also reported that the slow acetylator phenotype had better blood pressure reduction with hydralazine, although another study indirectly showed that an adverse drug reaction (ADR) in the form of lupus-like symptoms appear with the slow acetylator phenotype.*

Anti-arrythmic drugs

As with beta-blockers, anti-arrythmic drugs are also metabolised by CYP2D6. Caution in the use of propafenone amongst patients with CYP2D6

deficiency has been released by the FDA when combined with CYP3A4 inhibition.

Angiotensin-converting enzyme (ACE)-inhibitors

ACE inhibitors are another class of medications for treating hypertension. The polymorphism most commonly studied in the pharmacogenetics of this drug is rs4646994. However, results of research studies are inconsistent in terms of mortality risk and therapeutic response among others.

5.D.III Pharmacogenetics in T2D and CVC: The Contribution of this Study in A Nutshell

Based on the aforementioned studies and comprehensive reviews of pharmacogenetics in T2D and CVC, it is clear that *LPL* is not a significant player currently. For instance, dosing requirements for individuals suffering from T2D or CVC with *LPL* SNPs, or the risks involved in prescribing the various drug classes to individuals with these SNPs, are not yet elucidated nor known, to the best of the researcher's knowledge and latest electronic search.

In earlier studies conducted in 2002 (Brisson et al., 2002) and 2014 (Gao et al., 2014), the response to fenofibrate therapy amongst individuals with LPL genetic variants was investigated, with findings suggesting that these variants may modulate the response to this therapy (e.g. attenuated response in the 2014 study). Another study in 2004 (Brousseau et al., 2004) evaluated the response to gemfibrozil and found that this medication was

associated with LDL subclass response. Research by Munshi in 2012 (Munshi, 2012) reported that atorvastatin use was associated with poorer outcomes in stroke patients with LPL gene variants. No further significant studies were identified via electronic search in Pubmed beyond these.

Consequently, the body of knowledge from this research project may serve as a notable contributor to the relatively small amount of information known about *LPL*, particularly with its role in the development of MetS and diagnoses of both T2D or CVC as well as preventive and therapeutic measures relating to pharmacogenomics. This includes stratified medicine or SM, as previously introduced in Section 2.A.VII – this is defined as grouping patients based on disease risk or response to treatments (Bell, 2014). In this context, for example, individuals with the *LPL* SNPs, or as described in the medication classes for T2D and CVC, may benefit from SM. SM offers several potential benefits including obvious clinical (primarily patient care) and economic advantages, but it also comes with many challenges or disadvantages – these include timing mismatch (i.e. predictive biomarker science trails the therapeutic,), and that economic value may be questionable given cascades of developmental, regulatory, and commercial considerations (Trusheim and Berndt, 2015, Trusheim et al., 2011). Other difficulties in this field are complex methods and lack of consistency or consensus in terms of definitions and strategies (Attar et al., 2019). There are, however, some propositions on how to address these challenges, such as those from the study of Trusheim, et al., which outlined the use of multiple variable stimulations and the selection of optimal research, development and commercial approaches (Trusheim et al., 2011). Nevertheless, the

applications of this study and the utilization of all research publications on *LPL* to date are of definite importance to the further stratified medicine investigations for the chronic diseases in question.

Precision medicine is another approach, which is a step further considering individualized treatments - this was also introduced in the previous section. There are numerous publications in precision medicine specifically and integrally for obesity, T2D, and CVD. A recent study by Szcaerbinski and Florez in 2023 published in the Lancet Diabetes Endocrinology presented the development of a multi-disease management algorithm, which is primarily obesity-centred but also targets its comorbidities (Szczerbinski and Florez, 2023). This again highlights the centrality of obesity in these diseases and underscores the applicability of precision medicine for obesity even when the targets are, for example, T2D and/or CVD. Another study, which is very relevant to today's world, presented the use or addition of artificial intelligence (AI) to genotyping and deep phenotyping, where AI or machine learning is used for data integration and relationship exploration (Subramanian et al., 2020). The UK Biobank is one of the largest organizational bodies providing a significant platform for such approaches. In addition to whole genome sequencing, one study has also made use of other techniques, such as imaging techniques (e.g. magnetic resonance imaging (MRI)), global metabolomics, a new blood test for prediabetes, echocardiography (ECHO), electrocardiogram (ECG), and cardiac rhythm monitoring for the identification of age-related chronic disease risks, signifying the importance of multidisciplinary approaches in understanding multifactorial disorders (Perkins et al., 2018).

Personalised or precision medicine goes beyond the conventional treatment measures, however, and may also be taken from the preventive approaches. An interesting new study involving personalised nutrition and *LPL* rs268 for the lipid cluster (i.e. one of the study groups in the research project) presented a protocol for a parallel double-blinded randomised intervention trial to investigate biomarker-based nutrition plans for weight loss, with the goal of empowering consumers to prevent diet-related diseases through omics sciences (which the authors called PREVENTOMICS) - this was also the article's title (Aldubayan et al., 2022). The authors concluded that the study results are proof of the use of metabolic and genetic biomarkers in providing personalised dietary treatments for overall health benefits including reduction in body fat mass. In line with this, on the side of lifestyle modification, personalised exercise or physical activity may also be a potential consideration for further study in the management of obesity and comorbidities. In addition, a combination of both lifestyle and therapeutic approaches to the stratification or personalization of treatments may prove to be the optimal strategy in developing treatment modalities for different diseases including obesity, T2D, and CVD, thus requiring multidisciplinary healthcare involvement.

In summary, the outcomes of this research contribute valuably not just on the *LPL* variants studied, the clinical associations presented, and conclusions generated, but also to the current understanding on stratified or personalised medicine for obesity, dyslipidaemia, T2D, and CVC.

5.E. Strengths and Limitations

The main notable strength of this study is the large number of participants included in the data analysed. As a result, several statistical analyses were performed, including the inclusion of individuals with relatively rare *LPL* SNPs. Furthermore, the source of the data, the UKB, is highly reputable source, with regulations and systems in place which are in accordance and in compliance with universal ethical and regulatory guidelines (copies of the UK Biobank Material Transfer Agreement and the University of Staffordshire Research Ethics Proportionate Review are presented in Appendix 7 and Appendix 8, respectively). Although the majority (90.0%) of the population was of British ethnicity, the study also involved other ethnic groups, which may also be an advantage for the applicability of the results. The remaining 10% comprised of African, Bangladeshi, Black or Black British, Carribean, Chinese, Indian, Irish, Mixed, Pakistani, White, White and Asian, White and Black African, White and Black Carribean, and any other (Asian, Black, Mixed, White) background, with 0.41% who preferred not to say, do not know, and with no ethnicity information. These classifications generally correspond to the high-level (five broad groups: White; Asian or Asian British; Black, Black British, Carribean or African; Mixed or Multiple ethnic group; Other ethnic group) and detailed (19 categories) ethnic classifications as being used by the Office for National Statistics (ONS) for England and Wales (eg. per published ONS 2021 census). The British ethnicity specified as majority (i.e. 90.0%) as presented throughout this study was White British, noting that there were, in addition, 11 (<0.10%) individuals specified as White without further category. Nevertheless, these factors result in concrete

outcomes and conclusions overall, which may significantly contribute to the study of obesity and related comorbidities, with a recommendation for future studies on different ethnicities.

One of the limitations of this study is the use of surrogate marker for CVD - that is, CVC, which was collectively presented as heart or vascular problems (HVP) in the UKB data. HVP, as previously discussed, is any or combination of heart attack, angina, stroke, and high blood pressure or hypertension. The National Health Service (NHS) in the UK specifies CVD as four types- coronary heat disease (may be in the form of angina, hear attacks, heart failure), strokes and transchemic-ischemic attack (TIA), peripheral arterial disease, and aortic disease. Although it is clear that the components of HVP are within these types or sub-types, the inclusion of hypertension as an HVP or CVC, which in turn is also a MetS parameter in itself, may be a limitation. It is important to note here, however, that hypertension alone cannot be considered as MetS by itself in binary criteria. In addition, hypertension is known as a major CVD risk factor, with several of the discussed current CVD pharmacotherapy classes directly preventing or treating this condition.

Another minor limitation of the study is the exclusion of serious medical conditions. In the UKB data, the latter was collectively presented based on a binary Yes/No information; the details on how this was generated by the UKB is presented here. In the UKB assessment centre, a verbal interview included specifying "medical conditions" then "non-cancer illness code, self reported." This then points to data coding (i.e. data coding 6 via UKB), which is a list of the codes and meanings (freely retrievable via UKB online site). Non-cancer

medical condition or disability was defined when one or a combination of the list of the serious diseases (classified under infections, immunological/systemic disorders, gynaecology/breast, haematology/dermatology, musculoskeletal/trauma, neurology/eye/psychiatry, endocrine/diabetes, renal/urology, gastrointestinal/abdominal, cardiovascular, and respiratory/ENT) was identified (i.e. “Yes,” pertaining to the presence of other serious non-cancer medical condition or disability). To avoid the effect of those on the results of this study, this group has been collectively excluded in the study. Although some CVD or CVC and diabetes cases were included, given the collective entity for presence of non-cancer disease, these were excluded; only confirmed T2D and CVC diagnosed by a doctor as described in the UKB data were retained as included in the analyses.

Also worth mentioning as a minor limitation is the minimum age of participants that UKB includes (i.e. from 40 years old). In the recent age, more people are receiving a confirmed diagnosis of T2D or CVC at a younger age (i.e. under 40 years old) (Barker et al., 2022). In hindsight, had the younger age group been included, the primary impact might have been on the age of diagnosis and modifiable risk factors rather than genetics, which remain constant for any individual. Nevertheless, research on younger population with the *LPL* SNPs evaluated in this study would certainly add value to the findings reported on this project.

6. CONCLUSIONS

This retrospective study investigated the association between MetS-associated *LPL* SNPs and the progression to T2D and CVC. Significant

associations were identified between the SNPs rs268, rs11542065, rs116403115, rs118204057, rs118204061, rs144466625, and rs54764495 and diagnosis of both T2D and CVC. In addition to previous publications with similar findings as discussed in earlier sections, this strengthens the crucial role that these SNPs play in the pathogenesis of T2D and CVC, irrespective of obesity.

There were statistically significant differences in weight, BMI, diastolic BP, total lipids in lipoprotein, HbA1c, WC, HDL, and LDL between groups, suggesting that these variants may have different effects on these clinical parameters. It is important to note, however, that the results of this study do not signify causation, but only highlight the significant differences in the specified clinical outcomes in the groups. BMI and WC were found to be significantly higher in individuals who were diagnosed with both T2D and CVC. The role of obesity in the development of these two diseases, as well as the metabolic interrelatedness of obesity, T2D, and CVC are once more evident in these results. It is paramount to combat obesity or excess weight as preventive and treatment strategies, with or without pharmaceutical interventions, for both T2D and CVC. When sexes were compared, men who were diagnosed with T2D and CVC had slightly increased BMI and WC compared to women, showing that women may be more prone to developing T2D and CVC with lesser weight gain. Ethnicity may play a crucial role in this aspect, however, as well as other factors, and further investigations are needed to elucidate the outcome differences between sexes.

Models for predicting confirmed T2D and CVC diagnosis were explored using logistic regression on UKB data. The addition of genetic contribution

enhanced the AUC values; therefore, the models would better predict confirmed diagnosis of T2D or CVC. The additive effect of the *LPL* SNPs and relevant PRS was more pronounced in the CVC than in the T2D model. These results are supportive of the important role that genetics and genomics play in improving the prediction of T2D and CVC progression in addition to the much earlier established clinical parameters. Improved prediction measures are beneficial to help prevent the development of these diseases among high risk individuals.

Of the variants studied, rs547644955 had major significance for both T2D and CVC diagnoses, with an AUC of 1.0 and .910, respectively. The SNPs rs116403115 and rs118204057 both had an AUC of 1.0 for T2D diagnosis. These findings are highly important, and these SNPs are worth investigating much further in different races and larger cohorts to aid in better understanding on how they affect the development of T2D and CVC, in general. In particular, the genetic contribution of these SNPs versus the environmental and epigenetic factors of T2D and CVC pathogenesis need to be evaluated in varying populations.

Collectively, the results of this study have implications for stratified or personalised medicine amongst individuals with the investigated *LPL* SNPs. However, additional research is required to further elucidate the effects of these SNPs in the development of MetS and other obesity-related diseases.

7. PHILOSOPHICAL AND ETHICAL ISSUES

This part of the work is a reflective analysis on the major concerns surrounding philosophical and ethical issues in biobanks for research

purposes, particularly clinical research, with a focus on the UK Biobank Ethics and Governance Framework (EGF), as UKB data was used for this study; no other additional source of data or material was used. It is composed of an outline of the UKB EGF and critical appraisal.

7.A. UKB Salient Features

7.A.I. Responsibilities of UKB Staff and Stakeholders

The UK Biobank REG (research ethics and governance) framework has made clear-cut definitions, conditions, and details regarding the biobank's responsibilities to pertinent stakeholders as aforementioned, namely: relationship with participants, relationship with research users, and relationship with society. In fact, they have highlighted the significance of this by making them the main headings (3 out of 4 primary discussion points). Paramount subheadings are laid out with succinct yet sufficient information to serve the purpose of this section. They have stressed here that relationships with these sectors are the most important and is at the core of the biobank's commitment. This is likewise an efficient way to present an organized content, with the target audience finding it easy to locate what they seek. Hence, this particular feature of the UK Biobank is commendable and is recommended to be a model outline for other biobank REG frameworks to follow.

7.A.II Access

Decisions on access and use are stipulated in the UK Biobank framework, under relationship with research users, below research access to data and samples. Proposals are thoroughly reviewed and ensured that they

are consistent with the participants' consent and the biobank framework, and with relevant ethics approval. It is specified that the Board of Directors will have the overall decision-making authority, but may also delegate this responsibility to suitable groups or persons as necessary, such as an Access Committee. Assurance of public interest is underscored, with existing access policies and procedures to follow, handling of conflicts and prioritisation of the use of samples, as well as proper explanation to the public and participants. This, in this regard, is a good model for other biobanks to shadow.

7.A.III. Participant selection

The Biobank sought to recruit not only UK locals, but also various ethnicities in the country to maximize the potential of the project. In lieu of this, translations of relevant documents were made available for this purpose. This has a huge impact for epidemiological studies, as well as epigenetics and the role of environment on various diseases. Societal and cultural aspects are likewise considered, and may also be ground-breaking, not jut for biomedical studies, but also for social science investigations. This feature of the UK Biobank makes it a truly global resource for many scientists across the globe, in addition to the number of participants half a million strong.

7.B. On Research Ethics and Governance of Biobanks

Of the factors surrounding ethical concerns with biobanks, the primary aspects most often discussed in literature are: (1) informed consent and ethical approval; (2) sample and data collection; (3) data privacy and biospecimen security; and (4) policies and procedures (Caulfield and

Murdoch, 2017, Goisauf et al., 2019, Langhof et al., 2018, McGregor and Ott, 2019, Mikkelsen et al., 2019, Morrison et al., 2017, Rheeder, 2017, Thompson and McNamee, 2017, Verlinden et al., 2016).

7.B.I. Informed consent (IC) and ethical approval

Six applicable consent models (verbal consent, blanket consent, broad consent, meta consent, dynamic consent, and waived consent) may be used for research; of which, broad consent that needs to be deep, has been recommended for biobank use (Mikkelsen et al., 2019, Thompson and McNamee, 2017).

This broad but deep consent recommended for biobank use requiring the definition of unambiguous conditions (i.e. limitations, when to re-contact for special projects involving the use of data and biospecimen) appears to be agreeable and recommendable. Although no matter how deep this broad consent may be, there will still be cases or projects where obscurities may be apparent; in this case, it should be defined that participants need to be re-contacted and that re-consent will be sought. Conditions should also be precise when assent is required (specific age cut-off) or dealing with various vulnerable groups (properly enumerated, classified, and defined; e.g. minority ethnic/institutionalized groups, persons with disability, socially excluded groups). In this manner, autonomy and respect for individuals are somehow served to a higher extent. Additionally, the principles of justice, particularly when regarded at the context of beneficence, are evidently practiced in that the value to society far outweighs the burden to the participants in ways where each individual’s rights are also considered.

The approval of the IC or assent, study protocol, and other relevant documents pertaining to a research project is widely known to be one of the main responsibilities of the Research Ethics Committee (REC; mostly in European countries) or the Institutional Review Board (IRB, primarily in the United States of America (USA)), or similar group, which is also the case in biobanks. From here alone, disparity in stakeholder's appellation is evident, which may possibly lead in itself to confusion in the global context. Furthermore, the role of the REC/IRB in biobanks may also be challenged in that unlike conventional research where it is primarily "most active" so to speak, in the initial stages of research, the need for ongoing monitoring for biobanks appear essential. This is because numerous research projects may be proposed hence applied for by various researchers from a similar data set, unlike conventional research centres where RECs/IRBs know the organization's research projects fairly well. Stringent monitoring and control are thereby suggested.

7.B.II. Sample and data collection

The research participants' information sheet containing details of the study as may be suggested by the National Research Ethics Service (NRES) is read by the participants before the IC is presented, and normally includes sample and data collection specifics (Kirkby et al., 2012). Minimized risks, in agreement with the non-maleficence requisite of research, may be evident for basic data collection, anthropometric measurements, and non- to minimally invasive sample collection methods (e.g. urine and blood samples). On the same note, this is debatable for processes, which may involve harm or

increased discomfort to participants such as biopsies or lumbar puncture for specimen collection. If the samples are, however, derived from excess clinical specimens (i.e. from biorepositories) and consent is derived from the patients for research purposes, then that would be an opposite scenario. Nevertheless, proper benefit to risk analysis measures should be in place to assure solidarity in the processes involved.

In biobanks, sample collection may be a one-time event, but data collection may be longitudinal since participants' succeeding records can be derived from databases per participants' consents. In developed countries, for example, where medical data is centralized and made accessible as allowed, ethical concerns may be raised: "Is justice indeed served when the biobank can freely retrieve these 'non-basic' information, particularly in 'what-if' cases (e.g. What if the participant "forgot" about the consent and with that specific medical detail about himself, he would really wanted to keep it private? What if the patient instead wanted that unique genetic result from that collected blood sample be kept to himself, should he just have known?). Scenarios like these are typically not addressed, but anyone can argue that it is just impossible to tackle every possible case. Perhaps a group of experts from different relevant disciplines and from different regions may come up with a consensus to particularly tackle concerns as these, in general but comprehensive ways, as suggested by literature (Bernasconi et al., 2020).

7.B.III. Data privacy and biospecimen security

Privacy and security of patient details, medical reports, biospecimen and any aspect pertaining to the participants are crucial for many facets of

research, but is particularly applicable to data and sample management. In practice, access to information from electronic databases such as patient's clinical profiles or laboratory information management systems (LIMS) is often restricted to limited personnel within the research department whose responsibilities are defined and are often documented. Additionally, the use of passwords on computers and applications or systems is probably the most utilized method for this purpose. For biospecimen, similar access requirements are observed. Governing bodies typically composed of a team of experts who are responsible for reviewing, preparing, and granting access requests for data and/or samples, often referred to as custodians or gatekeepers, is pre-identified (Langhof et al., 2018).

Although it may be a given that these personnel are properly trained, it would be appropriate for transparency's sake to be explicit about procedural details on this matter, primarily for the information of the participants. The involvement of rigorous lines of approvals before data or samples are released, should also be stipulated. An easily understandable flow chart may be of good use for this purpose, with the appropriate approving post or unit specified. This may be included in the information sheet for the participants, as part of the invitation to participate which goes along with the IC form, or in the biobank website where it is accessible to them and to the public.

Another major element worth mentioning here, significantly applicable to biobanks and global networks, is data and sample exchange (DSE) across continents, with governing legal and ethical requirements reported by the International Clinical Trial Center Network (ICN) (Bernasconi et al., 2020). In addition to suggestions in literature to address the lack of globally harmonised

facilitation of DSE, perhaps the first issue to be resolved here is the gravity of necessity for this exchange. If generalizations or conclusions, for instance, are desired from various ethnic groups, review articles from already published research studies may work instead. This way, duplication of work, incurred expenses, and other associated unnecessary inputs will be avoided.

7.B.IV. Policies and procedures

The lack of globally harmonized policies on the access of bio-samples is also recognized as an ensuing concern, despite availability of international guidelines (Langhof et al., 2018). In addition, difficulty to gain access of these samples, as well as custodianship disputes, have been reported in several reports (Langhof et al., 2018, Verlinden et al., 2016). The currently available and internationally recognized or adapted legislations or recommendations for medical research which may also be applied to biobanks include the the Declaration of Helsinki, the Declaration of Taipei, the Council for International Organizations of Medical Sciences (CIOMS) international ethical guidelines for biomedical research involving human subjects, the Organization for Economic Co-operation and Development (OECD) Guidelines on human biobanks and genetic research databases, and the Universal Declaration of Bioethics and Human Rights (UDBHR) of the United Nations Educational, Scientific and Cultural Organization (UNESCO) which is in line with the Declaration on Ethical Considerations Regarding Health Databases and Biobanks of the World Medical Association (WMA) (Bernasconi et al., 2020, Rheeder, 2017). In Europe, the General Data Protection Regulation (GDPR) is accepted as the harmonized legislation, although some deviations are

allowed depending on the member states' discretion such as permissible range of minimum age for consent (Goisauf et al., 2019, Morrison et al., 2017).

In addition to these established principles, new and developing guidelines and standards are getting recognized such as the International Society for Biological and Environmental Repositories (ISBER) Best Practice and the International Organization for Standardization (ISO) biobanking standards (Henderson et al., 2019). Although not a guideline on its own, the proposal of the Bioresource Research Impact Factor (BRIF) is also worth mentioning here, which primarily aims to recognize the efforts of individual biobanks; this however, is not anywhere near wide acceptance (Langhof et al., 2018). In the UK, the Human Tissue Authority (HTA) is the governing body that issues licenses to biobanks, the Research Ethics Committees reviews projects, while data access committees may be in charge with data requisitions; however, there is no law which is specific for biobanks (Kaye et al., 2016). In addition, the question on which committee should approve requests for access requests remains controversial (Verlinden et al., 2016).

Overall, the main challenge that biobanks seem to face appears to be in the standardization and unification of policies, procedures, data management, and general legislative and ethical framework for all countries to follow (Bernasconi et al., 2020, Yang et al., 2016). To start off, suggestions from scholars include adaptation of outcomes from previous recommendations such as those from the Global Initiative for the Ethical Use of Human Specimens (GIFT), creation of a group of experts represented by various regions around the globe for preparation of the legislations, and

agreement on definition of biobank-specific terms (Bernasconi et al., 2020).

However, this undoubtedly comes with major challenges. Foremost, which organization or body would volunteer to initiate? Is it, in the first place, allowed to? From which country? How about funding and resources? And if any arises, how can cooperation be assured from many existing biobanks across the globe where various existing country-specific laws and regulations have also to be considered and approved first despite sincere willingness to participate? The role and passion of biobank forerunners here are hence extremely crucial to bring this to fruition. Although this may entail such a long, laborious process as with many other endeavors, this too shall materialize given the commitment of even a select few. It is important to stress here alike, that the voice of representatives from various regions, regardless of economic, cultural, or social background should always be heard and considered, to bar discrimination and for a truly harmonized consensus.

7.C. The UK Biobank Ethics and Governance Framework
Rooms for Improvement

7.C.I. Data Privacy

Concerns regarding data privacy in the UK Biobank may be similar to other biobanks in general in that ongoing data collection from centralized medical databases, from the NHS in this case, maybe retrieved as per participants' consents. The consent used here is clearly a broad one, and given the wide-ranging scope of NHS record systems (e.g. GP, hospital, dental, prescription) it can be said that it is a very broad IC indeed. The recommended broad but deep consent may be applicable here, to address elements of justice and autonomy among the participants (Mikkelsen et al.,

2019). Although the framework mentions that participants will be informed about progress with accessing of records, it also specifies that the consent shall cover access to full records. This is quite contradictory. When then will the participants be informed? This is blurry which requires clarification and further details.

7.C.II. Inclusion Criteria

The UK Biobank recruited participants with age 40-69 years only and excluded some vulnerable groups (such as those with diminished mental capacity, the sick, or those uncomfortable with any of the conditions of participation). Hence the general concerns about assent and these specific groups are addressed. However, other groups, which may not be eligible should have been further classified. The framework likewise specified that actions will be in accordance with the Data Protection Act as well as other legislations, and also stated that staff were trained to judge the capacity of potential participants' eligibility to be included in the project. However, details of this training as well as proficiency outcomes of the training should have been laid out. Was there a relevant examination for this training to assess the competency of the staff for this purpose? This shall be a significant contributing feature of the framework if indeed included.

7.C.III. Publication and Finances

The UK Biobank's REG framework is not very clear as well regarding publication. Although it is stated that outputs should be published in peer-reviewed journals, details on selection of proposals as well as pre-publication

verification (i.e. data analyses and interpretation in addition to peer reviews for enhanced quality assurance and control) are lacking. Adding this information to the framework may prove beneficial. For monetary gain to participants, they have made it precise that no monetary benefits will be given to research subjects except for request on reimbursement of justifiable expense such as travel. Any income generated from the resource, it has also been stated, will be re-invested back to the biobank. A little more detail on this may be called for, but should also be made available upon request of the different stakeholders for transparency and accountability.

7.C.IV. Partnership with DNANexus

In September 2021, all UKB data and data access have been transferred to DNANexus, a cloud-based data management system. This has posed difficulties amongst users that an online community has been created where the UKB users may be able to ask questions and help each other to resolve issues encountered. This is in addition to several online tutorial guides that were often in the form of videos and written texts. Online assistance via email to the UKB DNANexus group was also made available. On another note, security issues in this type of system should be stringent given that very large confidential data are at stake.

8. REFLECTIVE ANALYSIS

8.A. Thesis Preparatory Requirements

The Professional Doctorate in Healthcare Science curriculum from University of Staffordshire, formerly Staffordshire University, includes a

doctoral research thesis and viva as one of the later requirements for degree completion. As a ProfDoc (professional doctorate) program, it differs from a PhD (Doctor of Philosophy), in that it is mainly aimed to improve clinical practice (Ellis, 2007) while a PhD mainly focuses on academic research. There appears to be more varied perception on doctorates rather than PhD though, and the applicability of doctoral degrees to clinical settings has been suggested to have the need for further clarity and be more streamlined (Ellis, 2007, Rosenfeld et al., 2022). Nevertheless, the Professional Doctorate in Healthcare Science offered by the university is aimed at developing leadership skills in the area of students' expertise, which is surely valuable particularly to individuals who wish to pursue higher-level administrative roles in the field of healthcare science and related disciplines.

As the research thesis topic is usually chosen within the field of expertise of the student, it is often a norm to derive the data from his/her place of work or employment- this is also laid out in the online information for the program within the university website. There are cases though, that, this may not be feasible due to data sharing restrictions from the employer, for instance, which has been the case in this project. Thus, the use of the UKB data has been proven to be a valuable alternative, including the discounted access fee that UKB offers to students completing theses in universities around the world.

As expected, there were steps and approvals involved in the data access from UKB as well as the need for a Proportionate REC approval from the university. There have been not much bottlenecks nor issues involved in these processes, which included signatories of supervisors and university

representative, attesting administrative organization of both UKB and University of Staffordshire in this aspect - it has been a smooth process with these steps.

The particular knowledge gained from these steps include biobank access processes, research proposal preparation both for biobank and REC, ethical and practical considerations, high-level coordination with various stakeholders, efficient communication, and effective organization.

8.B. Conduct of the Study: Bottlenecks and Challenges

The major issue encountered in the conduct of this project was in relation to data access, particularly data download. Initially, data from the UKB online resource has been used; however, how the data can be translated to readable or usable format has been a chief concern, where how-to guides did not seem to work. In 2021, as described, the data management system has been transferred via the cloud-based platform from DNANexus – this has posed another layer of issues which caused further delay in the progress of the study. As the system was new, many researchers accessing the data faced various difficulties that personally, different resources, DNANexus and UKB access teams, as well as resource persons internationally have been reached out for assistance. With persistence and diligence as well as support and encouragement from supervisors, the data required for the study has been eventually properly accessed and downloaded.

There were also other personal situations and circumstances that were not necessarily hindrances, but are a part of life, wherein formal pauses from the study were needed. This included pregnancy and childbirth, and migration

to a new country as well as a new job. Although this caused delay in the completion of the program, it has been proven to be a healthy approach in terms of physical and mental well-being, finances and resources, balance with family and other relational requirements, and other factors. The support of the supervisors as well as the university's program for intermission during the course of the study have been demonstrated to be helpful and are well-appreciated.

Also a part of life in general and career move in particular, the transfer of two main supervisors to different universities while the thesis is being completed has happened. This did not present significant challenge though, largely because the main supervisor has been very hands-on as usual. A new within-the-university supervisor has also been assigned, and has been very helpful likewise. Hence, collaboration and advice from three supervisors in three large universities have made the supervision and advice even stronger and again, well-treasured. The unswerving support of the university administration and the new program lead has likewise been consistent for which any student would be grateful. In addition, although on a different note, other forms of support from the university are also appreciated, including the provision of the SPSS program license to students for data analysis.

On the aspects laid out in this section, several skills have been honed including that of adaptability, flexibility, time management, diligence and coordination. These skills will surely be beneficial to perform duties as a researcher and in performing leadership positions, hence will surely be applicable in real-world scenarios not just in the workplace, but also in everyday life.

8.C. Research Study: Notes, Impact, and Future Enablers

Although the concept of MetS is known to have begun with its introduction by G.M. Reaven in the late 1980s, its history dates back way earlier as literature presents accounts of this disease for approximately a century (Matfin, 2010). The limited MetS GWAS reports appears to convey a message that stern scientific attention to MetS may indeed be lacking, particularly compared to studies involving its individual components such as obesity, which has much higher number of GWAS reports. This is considering that the first GWAS was published in 2005 (Loos, 2020).

The country of Korea appears to be the strongest player in MetS GWAS with the highest number of reports. Although MetS prevalence is growing in this country (Lim et al., 2011), this is, however, true globally. The challenge lies in the different MetS definition or criteria in assessing prevalence in different countries. Moving forward, the need for a unified MetS global definition may be found particularly beneficial. How this can be attained, may indeed be challenging, however, because organizations that proposed these are highly credible authorities in their own entitlements. For these institutions to, in fact, agree with a MetS definition consensus may be far-fetched, but as with other proposals, essentially not impossible. A strong will and commitment from international advocates may indeed be called for.

Nevertheless, Korea seems to be on the right track in the study of MetS in this sense. Other countries, particularly the first world which are resource-rich and are already doing intensive molecular studies may need to step up further. The absence of GWAS from numerous other nationalities is also

worth noting, with only very few countries and regions represented as yet accounting for the twelve reports from seven countries. Surprisingly, the total number of seven studies from Asia is the highest, with Europe as second with three studies, and North America with one study. Therefore, among the inhabited continents, no studies have been reported yet from South America, Africa, and Oceania. Again, this may be a call to countries such as Australia to pursue MetS GWAS studies.

On another note, advantages of the use of biobanks and similar large data repositories are evident with the largest number of novel SNPs reported from the UK Biobank. Tapping on these resources may be found particularly helpful in the greater understanding of MetS genetics, taking into account various ethnicities and populations. Certain genetic considerations may also be investigated such as the female-specific study reported from Korea. With over 120 biobanks in the world, the biggest ones such as Biobank Graz having millions of derived human samples, the possibilities are seemingly endless and therefore much more can certainly be done. Circulating adverts to researchers, students, the academe, and private organizations around the world about these resources may be found highly advantageous.

Another observation of note is how GWAS results are presented, not just for MetS studies, but also generally, for other diseases. To have consistent content, which is normally presented in tables may be found useful for the use of these resources, such as retrieving all SNPs reported globally. For instance, required parameters are recommended to always include gene or nearest gene, chromosome, and position at the least, as much as possible.

Consistent, stringent quality control procedures for GWAS are also expected to be in place and explicitly specified or referred to in published manuscripts.

On top of above notes and recommendations, it is important to mention here that although GWAS investigate genetic contributions of SNPs to human disorders, prediction of disease risk using GWAS-identified SNPs or SNP combinations is not well established (Patron et al., 2019). This is therefore another area that needs to be further explored, including that for MetS. Moreover, because MetS is linked with increased risk of T2D and CVC, the occurrence of the latter two chronic diseases subsequently after being diagnosed with MetS and/or having been tested with MetS-associated genotype may be interesting to examine.

On pharmacogenetics, the increasing availability and decreasing costs of genetic analysis may pave the way for greater usefulness of this molecular level-based technology for this purpose in the future (Gloyn and Drucker, 2018). However, it does not come without major barriers and considerable bottlenecks. One of the foremost necessities is the resolution of the problem with big data and its translation to actionable clinical decision system (Aghaei Meybodi et al., 2017). In addition, there is a big gap in our knowledge base regarding treatments in population subgroups such as adolescents, the elderly, and during pregnancy (Gloyn and Drucker, 2018). The use of other ‘omics’ (e.g. metagenomics and metabolomics) should likewise further be used to aid our understanding of other complex interacting networks in order to implement cost-effective and best-tolerated treatment strategies (Aghaei Meybodi et al., 2017, Scheen, 2016). Nevertheless, personalised medicine in T2D management is a truly promising tool for T2D treatment and the fruition

of its role in bedside patient care are dependent on extensive scientific investigations of varied specializations which are currently evident worldwide, and should also be similarly applicable to CVC.

Although it may be expected that the implementation of PM for T2D and CVC management in the future may begin to be a routine practice in first world countries, slow progress may be observed in third world countries primarily due to lack of resources and expertise. The role of healthcare front-runners, medical institutions, and government leaders shall be called for if so, focusing on strengthening linkages and partnerships with various international organizations including the private sector. Retention of scientists to their homeland should likewise be reinforced for this purpose. This is something that is somehow personally important, coming from a third-world country that due to lack of opportunities and promise of better generational advantages, have prompted the desire to move. Indeed, there are so many issues in this world – health, political, societal, the list goes on; nonetheless, it is with great hopes that the outputs of this research study would pave a way, in one way or another, in simpler or bigger ways, a healthier, better world.

8.D. The Course, the Degree: What's Next?

The findings from this study, the theoretical and practical learning gained from the thesis and the other courses from this program, the acquaintances, and all others in-between: what is next, what is next indeed?

Without a doubt, the writing of the thesis has been the most challenging, laborious, and time-consuming part of the program. But it can be said that the benefits arguably go beyond the knowledge gained and skills

developed- it has paved the way for the development of a well-rounded professional and individual. There is so much more to it than the title, the single components of the curriculum, or the anticipation of rewards or return of all that have been invested – collectively, it is somehow difficult to describe, but overall this achievement is something worth being very, very thankful for. It has been a huge blessing to embark on this journey, and it is with great hope and excitement to soon embody what it is to be, what it means to be, and be of significance as, a Doctor in Healthcare Science.

2016a. Pharmacologic Management of Type 2 Diabetes: 2016 Interim Update. *Can J Diabetes,* 40**,** 484-486.

2016b. Standards of Medical Care in Diabetes-2016: Summary of Revisions. *Diabetes Care,* 39 Suppl 1**,** S4-5.

ABBATE, R., STICCHI, E. & FATINI, C. 2008. Genetics of cardiovascular disease. *Clin Cases Miner Bone Metab,* 5**,** 63-6.

AGHAEI MEYBODI, H. R., HASANZAD, M. & LARIJANI, B. 2017. Path to Personalized Medicine for Type 2 Diabetes Mellitus: Reality and Hope. *Acta Med Iran,* 55**,** 166-174.

ALDUBAYAN, M. A., PIGSBORG, K., GORMSEN, S. M. O., SERRA, F., PALOU, M., MENA, P., WETZELS, M., CALLEJA, A., CAIMARI, A., DEL BAS, J., GUTIERREZ, B., MAGKOS, F. & HJORTH, M. F. 2022. Empowering consumers to PREVENT diet-related diseases through OMICS sciences (PREVENTOMICS): protocol for a parallel double-blinded randomised intervention trial to investigate biomarker-based nutrition plans for weight loss. *BMJ Open,* 12**,** e051285.

ALI, O. 2013. Genetics of type 2 diabetes. *World J Diabetes,* 4**,** 114-23.

ALINAGHIAN, N., ABDOLLAHI, E., TORAB, M., KHODAPARAST, M., ZAMANI, F. & RAHIMI-MOGHADDAM, P. 2019. Gender-related relation between metabolic syndrome and S447X and HindIII polymorphisms of lipoprotein lipase gene in northern Iran. *Gene,* 706**,** 13-18.

ANNA, A. & MONIKA, G. 2018. Splicing mutations in human genetic disorders: examples, detection, and confirmation. *J Appl Genet,* 59**,** 253-268.

ARIZA, M. J., SÁNCHEZ-CHAPARRO, M. A., BARÓN, F. J., HORNOS, A. M., CALVO-BONACHO, E., RIOJA, J., VALDIVIELSO, P., GELPI, J. A. & GONZÁLEZ-SANTOS, P. 2010. Additive effects of LPL, APOA5 and APOE variant combinations on triglyceride levels and hypertriglyceridemia: results of the ICARIA genetic sub-study. *BMC Med Genet,* 11**,** 66.

ARNOLD, N. & KOENIG, W. 2021. Polygenic Risk Score: Clinically Useful Tool for Prediction of Cardiovascular Disease and Benefit from Lipid-Lowering Therapy? *Cardiovasc Drugs Ther,* 35**,** 627-635.

ATTAR, S. G., POUSTIE, V. J., SMYE, S. W., BEETY, J. M., HAWCUTT, D. B., LITTLEWOOD, S., ONI, L., PIRMOHAMED, M. & BERESFORD, M. W. 2019. Working together to deliver stratified medicine research effectively. *Br Med Bull,* 129**,** 107-116.

AUTON, A., BROOKS, L. D., DURBIN, R. M., GARRISON, E. P., KANG, H. M., KORBEL, J. O., MARCHINI, J. L., MCCARTHY, S., MCVEAN, G. A. & ABECASIS, G. R. 2015. A global reference for human genetic variation. *Nature,* 526**,** 68-74.

BABAEV, V. R., PATEL, M. B., SEMENKOVICH, C. F., FAZIO, S. & LINTON, M. F. 2000. Macrophage lipoprotein lipase promotes foam cell formation and atherosclerosis in low density lipoprotein receptor-deficient mice. *J Biol Chem,* 275**,** 26293-9.

BARKER, M. M., ZACCARDI, F., BRADY, E. M., GULSIN, G. S., HALL, A. P., HENSON, J., HTIKE, Z. Z., KHUNTI, K., MCCANN, G. P., REDMAN, E. L., WEBB, D. R., WILMOT, E. G., YATES, T., YEO, J., DAVIES, M. J. & SARGEANT, J. A. 2022. Age at diagnosis of type 2 diabetes and cardiovascular risk factor profile: A pooled analysis. *World J Diabetes,* 13**,** 260-271.

BELL, C. G., FINER, S., LINDGREN, C. M., WILSON, G. A., RAKYAN, V. K., TESCHENDORFF, A. E., AKAN, P., STUPKA, E., DOWN, T. A., PROKOPENKO, I., MORISON, I. M., MILL, J., PIDSLEY, R., DELOUKAS, P., FRAYLING, T. M., HATTERSLEY, A. T., MCCARTHY, M. I., BECK, S. & HITMAN, G. A. 2010. Integrated genetic and epigenetic analysis identifies haplotype-specific methylation in the FTO type 2 diabetes and obesity susceptibility locus. *PLoS One,* 5**,** e14040.

BELL, J. 2014. Stratified medicines: towards better treatment for disease. *Lancet,* 383 Suppl 1**,** S3-5.

BELL, J. I. 2002. Single nucleotide polymorphisms and disease gene mapping. *Arthritis Res,* 4 Suppl 3**,** S273-8.

BENAIGES, D., GODAY, A., PEDRO-BOTET, J., MAS, A., CHILLARON, J. J. & FLORES-LE ROUX, J. A. 2015. Bariatric surgery: to whom and when? *Minerva Endocrinol,* 40**,** 119-28.

BERNASCONI, L., SEN, S., ANGERAME, L., BALYEGISAWA, A. P., HONG YEW HUI, D., HOTTER, M., HSU, C. Y., ITO, T., JORGER, F., KRASSNITZER, W., PHILLIPS, A. T., LI, R., STOCKLEY, L., TAY, F., VON HEIJNE WIDLUND, C., WAN, M., WONG, C., YAU, H., HIEMSTRA, T. F., URESIN, Y. & SENTI, G. 2020. Legal and ethical framework for global health information and biospecimen exchange - an international perspective. *BMC Med Ethics,* 21**,** 8.

BODEN, G. & SHULMAN, G. I. 2002. Free fatty acids in obesity and type 2 diabetes: defining their role in the development of insulin resistance and beta-cell dysfunction. *Eur J Clin Invest,* 32 Suppl 3**,** 14-23.

BOS, G., SCHEFFER, P. G., VIEIRA, D., DEKKER, J. M., NIJPELS, G., DIAMANT, M., TEERLINK, T., STEHOUWER, C. D., BOUTER, L. M., HEINE, R. J. & JANSEN, H. 2004. The relationship of lipoprotein lipase activity and LDL size is dependent on glucose metabolism in an elderly population: the Hoorn Study. *Diabetes Care,* 27**,** 796-8.

BOZINA, T., SIMIĆ, I., LOVRIĆ, J., PEĆIN, I., JELAKOVIĆ, B., SERTIĆ, J. & REINER, Z. 2013. Effects of lipoprotein lipase and peroxisome proliferator-activated receptor-gamma gene variants on metabolic syndrome traits. *Coll Antropol,* 37**,** 801-8.

BRISSON, D., LEDOUX, K., BOSSÉ, Y., ST-PIERRE, J., JULIEN, P., PERRON, P., HUDSON, T. J., VOHL, M. C. & GAUDET, D. 2002. Effect of apolipoprotein E, peroxisome proliferator-activated receptor alpha and lipoprotein lipase gene mutations on the ability of fenofibrate to improve lipid profiles and reach clinical guideline targets among hypertriglyceridemic patients. *Pharmacogenetics,* 12**,** 313-20.

BROEKSTRA, R., ARIS-MEIJER, J., MAECKELBERGHE, E., STOLK, R. & OTTEN, S. 2019. Trust in Centralized Large-Scale Data Repository: A Qualitative Analysis. *J Empir Res Hum Res Ethics***,** 1556264619888365.

BROUSSEAU, M. E., GOLDKAMP, A. L., COLLINS, D., DEMISSIE, S., CONNOLLY, A. C., CUPPLES, L. A., ORDOVAS, J. M., BLOOMFIELD, H. E., ROBINS, S. J. & SCHAEFER, E. J. 2004. Polymorphisms in the gene encoding lipoprotein lipase in men with low HDL-C and coronary heart disease: the Veterans Affairs HDL Intervention Trial. *J Lipid Res,* 45**,** 1885-91.

BROWN, A. E. & WALKER, M. 2016. Genetics of Insulin Resistance and the Metabolic Syndrome. *Curr Cardiol Rep,* 18**,** 75.

BUCHWALD, H. & BUCHWALD, J. N. 2019. Metabolic (Bariatric and Nonbariatric) Surgery for Type 2 Diabetes: A Personal Perspective Review. *Diabetes Care,* 42**,** 331-340.

CAHUA-PABLO, J., CRUZ, M., MÉNDEZ-PALACIOS, A., ANTÚNEZ-ORTIZ, D. L., VENCES-VELÁZQUEZ, A., DEL CARMEN ALARCÓN-ROMERO, L., PARRA, E. J., TELLO-FLORES, V. A., LEYVA-VÁZQUEZ, M. A., VALLADARES-SALGADO, A., PÉREZ-MACEDONIO, C. P. & FLORES-ALFARO, E. 2015. Polymorphisms in the LPL and CETP Genes and Haplotype in the ESR1 Gene Are Associated with Metabolic Syndrome in Women from Southwestern Mexico. *Int J Mol Sci,* 16**,** 21539-54.

CASTELLANO-CASTILLO, D., MORENO-INDIAS, I., FERNÁNDEZ-GARCÍA, J. C., ALCAIDE-TORRES, J., MORENO-SANTOS, I., OCAÑA, L., GLUCKMAN, E., TINAHONES, F., QUEIPO-ORTUÑO, M. I. & CARDONA, F. 2018. Adipose Tissue LPL Methylation is Associated with Triglyceride Concentrations in the Metabolic Syndrome. *Clin Chem,* 64**,** 210-218.

CAULFIELD, T. & MURDOCH, B. 2017. Genes, cells, and biobanks: Yes, there's still a consent problem. *PLoS Biol,* 15**,** e2002654.

CHAHWALA, V. & ARORA, R. 2009. Cardiovascular manifestations of insulin resistance. *Am J Ther,* 16**,** e14-28.

CHATTERJEE, S., KHUNTI, K. & DAVIES, M. J. 2017. Type 2 diabetes. *Lancet,* 389**,** 2239-2251.

CHOBOT, A., GOROWSKA-KOWOLIK, K., SOKOLOWSKA, M. & JAROSZ-CHOBOT, P. 2018. Obesity and diabetes-Not only a simple link between two epidemics. *Diabetes Metab Res Rev,* 34**,** e3042.

CHOI, S. W., MAK, T. S. & O'REILLY, P. F. 2020. Tutorial: a guide to performing polygenic risk score analyses. *Nat Protoc,* 15**,** 2759-2772.

CHULUUN-ERDENE, A., SENGERAGCHAA, O., ALTANGEREL, T. A., SANJMYATAV, P., DAGDAN, B., BATTULGA, S., ENKHBAT, L., BYAMBASUREN, N., MALCHINKHUU, M. & JANLAV, M. 2020. Association of Candidate Gene Polymorphism with Metabolic Syndrome among Mongolian Subjects: A Case-Control Study. *Med Sci (Basel),* 8.

COELHO, O. G. L., DA SILVA, B. P., ROCHA, D., LOPES, L. L. & ALFENAS, R. C. G. 2017. Polyunsaturated fatty acids and type 2 diabetes: Impact on the glycemic control mechanism. *Crit Rev Food Sci Nutr,* 57**,** 3614-3619.

COOPER, D. N. 2010. Functional intronic polymorphisms: Buried treasure awaiting discovery within our genes. *Hum Genomics,* 4**,** 284-8.

CORDOVADO, S. K., HENDRIX, M., GREENE, C. N., MOCHAL, S., EARLEY, M. C., FARRELL, P. M., KHARRAZI, M., HANNON, W. H. &

MUELLER, P. W. 2012. CFTR mutation analysis and haplotype associations in CF patients. *Mol Genet Metab,* 105**,** 249-54.

CRUZ, W. S., KWON, G., MARSHALL, C. A., MCDANIEL, M. L. & SEMENKOVICH, C. F. 2001. Glucose and insulin stimulate heparin-releasable lipoprotein lipase activity in mouse islets and INS-1 cells. A potential link between insulin resistance and beta-cell dysfunction. *J Biol Chem,* 276**,** 12162-8.

CUMMINGS, D. E. & RUBINO, F. 2018. Metabolic surgery for the treatment of type 2 diabetes in obese individuals. *Diabetologia,* 61**,** 257-264.

CZYZEWSKA, M., WOLSKA, A., CWIKLIŃSKA, A., KORTAS-STEMPAK, B. & WRÓBLEWSKA, M. 2010. [Disturbances of lipoprotein metabolism in metabolic syndrome]. *Postepy Hig Med Dosw (Online),* 64**,** 1-10.

DAL CANTO, E., CERIELLO, A., RYDÉN, L., FERRINI, M., HANSEN, T. B., SCHNELL, O., STANDL, E. & BEULENS, J. W. 2019. Diabetes as a cardiovascular risk factor: An overview of global trends of macro and micro vascular complications. *Eur J Prev Cardiol,* 26**,** 25-32.

DALLALI, H., HECHMI, M., MORJANE, I., ELOUEJ, S., JMEL, H., BEN HALIMA, Y., ABID, A., BAHLOUS, A., BARAKAT, A., JAMOUSSI, H., ABDELHAK, S. & KEFI, R. 2022. Association of HNF1A gene variants and haplotypes with metabolic syndrome: a case-control study in the Tunisian population and a meta-analysis. *Diabetol Metab Syndr,* 14**,** 25.

DAMASCENO, A., MAYOSI, B. M., SANI, M., OGAH, O. S., MONDO, C., OJJI, D., DZUDIE, A., KOUAM, C. K., SULIMAN, A., SCHRUEDER, N., YONGA, G., BA, S. A., MARU, F., ALEMAYEHU, B., EDWARDS, C., DAVISON, B. A., COTTER, G. & SLIWA, K. 2012. The causes, treatment, and outcome of acute heart failure in 1006 Africans from 9 countries. *Arch Intern Med,* 172**,** 1386-94.

DAMASK, A., STEG, P. G., SCHWARTZ, G. G., SZAREK, M., HAGSTRÖM, E., BADIMON, L., CHAPMAN, M. J., BOILEAU, C., TSIMIKAS, S., GINSBERG, H. N., BANERJEE, P., MANVELIAN, G., PORDY, R., HESS, S., OVERTON, J. D., LOTTA, L. A., YANCOPOULOS, G. D., ABECASIS, G. R., BARAS, A. & PAULDING, C. 2020. Patients With High Genome-Wide Polygenic Risk Scores for Coronary Artery Disease May Receive Greater Clinical Benefit From Alirocumab Treatment in the ODYSSEY OUTCOMES Trial. *Circulation,* 141**,** 624-636.

DAWED, A. Y., ZHOU, K. & PEARSON, E. R. 2016. Pharmacogenetics in type 2 diabetes: influence on response to oral hypoglycemic agents. *Pharmgenomics Pers Med,* 9**,** 17-29.

DE BRUIN, T. W., MAILLY, F., VAN BARLINGEN, H. H., FISHER, R., CASTRO CABEZAS, M., TALMUD, P., DALLINGA-THIE, G. M. & HUMPHRIES, S. E. 1996. Lipoprotein lipase gene mutations D9N and N291S in four pedigrees with familial combined hyperlipidaemia. *Eur J Clin Invest,* 26**,** 631-9.

DEFRONZO, R. A., FERRANNINI, E., GROOP, L., HENRY, R. R., HERMAN, W. H., HOLST, J. J., HU, F. B., KAHN, C. R., RAZ, I., SHULMAN, G. I., SIMONSON, D. C., TESTA, M. A. & WEISS, R. 2015. Type 2 diabetes mellitus. *Nat Rev Dis Primers,* 1**,** 15019.

DELEZIE, J., DUMONT, S., DARDENTE, H., OUDART, H., GRÉCHEZ-CASSIAU, A., KLOSEN, P., TEBOUL, M., DELAUNAY, F., PÉVET, P. & CHALLET, E. 2012. The nuclear receptor REV-ERBα is required for the daily balance of carbohydrate and lipid metabolism. *Faseb j,* 26**,** 3321-35.

DENDUP, T., FENG, X., CLINGAN, S. & ASTELL-BURT, T. 2018. Environmental Risk Factors for Developing Type 2 Diabetes Mellitus: A Systematic Review. *Int J Environ Res Public Health,* 15.

DERE, W. H. & SUTO, T. S. 2009. The role of pharmacogenetics and pharmacogenomics in improving translational medicine. *Clin Cases Miner Bone Metab,* 6**,** 13-6.

DINH, A., MIERTSCHIN, S., YOUNG, A. & MOHANTY, S. D. 2019. A data-driven approach to predicting diabetes and cardiovascular disease with machine learning. *BMC Med Inform Decis Mak,* 19**,** 211.

DRON, J. S., WANG, J., CAO, H., MCINTYRE, A. D., IACOCCA, M. A., MENARD, J. R., MOVSESYAN, I., MALLOY, M. J., PULLINGER, C. R., KANE, J. P. & HEGELE, R. A. 2019. Severe hypertriglyceridemia is primarily polygenic. *J Clin Lipidol,* 13**,** 80-88.

EDLITZ, Y. & SEGAL, E. 2022. Prediction of type 2 diabetes mellitus onset using logistic regression-based scorecards. *Elife,* 11.

EINARSON, T. R., ACS, A., LUDWIG, C. & PANTON, U. H. 2018. Economic Burden of Cardiovascular Disease in Type 2 Diabetes: A Systematic Review. *Value Health,* 21**,** 881-890.

ELK, N. & IWUCHUKWU, O. F. 2017. Using Personalized Medicine in the Management of Diabetes Mellitus. *Pharmacotherapy,* 37**,** 1131-1149.

ELLIS, L. 2007. Academics' perceptions of the professional or clinical doctorate: findings of a national survey. *J Clin Nurs,* 16**,** 2272-9.

ESTAMPADOR, A. C. & FRANKS, P. W. 2018. Precision Medicine in Obesity and Type 2 Diabetes: The Relevance of Early-Life Exposures. *Clin Chem,* 64**,** 130-141.

FAHED, G., AOUN, L., BOU ZERDAN, M., ALLAM, S., BOU ZERDAN, M., BOUFERRAA, Y. & ASSI, H. I. 2022. Metabolic Syndrome: Updates on Pathophysiology and Management in 2021. *Int J Mol Sci,* 23.

FEINGOLD, K. R. 2000. Introduction to Lipids and Lipoproteins. *In:* FEINGOLD, K. R., ANAWALT, B., BLACKMAN, M. R., BOYCE, A., CHROUSOS, G., CORPAS, E., DE HERDER, W. W., DHATARIYA, K., DUNGAN, K., HOFLAND, J., KALRA, S., KALTSAS, G., KAPOOR, N., KOCH, C., KOPP, P., KORBONITS, M., KOVACS, C. S., KUOHUNG, W., LAFERRÈRE, B., LEVY, M., MCGEE, E. A., MCLACHLAN, R., NEW, M., PURNELL, J., SAHAY, R., SHAH, A. S., SINGER, F., SPERLING, M. A., STRATAKIS, C. A., TRENCE, D. L. & WILSON, D. P. (eds.) *Endotext.* South Dartmouth (MA): MDText.com, Inc.

FEOLI-FONSECA, J. C., LÉVY, E., GODARD, M. & LAMBERT, M. 1998. Familial lipoprotein lipase deficiency in infancy: clinical, biochemical, and molecular study. *J Pediatr,* 133**,** 417-23.

FERLAND, A., CHÂTEAU-DEGAT, M. L., HERNANDEZ, T. L. & ECKEL, R. H. 2012. Tissue-specific responses of lipoprotein lipase to dietary macronutrient composition as a predictor of weight gain over 4 years. *Obesity (Silver Spring),* 20**,** 1006-11.

FERREIRA, L. D., PULAWA, L. K., JENSEN, D. R. & ECKEL, R. H. 2001. Overexpressing human lipoprotein lipase in mouse skeletal muscle is associated with insulin resistance. *Diabetes,* 50**,** 1064-8.

FLOREZ, J. C. 2008. Newly identified loci highlight beta cell dysfunction as a key cause of type 2 diabetes: where are the insulin resistance genes? *Diabetologia,* 51**,** 1100-10.

FLOREZ, J. C. 2017. Pharmacogenetics in type 2 diabetes: precision medicine or discovery tool? *Diabetologia,* 60**,** 800-807.

FRANCESCHINI, N., CARTY, C., BŮZKOVÁ, P., REINER, A. P., GARRETT, T., LIN, Y., VÖCKLER, J. S., HINDORFF, L. A., COLE, S. A., BOERWINKLE, E., LIN, D. Y., BOOKMAN, E., BEST, L. G., BELLA, J. N., EATON, C., GREENLAND, P., JENNY, N., NORTH, K. E., TAVERNA, D., YOUNG, A. M., DEELMAN, E., KOOPERBERG, C., PSATY, B. & HEISS, G. 2011. Association of genetic variants and incident coronary heart disease in multiethnic cohorts: the PAGE study. *Circ Cardiovasc Genet,* 4**,** 661-72.

FRANEK, E., PAIS, P., BASILE, J., NICOLAY, C., RAHA, S., HICKEY, A., AHMAD, N. N., KONIG, M., KAN, H. & GERSTEIN, H. C. 2023. General versus central adiposity as risk factors for cardiovascular-related outcomes in a high-risk population with type 2 diabetes: a post hoc analysis of the REWIND trial. *Cardiovasc Diabetol,* 22**,** 52.

FREEMAN, A. M., ACEVEDO, L. A. & PENNINGS, N. 2024. Insulin Resistance. *StatPearls.* Treasure Island (FL) ineligible companies. Disclosure: Luis Acevedo declares no relevant financial relationships with ineligible companies. Disclosure: Nicholas Pennings declares no relevant financial relationships with ineligible companies.: StatPearls Publishing

Copyright © 2024, StatPearls Publishing LLC.

FRIGGE, M., HOAGLIN, D. C. & IGLEWICZ, B. 1989. Some Implementations of the Boxplot. *The American Statistician,* 43**,** 50-54.

FROSTEGÅRD, J. 2013. Immunity, atherosclerosis and cardiovascular disease. *BMC Med,* 11**,** 117.

GAGNÉ, C., BRUN, L. D., JULIEN, P., MOORJANI, S. & LUPIEN, P. J. 1989. Primary lipoprotein-lipase-activity deficiency: clinical investigation of a French Canadian population. *Cmaj,* 140**,** 405-11.

GAO, F., BALLANTYNE, C., MA, L., VIRANI, S. S., KEINAN, A. & BRAUTBAR, A. 2014. Rare LPL gene variants attenuate triglyceride reduction and HDL cholesterol increase in response to fenofibric acid therapy in individuals with mixed dyslipidemia. *Atherosclerosis,* 234**,** 249-53.

GARBER, A. J., ABRAHAMSON, M. J., BARZILAY, J. I., BLONDE, L., BLOOMGARDEN, Z. T., BUSH, M. A., DAGOGO-JACK, S., DAVIDSON, M. B., EINHORN, D., GARBER, J. R., GARVEY, W. T., GRUNBERGER, G., HANDELSMAN, Y., HIRSCH, I. B., JELLINGER, P. S., MCGILL, J. B., MECHANICK, J. I., ROSENBLIT, P. D. & UMPIERREZ, G. E. 2015. CONSENSUS STATEMENT BY THE AMERICAN ASSOCIATION OF CLINICAL ENDOCRINOLOGISTS AND AMERICAN COLLEGE OF ENDOCRINOLOGY ON THE COMPREHENSIVE TYPE 2 DIABETES MANAGEMENT

ALGORITHM--2015 EXECUTIVE SUMMARY. *Endocr Pract,* 21**,** 1403-14.

GENEST, J. J., JR., MARTIN-MUNLEY, S. S., MCNAMARA, J. R., ORDOVAS, J. M., JENNER, J., MYERS, R. H., SILBERMAN, S. R., WILSON, P. W., SALEM, D. N. & SCHAEFER, E. J. 1992. Familial lipoprotein disorders in patients with premature coronary artery disease. *Circulation,* 85**,** 2025-33.

GENTILELLA, R., PECHTNER, V., CORCOS, A. & CONSOLI, A. 2019. Glucagon-like peptide-1 receptor agonists in type 2 diabetes treatment: are they all the same? *Diabetes Metab Res Rev,* 35**,** e3070.

GIBSON, G. 2012. Rare and common variants: twenty arguments. *Nat Rev Genet,* 13**,** 135-45.

GLOYN, A. L. & DRUCKER, D. J. 2018. Precision medicine in the management of type 2 diabetes. *Lancet Diabetes Endocrinol,* 6**,** 891-900.

GOISAUF, M., MARTIN, G., BENTZEN, H. B., BUDIN-LJOSNE, I., URSIN, L., DURNOVA, A., LEITSALU, L., SMITH, K., CASATI, S., LAVITRANO, M., MASCALZONI, D., BOECKHOUT, M. & MAYRHOFER, M. T. 2019. Data in question: A survey of European biobank professionals on ethical, legal and societal challenges of biobank research. *PLoS One,* 14**,** e0221496.

GOLLAMUDI, J., SARTAIN, S. E., NAVAEI, A. H., ANEJA, S., KAUR DHAWAN, P., TRAN, D., JOSHI, J., GIDUDU, J., GOLLAMUDI, J., CHIAPPINI, E., VARRICCHIO, F., LAW, B. & MUNOZ, F. M. 2022. Thrombosis and thromboembolism: Brighton collaboration case definition and guidelines for data collection, analysis, and presentation of immunization safety data. *Vaccine,* 40**,** 6431-6444.

GOODARZI, M. O., GUO, X., TAYLOR, K. D., QUIÑONES, M. J., SAAD, M. F., YANG, H., HSUEH, W. A. & ROTTER, J. I. 2004. Lipoprotein lipase is a gene for insulin resistance in Mexican Americans. *Diabetes,* 53**,** 214-20.

GREGG, E. W. 2017. The Changing Tides of the Type 2 Diabetes Epidemic-Smooth Sailing or Troubled Waters Ahead? Kelly West Award Lecture 2016. *Diabetes Care,* 40**,** 1289-1297.

GRIFFIN, M. E., MARCUCCI, M. J., CLINE, G. W., BELL, K., BARUCCI, N., LEE, D., GOODYEAR, L. J., KRAEGEN, E. W., WHITE, M. F. & SHULMAN, G. I. 1999. Free fatty acid-induced insulin resistance is associated with activation of protein kinase C theta and alterations in the insulin signaling cascade. *Diabetes,* 48**,** 1270-4.

GROENEMEIJER, B. E., HALLMAN, M. D., REYMER, P. W., GAGNÉ, E., KUIVENHOVEN, J. A., BRUIN, T., JANSEN, H., LIE, K. I., BRUSCHKE, A. V., BOERWINKLE, E., HAYDEN, M. R. & KASTELEIN, J. J. 1997. Genetic variant showing a positive interaction with beta-blocking agents with a beneficial influence on lipoprotein lipase activity, HDL cholesterol, and triglyceride levels in coronary artery disease patients. The Ser447-stop substitution in the lipoprotein lipase gene. REGRESS Study Group. *Circulation,* 95**,** 2628-35.

GRUNDY, S. M. 2008. Metabolic syndrome pandemic. *Arterioscler Thromb Vasc Biol,* 28**,** 629-36.

HAHN, S. J., KIM, S., CHOI, Y. S., LEE, J. & KANG, J. 2022. Prediction of type 2 diabetes using genome-wide polygenic risk score and metabolic profiles: A machine learning analysis of population-based 10-year prospective cohort study. *EBioMedicine,* 86**,** 104383.

HAJIAN-TILAKI, K. 2013. Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. *Caspian J Intern Med,* 4**,** 627-35.

HAVEL, R. J. & GORDON, R. S., JR. 1960. Idiopathic hyperlipemia: metabolic studies in an affected family. *J Clin Invest,* 39**,** 1777-90.

HE, J., OGDEN, L. G., BAZZANO, L. A., VUPPUTURI, S., LORIA, C. & WHELTON, P. K. 2001. Risk factors for congestive heart failure in US men and women: NHANES I epidemiologic follow-up study. *Arch Intern Med,* 161**,** 996-1002.

HENDERSON, H. E., HASSAN, F., BERGER, G. M. & HAYDEN, M. R. 1992. The lipoprotein lipase Gly188----Glu mutation in South Africans of Indian descent: evidence suggesting common origins and an increased frequency. *J Med Genet,* 29**,** 119-22.

HENDERSON, H. E., MA, Y., HASSAN, M. F., MONSALVE, M. V., MARAIS, A. D., WINKLER, F., GUBERNATOR, K., PETERSON, J., BRUNZELL, J. D. & HAYDEN, M. R. 1991. Amino acid substitution (Ile194----Thr) in exon 5 of the lipoprotein lipase gene causes lipoprotein lipase deficiency in three unrelated probands. Support for a multicentric origin. *J Clin Invest,* 87**,** 2005-11.

HENDERSON, M. K., GOLDRING, K. & SIMEON-DUBACH, D. 2019. Advancing Professionalization of Biobank Business Operations: Performance and Utilization. *Biopreserv Biobank,* 17**,** 213-218.

HIERONYMUS, L. & GRIFFIN, S. 2015. Role of Amylin in Type 1 and Type 2 Diabetes. *Diabetes Educ,* 41**,** 47s-56s.

HIMANSHU, D., ALI, W. & WAMIQUE, M. 2020. Type 2 diabetes mellitus: pathogenesis and genetic diagnosis. *J Diabetes Metab Disord,* 19**,** 1959-1966.

HIVERT, M. F., JABLONSKI, K. A., PERREAULT, L., SAXENA, R., MCATEER, J. B., FRANKS, P. W., HAMMAN, R. F., KAHN, S. E., HAFFNER, S., MEIGS, J. B., ALTSHULER, D., KNOWLER, W. C. & FLOREZ, J. C. 2011. Updated genetic score based on 34 confirmed type 2 diabetes Loci is associated with diabetes incidence and regression to normoglycemia in the diabetes prevention program. *Diabetes,* 60**,** 1340-8.

HO, W. M., WU, Y. Y. & CHEN, Y. C. 2020. Genetic Variants behind Cardiovascular Diseases and Dementia. *Genes (Basel),* 11.

HSU, C. S., CHANG, S. T., NFOR, O. N., LEE, K. J., HO, C. C., LIU, C. C., LEE, S. S. & LIAW, Y. P. 2021. Association of Metabolic Syndrome with Aerobic Exercise and LPL rs3779788 Polymorphism in Taiwan Biobank Individuals. *Diabetes Metab Syndr Obes,* 14**,** 3997-4004.

HUANG, H., CHEN, G., DONG, Y., ZHU, Y. & CHEN, H. 2018. Chromium supplementation for adjuvant treatment of type 2 diabetes mellitus: Results from a pooled analysis. *Mol Nutr Food Res,* 62.

HUANG, P. L. 2009. A comprehensive definition for metabolic syndrome. *Dis Model Mech,* 2**,** 231-7.

HUANG, X., GONG, R., LIN, J., LI, R., XIAO, L., DUAN, W. & FANG, D. 2011. Effects of lipoprotein lipase gene variations, a high-carbohydrate low-fat diet, and gender on serum lipid profiles in healthy Chinese Han youth. *Biosci Trends,* 5**,** 198-204.

HUNT, S. C., PROVINCE, M. A., ATWOOD, L. D., SHOLINSKY, P., LALOUEL, J. M., RAO, D. C., WILLIAMS, R. R. & LEPPERT, M. F. 1999. No linkage of the lipoprotein lipase locus to hypertension in Caucasians. *J Hypertens,* 17**,** 39-43.

IGLAY, K., HANNACHI, H., JOSEPH HOWIE, P., XU, J., LI, X., ENGEL, S. S., MOORE, L. M. & RAJPATHAK, S. 2016. Prevalence and co-prevalence of comorbidities among patients with type 2 diabetes mellitus. *Curr Med Res Opin,* 32**,** 1243-52.

INGELMAN-SUNDBERG, M. & PIRMOHAMED, M. 2024. Precision medicine in cardiovascular therapeutics: Evaluating the role of pharmacogenetic analysis prior to drug treatment. *J Intern Med,* 295**,** 583-598.

JEMAA, R., FUMERON, F., POIRIER, O., LECERF, L., EVANS, A., ARVEILER, D., LUC, G., CAMBOU, J. P., BARD, J. M., FRUCHART, J. C. & ET AL. 1995. Lipoprotein lipase gene polymorphisms: associations with myocardial infarction and lipoprotein levels, the ECTIM study. Etude Cas Témoin sur l'Infarctus du Myocarde. *J Lipid Res,* 36**,** 2141-6.

JOHANSEN, C. T., WANG, J., LANKTREE, M. B., CAO, H., MCINTYRE, A. D., BAN, M. R., MARTINS, R. A., KENNEDY, B. A., HASSELL, R. G., VISSER, M. E., SCHWARTZ, S. M., VOIGHT, B. F., ELOSUA, R., SALOMAA, V., O'DONNELL, C. J., DALLINGA-THIE, G. M., ANAND, S. S., YUSUF, S., HUFF, M. W., KATHIRESAN, S. & HEGELE, R. A. 2010. Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. *Nat Genet,* 42**,** 684-7.

JOSEPH, P., LEONG, D., MCKEE, M., ANAND, S. S., SCHWALM, J. D., TEO, K., MENTE, A. & YUSUF, S. 2017. Reducing the Global Burden of Cardiovascular Disease, Part 1: The Epidemiology and Risk Factors. *Circ Res,* 121**,** 677-694.

JOSHI, R. D. & DHAKAL, C. K. 2021. Predicting Type 2 Diabetes Using Logistic Regression and Machine Learning Approaches. *Int J Environ Res Public Health,* 18.

KAHN, S. E. 2003. The relative contributions of insulin resistance and beta-cell dysfunction to the pathophysiology of Type 2 diabetes. *Diabetologia,* 46**,** 3-19.

KANETO, H. 2015. [Pathophysiology of type 2 diabetes mellitus]. *Nihon Rinsho,* 73**,** 2003-7.

KANG, S. W., KIM, S. K., KIM, Y. S. & PARK, M. S. 2023. Risk prediction of the metabolic syndrome using TyG Index and SNPs: a 10-year longitudinal prospective cohort study. *Mol Cell Biochem,* 478**,** 39-45.

KARKI, R., PANDYA, D., ELSTON, R. C. & FERLINI, C. 2015. Defining "mutation" and "polymorphism" in the era of personal genomics. *BMC Med Genomics,* 8**,** 37.

KASSI, E., PERVANIDOU, P., KALTSAS, G. & CHROUSOS, G. 2011. Metabolic syndrome: definitions and controversies. *BMC Med,* 9**,** 48.

KATHIRESAN, S. & SRIVASTAVA, D. 2012. Genetics of human cardiovascular disease. *Cell,* 148**,** 1242-57.

KAUTZKY-WILLER, A., HARREITER, J. & PACINI, G. 2016. Sex and Gender Differences in Risk, Pathophysiology and Complications of Type 2 Diabetes Mellitus. *Endocr Rev,* 37**,** 278-316.

KAVAZARAKIS, E., STABOULI, S., GOURGIOTIS, D., ROUMELIOTOU, K., TRAEGER-SYNODINOS, J., BOSSIOS, A., FRETZAYAS, A. & KANAVAKIS, E. 2004. Severe hypertriglyceridaemia in a Greek infant: a clinical, biochemical and genetic study. *Eur J Pediatr,* 163**,** 462-6.

KAWAJIRI, T., OSAKI, Y. & KISHIMOTO, T. 2012. Association of gene polymorphism of the fat mass and obesity associated gene with metabolic syndrome: a retrospective cohort study in Japanese workers. *Yonago Acta Med,* 55**,** 29-40.

KAYE, J., BELL, J., BRICENO, L. & MITCHELL, C. 2016. Biobank Report: United Kingdom. *J Law Med Ethics,* 44**,** 96-105.

KELSEY, M. D., NELSON, A. J., GREEN, J. B., GRANGER, C. B., PETERSON, E. D., MCGUIRE, D. K. & PAGIDIPATI, N. J. 2022. Guidelines for Cardiovascular Risk Reduction in Patients With Type 2 Diabetes: JACC Guideline Comparison. *J Am Coll Cardiol,* 79**,** 1849-1857.

KHANAM, P. A., HOQUE, S., BEGUM, T., HABIB, S. H. & LATIF, Z. A. 2017. Microvascular complications and their associated risk factors in type 2 diabetes mellitus. *Diabetes Metab Syndr,* 11 Suppl 2**,** S577-s581.

KHERA, A. V., EMDIN, C. A., DRAKE, I., NATARAJAN, P., BICK, A. G., COOK, N. R., CHASMAN, D. I., BABER, U., MEHRAN, R., RADER, D. J., FUSTER, V., BOERWINKLE, E., MELANDER, O., ORHO-MELANDER, M., RIDKER, P. M. & KATHIRESAN, S. 2016. Genetic Risk, Adherence to a Healthy Lifestyle, and Coronary Disease. *N Engl J Med,* 375**,** 2349-2358.

KING, H., AUBERT, R. E. & HERMAN, W. H. 1998. Global burden of diabetes, 1995-2025: prevalence, numerical estimates, and projections. *Diabetes Care,* 21**,** 1414-31.

KIRKBY, H. M., CALVERT, M., DRAPER, H., KEELEY, T. & WILSON, S. 2012. What potential research participants want to know about research: a systematic review. *BMJ Open,* 2.

KODERA, M., HAYAKAWA, I., KOMURA, K., YANABA, K., HASEGAWA, M., TAKEHARA, K. & SATO, S. 2005. Anti-lipoprotein lipase antibody in systemic sclerosis: association with elevated serum triglyceride concentrations. *J Rheumatol,* 32**,** 629-36.

KOLB, H. & MARTIN, S. 2017. Environmental/lifestyle factors in the pathogenesis and prevention of type 2 diabetes. *BMC Med,* 15**,** 131.

KONG, S. & CHO, Y. S. 2019. Identification of female-specific genetic variants for metabolic syndrome and its component traits to improve the prediction of metabolic syndrome in females. *BMC Med Genet,* 20**,** 99.

KRAJA, A. T., VAIDYA, D., PANKOW, J. S., GOODARZI, M. O., ASSIMES, T. L., KULLO, I. J., SOVIO, U., MATHIAS, R. A., SUN, Y. V., FRANCESCHINI, N., ABSHER, D., LI, G., ZHANG, Q., FEITOSA, M. F., GLAZER, N. L., HARITUNIANS, T., HARTIKAINEN, A. L., KNOWLES, J. W., NORTH, K. E., IRIBARREN, C., KRAL, B., YANEK, L., O'REILLY, P. F., MCCARTHY, M. I., JAQUISH, C., COUPER, D. J., CHAKRAVARTI, A., PSATY, B. M., BECKER, L. C., PROVINCE, M.

A., BOERWINKLE, E., QUERTERMOUS, T., PALOTIE, L., JARVELIN, M. R., BECKER, D. M., KARDIA, S. L., ROTTER, J. I., CHEN, Y. D. & BORECKI, I. B. 2011. A bivariate genome-wide approach to metabolic syndrome: STAMPEED consortium. *Diabetes,* 60**,** 1329-39.

KRISTIANSSON, K., PEROLA, M., TIKKANEN, E., KETTUNEN, J., SURAKKA, I., HAVULINNA, A. S., STANCÁKOVÁ, A., BARNES, C., WIDEN, E., KAJANTIE, E., ERIKSSON, J. G., VIIKARI, J., KÄHÖNEN, M., LEHTIMÄKI, T., RAITAKARI, O. T., HARTIKAINEN, A. L., RUOKONEN, A., POUTA, A., JULA, A., KANGAS, A. J., SOININEN, P., ALA-KORPELA, M., MÄNNISTÖ, S., JOUSILAHTI, P., BONNYCASTLE, L. L., JÄRVELIN, M. R., KUUSISTO, J., COLLINS, F. S., LAAKSO, M., HURLES, M. E., PALOTIE, A., PELTONEN, L., RIPATTI, S. & SALOMAA, V. 2012. Genome-wide screen for metabolic syndrome susceptibility Loci reveals strong lipid gene contribution but no evidence for common genetic basis for clustering of metabolic syndrome traits. *Circ Cardiovasc Genet,* 5**,** 242-9.

KUMARI, A., KRISTENSEN, K. K., PLOUG, M. & WINTHER, A. L. 2021. The Importance of Lipoprotein Lipase Regulation in Atherosclerosis. *Biomedicines,* 9.

LAMARCHE, B., TCHERNOF, A., MOORJANI, S., CANTIN, B., DAGENAIS, G. R., LUPIEN, P. J. & DESPRÉS, J. P. 1997. Small, dense low-density lipoprotein particles as a predictor of the risk of ischemic heart disease in men. Prospective results from the Québec Cardiovascular Study. *Circulation,* 95**,** 69-75.

LANDER, E. S., LINTON, L. M., BIRREN, B., NUSBAUM, C., ZODY, M. C., BALDWIN, J., DEVON, K., DEWAR, K., DOYLE, M., FITZHUGH, W., FUNKE, R., GAGE, D., HARRIS, K., HEAFORD, A., HOWLAND, J., KANN, L., LEHOCZKY, J., LEVINE, R., MCEWAN, P., MCKERNAN, K., MELDRIM, J., MESIROV, J. P., MIRANDA, C., MORRIS, W., NAYLOR, J., RAYMOND, C., ROSETTI, M., SANTOS, R., SHERIDAN, A., SOUGNEZ, C., STANGE-THOMANN, Y., STOJANOVIC, N., SUBRAMANIAN, A., WYMAN, D., ROGERS, J., SULSTON, J., AINSCOUGH, R., BECK, S., BENTLEY, D., BURTON, J., CLEE, C., CARTER, N., COULSON, A., DEADMAN, R., DELOUKAS, P., DUNHAM, A., DUNHAM, I., DURBIN, R., FRENCH, L., GRAFHAM, D., GREGORY, S., HUBBARD, T., HUMPHRAY, S., HUNT, A., JONES, M., LLOYD, C., MCMURRAY, A., MATTHEWS, L., MERCER, S., MILNE, S., MULLIKIN, J. C., MUNGALL, A., PLUMB, R., ROSS, M., SHOWNKEEN, R., SIMS, S., WATERSTON, R. H., WILSON, R. K., HILLIER, L. W., MCPHERSON, J. D., MARRA, M. A., MARDIS, E. R., FULTON, L. A., CHINWALLA, A. T., PEPIN, K. H., GISH, W. R., CHISSOE, S. L., WENDL, M. C., DELEHAUNTY, K. D., MINER, T. L., DELEHAUNTY, A., KRAMER, J. B., COOK, L. L., FULTON, R. S., JOHNSON, D. L., MINX, P. J., CLIFTON, S. W., HAWKINS, T., BRANSCOMB, E., PREDKI, P., RICHARDSON, P., WENNING, S., SLEZAK, T., DOGGETT, N., CHENG, J. F., OLSEN, A., LUCAS, S., ELKIN, C., UBERBACHER, E., FRAZIER, M., et al. 2001. Initial sequencing and analysis of the human genome. *Nature,* 409**,** 860-921.

LANDRUM, M. J., LEE, J. M., BENSON, M., BROWN, G. R., CHAO, C., CHITIPIRALLA, S., GU, B., HART, J., HOFFMAN, D., JANG, W.,

KARAPETYAN, K., KATZ, K., LIU, C., MADDIPATLA, Z., MALHEIRO, A., MCDANIEL, K., OVETSKY, M., RILEY, G., ZHOU, G., HOLMES, J. B., KATTMAN, B. L. & MAGLOTT, D. R. 2018. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res,* 46**,** D1062-d1067.

LANGHOF, H., KAHRASS, H., ILLIG, T., JAHNS, R. & STRECH, D. 2018. Current practices for access, compensation, and prioritization in biobanks. Results from an interview study. *Eur J Hum Genet,* 26**,** 1572-1581.

LARSSON, S. C. & BURGESS, S. 2021. Causal role of high body mass index in multiple chronic diseases: a systematic review and meta-analysis of Mendelian randomization studies. *BMC Med,* 19**,** 320.

LAWLOR, N., KHETAN, S., UCAR, D. & STITZEL, M. L. 2017. Genomics of Islet (Dys)function and Type 2 Diabetes. *Trends Genet,* 33**,** 244-255.

LAZAREVA, T. E., BARBITOFF, Y. A., NASYKHOVA, Y. A. & GLOTOV, A. S. 2024. Major Causes of Conflicting Interpretations of Variant Pathogenicity in Rare Disease: A Systematic Analysis. *J Pers Med,* 14.

LEAN, M. E., LESLIE, W. S., BARNES, A. C., BROSNAHAN, N., THOM, G., MCCOMBIE, L., PETERS, C., ZHYZHNEUSKAYA, S., AL-MRABEH, A., HOLLINGSWORTH, K. G., RODRIGUES, A. M., REHACKOVA, L., ADAMSON, A. J., SNIEHOTTA, F. F., MATHERS, J. C., ROSS, H. M., MCILVENNA, Y., STEFANETTI, R., TRENELL, M., WELSH, P., KEAN, S., FORD, I., MCCONNACHIE, A., SATTAR, N. & TAYLOR, R. 2018. Primary care-led weight management for remission of type 2 diabetes (DiRECT): an open-label, cluster-randomised trial. *Lancet,* 391**,** 541-551.

LEE, H. S., KIM, Y. & PARK, T. 2018. New Common and Rare Variants Influencing Metabolic Syndrome and Its Individual Components in a Korean Population. *Sci Rep,* 8**,** 5701.

LEE, S. H., HWANG, S. M., KANG, D. H. & YANG, H. J. 2019. Brain education-based meditation for patients with hypertension and/or type 2 diabetes: A pilot randomized controlled trial. *Medicine (Baltimore),* 98**,** e15574.

LEMIEUX, I. & DESPRÉS, J. P. 2020. Metabolic Syndrome: Past, Present and Future. *Nutrients,* 12.

LENT-SCHOCHET, D. & JIALAL, I. 2024. Biochemistry, Lipoprotein Metabolism. *StatPearls.* Treasure Island (FL) with ineligible companies. Disclosure: Ishwarlal Jialal declares no relevant financial relationships with ineligible companies.: StatPearls Publishing

LEUNG, A., SCHONES, D. E. & NATARAJAN, R. 2012. Using epigenetic mechanisms to understand the impact of common disease causing alleles. *Curr Opin Immunol,* 24**,** 558-63.

LI, J. H., SZCZERBINSKI, L., DAWED, A. Y., KAUR, V., TODD, J. N., PEARSON, E. R. & FLOREZ, J. C. 2021. A Polygenic Score for Type 2 Diabetes Risk Is Associated With Both the Acute and Sustained Response to Sulfonylureas. *Diabetes,* 70**,** 293-300.

LI, M. M., DATTO, M., DUNCAVAGE, E. J., KULKARNI, S., LINDEMAN, N. I., ROY, S., TSIMBERIDOU, A. M., VNENCAK-JONES, C. L., WOLFF, D. J., YOUNES, A. & NIKIFOROVA, M. N. 2017. Standards and

Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. *J Mol Diagn,* 19**,** 4-23.

LIM, S., SHIN, H., SONG, J. H., KWAK, S. H., KANG, S. M., WON YOON, J., CHOI, S. H., CHO, S. I., PARK, K. S., LEE, H. K., JANG, H. C. & KOH, K. K. 2011. Increasing prevalence of metabolic syndrome in Korea: the Korean National Health and Nutrition Examination Survey for 1998-2007. *Diabetes Care,* 34**,** 1323-8.

LIN, E., KUO, P. H., LIU, Y. L., YANG, A. C. & TSAI, S. J. 2017. Detection of susceptibility loci on APOA5 and COLEC12 associated with metabolic syndrome using a genome-wide association study in a Taiwanese population. *Oncotarget,* 8**,** 93349-93359.

LIND, L. 2019. Genome-Wide Association Study of the Metabolic Syndrome in UK Biobank. *Metab Syndr Relat Disord,* 17**,** 505-511.

LIU, A., LEE, L., ZHAN, S., CAO, W., LV, J., GUO, X. & HU, Y. 2004. The S447X polymorphism of the lipoprotein lipase gene is associated with lipoprotein lipid and blood pressure levels in Chinese patients with essential hypertension. *J Hypertens,* 22**,** 1503-9.

LIU, N., SANG, Y., CHEN, S. & LIU, X. 2021. Associations of the LPL S447X and Hind III Polymorphism with Type 2 Diabetes Mellitus Risk: A Meta-Analysis. *Horm Metab Res,* 53**,** 49-55.

LOOS, R. J. F. 2020. 15 years of genome-wide association studies and no signs of slowing down. *Nat Commun,* 11**,** 5900.

MAILLY, F., PALMEN, J., MULLER, D. P., GIBBS, T., LLOYD, J., BRUNZELL, J., DURRINGTON, P., MITROPOULOS, K., BETTERIDGE, J., WATTS, G., LITHELL, H., ANGELICO, F., HUMPHRIES, S. E. & TALMUD, P. J. 1997. Familial lipoprotein lipase (LPL) deficiency: a catalogue of LPL gene mutations identified in 20 patients from the UK, Sweden, and Italy. *Hum Mutat,* 10**,** 465-73.

MALANDRINO, N. & SMITH, R. J. 2011. Personalized medicine in diabetes. *Clin Chem,* 57**,** 231-40.

MALECKAS, A., VENCLAUSKAS, L., WALLENIUS, V., LONROTH, H. & FANDRIKS, L. 2015. Surgery in the treatment of type 2 diabetes mellitus. *Scand J Surg,* 104**,** 40-7.

MAMPUTU, J. C., LEVESQUE, L. & RENIER, G. 2000. Proliferative effect of lipoprotein lipase on human vascular smooth muscle cells. *Arterioscler Thromb Vasc Biol,* 20**,** 2212-9.

MANNINO, G. C., ANDREOZZI, F. & SESTI, G. 2019. Pharmacogenetics of type 2 diabetes mellitus, the route toward tailored medicine. *Diabetes Metab Res Rev,* 35**,** e3109.

MARS, N., KOSKELA, J. T., RIPATTI, P., KIISKINEN, T. T. J., HAVULINNA, A. S., LINDBOHM, J. V., AHOLA-OLLI, A., KURKI, M., KARJALAINEN, J., PALTA, P., NEALE, B. M., DALY, M., SALOMAA, V., PALOTIE, A., WIDÉN, E. & RIPATTI, S. 2020. Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers. *Nat Med,* 26**,** 549-557.

MATFIN, G. 2010. The metabolic syndrome: what's in a name? *Ther Adv Endocrinol Metab,* 1**,** 39-45.

MCGILL, R., TUKEY, J. W. & LARSEN, W. A. 1978. Variations of Box Plots. *The American Statistician,* 32**,** 12-16.
MCGREGOR, K. A. & OTT, M. A. 2019. Banking the Future: Adolescent Capacity to Consent to Biobank Research. *Ethics Hum Res,* 41**,** 15-22.
MEAD, J. R., IRVINE, S. A. & RAMJI, D. P. 2002. Lipoprotein lipase: structure, function, regulation, and role in disease. *J Mol Med (Berl),* 80**,** 753-69.
MEAD, J. R. & RAMJI, D. P. 2002. The pivotal role of lipoprotein lipase in atherosclerosis. *Cardiovasc Res,* 55**,** 261-9.
MEIGS, J. B., SHRADER, P., SULLIVAN, L. M., MCATEER, J. B., FOX, C. S., DUPUIS, J., MANNING, A. K., FLOREZ, J. C., WILSON, P. W., D'AGOSTINO, R. B., SR. & CUPPLES, L. A. 2008. Genotype score in addition to common risk factors for prediction of type 2 diabetes. *N Engl J Med,* 359**,** 2208-19.
MEIGS, J. B., WILSON, P. W., FOX, C. S., VASAN, R. S., NATHAN, D. M., SULLIVAN, L. M. & D'AGOSTINO, R. B. 2006. Body mass index, metabolic syndrome, and risk of type 2 diabetes or cardiovascular disease. *J Clin Endocrinol Metab,* 91**,** 2906-12.
MERKEL, M., HEEREN, J., DUDECK, W., RINNINGER, F., RADNER, H., BRESLOW, J. L., GOLDBERG, I. J., ZECHNER, R. & GRETEN, H. 2002. Inactive lipoprotein lipase (LPL) alone increases selective cholesterol ester uptake in vivo, whereas in the presence of active LPL it also increases triglyceride hydrolysis and whole particle lipoprotein uptake. *J Biol Chem,* 277**,** 7405-11.
MERONE, L. & MCDERMOTT, R. 2017. Nutritional anti-inflammatories in the treatment and prevention of type 2 diabetes mellitus and the metabolic syndrome. *Diabetes Res Clin Pract,* 127**,** 238-253.
METZ, C. E. 1978. Basic principles of ROC analysis. *Semin Nucl Med,* 8**,** 283-98.
MIKKELSEN, R. B., GJERRIS, M., WALDEMAR, G. & SANDOE, P. 2019. Broad consent for biobanks is best - provided it is also deep. *BMC Med Ethics,* 20**,** 71.
MOLINA-LUQUE, R., ULLOA, N., ROMERO-SALDAÑA, M., ZILIC, M., GLEISNER, A., LANUZA, F. & MOLINA-RECIO, G. 2021. Association between the FTO SNP rs9939609 and Metabolic Syndrome in Chilean Children. *Nutrients,* 13.
MONDA, K. L., NORTH, K. E., HUNT, S. C., RAO, D. C., PROVINCE, M. A. & KRAJA, A. T. 2010. The genetics of obesity and the metabolic syndrome. *Endocr Metab Immune Disord Drug Targets,* 10**,** 86-108.
MONSALVE, M. V., HENDERSON, H., ROEDERER, G., JULIEN, P., DEEB, S., KASTELEIN, J. J., PERITZ, L., DEVLIN, R., BRUIN, T., MURTHY, M. R. & ET AL. 1990. A missense mutation at codon 188 of the human lipoprotein lipase gene is a frequent cause of lipoprotein lipase deficiency in persons of different ancestries. *J Clin Invest,* 86**,** 728-34.
MOON, S., LEE, Y., WON, S. & LEE, J. 2018. Multiple genotype-phenotype association study reveals intronic variant pair on SIDT2 associated with metabolic syndrome in a Korean population. *Hum Genomics,* 12**,** 48.
MORABIA, A., CAYANIS, E., COSTANZA, M. C., ROSS, B. M., FLAHERTY, M. S., ALVIN, G. B., DAS, K. & GILLIAM, T. C. 2003. Association of extreme blood lipid profile phenotypic variation with 11 reverse

cholesterol transport genes and 10 non-genetic cardiovascular disease risk factors. *Hum Mol Genet,* 12**,** 2733-43.

MORRISON, M., BELL, J., GEORGE, C., HARMON, S., MUNSIE, M. & KAYE, J. 2017. The European General Data Protection Regulation: challenges and considerations for iPSC researchers and biobanks. *Regen Med,* 12**,** 693-703.

MUNSHI, A. 2012. Genetic variation in MDR1, LPL and eNOS genes and the response to atorvastatin treatment in ischemic stroke. *Hum Genet,* 131**,** 1775-81.

MURAKAMI, T., MICHELAGNOLI, S., LONGHI, R., GIANFRANCESCHI, G., PAZZUCCONI, F., CALABRESI, L., SIRTORI, C. R. & FRANCESCHINI, G. 1995. Triglycerides are major determinants of cholesterol esterification/transfer and HDL remodeling in human plasma. *Arterioscler Thromb Vasc Biol,* 15**,** 1819-28.

MUS, R., SADEWA, A. H., HASTUTI, P., PUSPASARI, A., MAHARANI, C. & SETYAWATI, I. 2019. Age Stratification in Genetic Variation of Lipoprotein Lipase in Metabolic Syndrome Javanese Ethnics of Indonesia. *Open Access Maced J Med Sci,* 7**,** 3540-3545.

NAGRANI, R., FORAITA, R., GIANFAGNA, F., IACOVIELLO, L., MARILD, S., MICHELS, N., MOLNÁR, D., MORENO, L., RUSSO, P., VEIDEBAUM, T., AHRENS, W. & MARRON, M. 2020. Common genetic variation in obesity, lipid transfer genes and risk of Metabolic Syndrome: Results from IDEFICS/I.Family study and meta-analysis. *Sci Rep,* 10**,** 7189.

NEJATI, M., ATLASI, M. A., KARIMIAN, M., NIKZAD, H. & AZAMI TAMEH, A. 2018. Lipoprotein lipase gene polymorphisms as risk factors for stroke: a computational and meta-analysis. *Iran J Basic Med Sci,* 21**,** 701-708.

NGUYEN, N. T. & VARELA, J. E. 2017. Bariatric surgery for obesity and metabolic disorders: state of the art. *Nat Rev Gastroenterol Hepatol,* 14**,** 160-169.

NGUYEN, T. T. B., JIN, Y. Y., CHUNG, H. J. & HONG, S. T. 2017. Pharmabiotics as an Emerging Medication for Metabolic Syndrome and Its Related Diseases. *Molecules,* 22.

NOUBIAP, J. J., NANSSEU, J. R., LONTCHI-YIMAGOU, E., NKECK, J. R., NYAGA, U. F., NGOUO, A. T., TOUNOUGA, D. N., TIANYI, F. L., FOKA, A. J., NDOADOUMGUE, A. L. & BIGNA, J. J. 2022. Geographic distribution of metabolic syndrome and its components in the general adult population: A meta-analysis of global data from 28 million individuals. *Diabetes Res Clin Pract,* 188**,** 109924.

NUSSBAUM, R. L., YANG, S. & LINCOLN, S. E. 2017. Clinical Genetics Testing Laboratories Have a Remarkably Low Rate of Clinically Significant Discordance When Interpreting Variants in Hereditary Cancer Syndrome Genes. *J Clin Oncol,* 35**,** 1259-1261.

O'NEILL, S. & O'DRISCOLL, L. 2015. Metabolic syndrome: a closer look at the growing epidemic and its associated pathologies. *Obes Rev,* 16**,** 1-12.

O'SULLIVAN, J. W., RAGHAVAN, S., MARQUEZ-LUNA, C., LUZUM, J. A., DAMRAUER, S. M., ASHLEY, E. A., O'DONNELL, C. J., WILLER, C. J. & NATARAJAN, P. 2022. Polygenic Risk Scores for Cardiovascular Disease: A Scientific Statement From the American Heart Association. *Circulation,* 146**,** e93-e118.

OH, S. W., LEE, J. E., SHIN, E., KWON, H., CHOE, E. K., CHOI, S. Y., RHEE, H. & CHOI, S. H. 2020. Genome-wide association study of metabolic syndrome in Korean populations. *PLoS One,* 15**,** e0227357.

OLIVECRONA, G. 2016. Role of lipoprotein lipase in lipid metabolism. *Curr Opin Lipidol,* 27**,** 233-41.

OSMAN, W., HASSOUN, A., JELINEK, H. F., ALMAHMEED, W., AFANDI, B., TAY, G. K. & ALSAFAR, H. 2020. Genetics of type 2 diabetes and coronary artery disease and their associations with twelve cardiometabolic traits in the United Arab Emirates population. *Gene,* 750**,** 144722.

PADMANABHAN, S., DU TOIT, C. & DOMINICZAK, A. F. 2023. Cardiovascular precision medicine - A pharmacogenomic perspective. *Camb Prism Precis Med,* 1**,** e28.

PATRON, J., SERRA-CAYUELA, A., HAN, B., LI, C. & WISHART, D. S. 2019. Assessing the performance of genome-wide association studies for predicting disease risk. *PLoS One,* 14**,** e0220215.

PAULWEBER, B., WIEBUSCH, H., MIESENBOECK, G., FUNKE, H., ASSMANN, G., HOELZL, B., SIPPL, M. J., FRIEDL, W., PATSCH, J. R. & SANDHOFER, F. 1991. Molecular basis of lipoprotein lipase deficiency in two Austrian families with type I hyperlipoproteinemia. *Atherosclerosis,* 86**,** 239-50.

PERERA, S. D., WANG, J., MCINTYRE, A. D. & HEGELE, R. A. 2025. Lipoprotein Lipase: Structure, Function, and Genetic Variation. *Genes (Basel),* 16.

PERKINS, B. A., CASKEY, C. T., BRAR, P., DEC, E., KAROW, D. S., KAHN, A. M., HOU, Y. C., SHAH, N., BOELDT, D., COUGHLIN, E., HANDS, G., LAVRENKO, V., YU, J., PROCKO, A., APPIS, J., DALE, A. M., GUO, L., JÖNSSON, T. J., WITTMANN, B. M., BARTHA, I., RAMAKRISHNAN, S., BERNAL, A., BREWER, J. B., BREWERTON, S., BIGGS, W. H., TURPAZ, Y. & VENTER, J. C. 2018. Precision medicine screening using whole-genome sequencing and advanced imaging to identify disease risk in adults. *Proc Natl Acad Sci U S A,* 115**,** 3686-3691.

PIRAHANCHI, Y., ANORUO, M. & SHARMA, S. 2023. Biochemistry, Lipoprotein Lipase. *StatPearls.* Treasure Island (FL) ineligible companies. Disclosure: McDamian Anoruo declares no relevant financial relationships with ineligible companies. Disclosure: Sandeep Sharma declares no relevant financial relationships with ineligible companies.: StatPearls Publishing

Copyright © 2023, StatPearls Publishing LLC.

POVEL, C. M., BOER, J. M., REILING, E. & FESKENS, E. J. 2011. Genetic variants and the metabolic syndrome: a systematic review. *Obes Rev,* 12**,** 952-67.

PULAWA, L. K. & ECKEL, R. H. 2002. Overexpression of muscle lipoprotein lipase and insulin sensitivity. *Curr Opin Clin Nutr Metab Care,* 5**,** 569-74.

QIAO, T., LUO, T., PEI, H., YIMINGNIYAZI, B., AILI, D., AIMUDULA, A., ZHAO, H., ZHANG, H., DAI, J. & WANG, D. 2022. Association between abdominal obesity indices and risk of cardiovascular events in

Chinese populations with type 2 diabetes: a prospective cohort study. *Cardiovasc Diabetol,* 21**,** 225.
QIN, D. 2019. Next-generation sequencing and its clinical application. *Cancer Biol Med,* 16**,** 4-10.
QUEITSCH, C., CARLSON, K. D. & GIRIRAJAN, S. 2012. Lessons from model organisms: phenotypic robustness and missing heritability in complex disease. *PLoS Genet,* 8**,** e1003041.
RAMZAN, S., TIMMINS, P., HASAN, S. S. & BABAR, Z. U. 2019. Cost analysis of type 2 diabetes mellitus treatment in economically developed countries. *Expert Rev Pharmacoecon Outcomes Res,* 19**,** 5-14.
REAVEN, G. M. 1988. Banting lecture 1988. Role of insulin resistance in human disease. *Diabetes,* 37**,** 1595-607.
REAVEN, G. M. 2001. Syndrome x: a short history. *Ochsner J,* 3**,** 124-5.
REGITZ-ZAGROSEK, V. & GEBHARD, C. 2023. Gender medicine: effects of sex and gender on cardiovascular disease manifestation and outcomes. *Nat Rev Cardiol,* 20**,** 236-247.
REYMER, P. W., GAGNÉ, E., GROENEMEYER, B. E., ZHANG, H., FORSYTH, I., JANSEN, H., SEIDELL, J. C., KROMHOUT, D., LIE, K. E., KASTELEIN, J. & ET AL. 1995a. A lipoprotein lipase mutation (Asn291Ser) is associated with reduced HDL cholesterol levels in premature atherosclerosis. *Nat Genet,* 10**,** 28-34.
REYMER, P. W., GROENEMEYER, B. E., GAGNÉ, E., MIAO, L., APPELMAN, E. E., SEIDEL, J. C., KROMHOUT, D., BIJVOET, S. M., VAN DE OEVER, K., BRUIN, T. & ET AL. 1995b. A frequently occurring mutation in the lipoprotein lipase gene (Asn291Ser) contributes to the expression of familial combined hyperlipidemia. *Hum Mol Genet,* 4**,** 1543-9.
RHEEDER, R. 2017. Biobanks in South Africa: A global perspective on privacy and confidentiality. *S Afr Med J,* 107**,** 390-393.
RIP, J., NIERMAN, M. C., ROSS, C. J., JUKEMA, J. W., HAYDEN, M. R., KASTELEIN, J. J., STROES, E. S. & KUIVENHOVEN, J. A. 2006. Lipoprotein lipase S447X: a naturally occurring gain-of-function mutation. *Arterioscler Thromb Vasc Biol,* 26**,** 1236-45.
ROBERTS, C. K., HEVENER, A. L. & BARNARD, R. J. 2013. Metabolic syndrome and insulin resistance: underlying causes and modification by exercise training. *Compr Physiol,* 3**,** 1-58.
RODBARD, H. W. 2018. The Clinical Impact of GLP-1 Receptor Agonists in Type 2 Diabetes: Focus on the Long-Acting Analogs. *Diabetes Technol Ther,* 20**,** S233-s241.
ROSENFELD, P., GLASSMAN, K., VETTER, M. & SMITH, B. 2022. A comparative study of PhD and DNP nurses in an integrated health care system. *Nurs Outlook,* 70**,** 145-153.
ROSOLOVÁ, H. 2017. [Metabolic syndrome and the role of the polypill in the prevention of its complications]. *Vnitr Lek,* 63**,** 712-716.
ROSS, C. J., TWISK, J., MEULENBERG, J. M., LIU, G., VAN DEN OEVER, K., MORAAL, E., HERMENS, W. T., RIP, J., KASTELEIN, J. J., KUIVENHOVEN, J. A. & HAYDEN, M. R. 2004. Long-term correction of murine lipoprotein lipase deficiency with AAV1-mediated gene transfer

of the naturally occurring LPL(S447X) beneficial mutation. *Hum Gene Ther,* 15**,** 906-19.

ROTH, G. A., JOHNSON, C., ABAJOBIR, A., ABD-ALLAH, F., ABERA, S. F., ABYU, G., AHMED, M., AKSUT, B., ALAM, T., ALAM, K., ALLA, F., ALVIS-GUZMAN, N., AMROCK, S., ANSARI, H., ÄRNLÖV, J., ASAYESH, H., ATEY, T. M., AVILA-BURGOS, L., AWASTHI, A., BANERJEE, A., BARAC, A., BÄRNIGHAUSEN, T., BARREGARD, L., BEDI, N., BELAY KETEMA, E., BENNETT, D., BERHE, G., BHUTTA, Z., BITEW, S., CARAPETIS, J., CARRERO, J. J., MALTA, D. C., CASTAÑEDA-ORJUELA, C. A., CASTILLO-RIVAS, J., CATALÁ-LÓPEZ, F., CHOI, J. Y., CHRISTENSEN, H., CIRILLO, M., COOPER, L., JR., CRIQUI, M., CUNDIFF, D., DAMASCENO, A., DANDONA, L., DANDONA, R., DAVLETOV, K., DHARMARATNE, S., DORAIRAJ, P., DUBEY, M., EHRENKRANZ, R., EL SAYED ZAKI, M., FARAON, E. J. A., ESTEGHAMATI, A., FARID, T., FARVID, M., FEIGIN, V., DING, E. L., FOWKES, G., GEBREHIWOT, T., GILLUM, R., GOLD, A., GONA, P., GUPTA, R., HABTEWOLD, T. D., HAFEZI-NEJAD, N., HAILU, T., HAILU, G. B., HANKEY, G., HASSEN, H. Y., ABATE, K. H., HAVMOELLER, R., HAY, S. I., HORINO, M., HOTEZ, P. J., JACOBSEN, K., JAMES, S., JAVANBAKHT, M., JEEMON, P., JOHN, D., JONAS, J., KALKONDE, Y., KARIMKHANI, C., KASAEIAN, A., KHADER, Y., KHAN, A., KHANG, Y. H., KHERA, S., KHOJA, A. T., KHUBCHANDANI, J., KIM, D., KOLTE, D., KOSEN, S., KROHN, K. J., KUMAR, G. A., KWAN, G. F., LAL, D. K., LARSSON, A., LINN, S., LOPEZ, A., LOTUFO, P. A., EL RAZEK, H. M. A., et al. 2017. Global, Regional, and National Burden of Cardiovascular Diseases for 10 Causes, 1990 to 2015. *J Am Coll Cardiol,* 70**,** 1-25.

SAGOO, G. S., TATT, I., SALANTI, G., BUTTERWORTH, A. S., SARWAR, N., VAN MAARLE, M., JUKEMA, J. W., WIMAN, B., KASTELEIN, J. J., BENNET, A. M., DE FAIRE, U., DANESH, J. & HIGGINS, J. P. 2008. Seven lipoprotein lipase gene polymorphisms, lipid fractions, and coronary disease: a HuGE association review and meta-analysis. *Am J Epidemiol,* 168**,** 1233-46.

SAID, M. A., VERWEIJ, N. & VAN DER HARST, P. 2018. Associations of Combined Genetic and Lifestyle Risks With Incident Cardiovascular Disease and Diabetes in the UK Biobank Study. *JAMA Cardiol,* 3**,** 693-702.

SAKLAYEN, M. G. 2018. The Global Epidemic of the Metabolic Syndrome. *Curr Hypertens Rep,* 20**,** 12.

SALEH, A., AL-ABCHA, A. & PEREIRA, N. 2023. Pharmacogenetics of cardiovascular drugs. *Curr Opin Cardiol,* 38**,** 207-214.

SALTIEL, A. R. 2021. Insulin signaling in health and disease. *J Clin Invest,* 131.

SAMPATH KUMAR, A., MAIYA, A. G., SHASTRY, B. A., VAISHALI, K., RAVISHANKAR, N., HAZARI, A., GUNDMI, S. & JADHAV, R. 2019. Exercise and insulin resistance in type 2 diabetes mellitus: A systematic review and meta-analysis. *Ann Phys Rehabil Med,* 62**,** 98-103.

SAPRA, A. & BHANDARI, P. 2024. Diabetes. *StatPearls.* Treasure Island (FL) ineligible companies. Disclosure: Priyanka Bhandari declares no

relevant financial relationships with ineligible companies.: StatPearls Publishing

SCHEEN, A. J. 2016. Precision medicine: The future in diabetes care? *Diabetes Res Clin Pract,* 117**,** 12-21.

SCHWARZE, K., BUCHANAN, J., TAYLOR, J. C. & WORDSWORTH, S. 2018. Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature. *Genet Med,* 20**,** 1122-1130.

SEIP, R. L. & SEMENKOVICH, C. F. 1998. Skeletal muscle lipoprotein lipase: molecular regulation and physiological effects in relation to exercise. *Exerc Sport Sci Rev,* 26**,** 191-218.

SHAH, H. S., GAO, H., MORIERI, M. L., SKUPIEN, J., MARVEL, S., PARÉ, G., MANNINO, G. C., BURANASUPKAJORN, P., MENDONCA, C., HASTINGS, T., MARCOVINA, S. M., SIGAL, R. J., GERSTEIN, H. C., WAGNER, M. J., MOTSINGER-REIF, A. A., BUSE, J. B., KRAFT, P., MYCHALECKYJ, J. C. & DORIA, A. 2016. Genetic Predictors of Cardiovascular Mortality During Intensive Glycemic Control in Type 2 Diabetes: Findings From the ACCORD Clinical Trial. *Diabetes Care,* 39**,** 1915-1924.

SHERRY, S. T., WARD, M. & SIROTKIN, K. 1999. dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation. *Genome Res,* 9**,** 677-9.

SHIM, U., KIM, H. N., SUNG, Y. A. & KIM, H. L. 2014. Pathway Analysis of Metabolic Syndrome Using a Genome-Wide Association Study of Korea Associated Resource (KARE) Cohorts. *Genomics Inform,* 12**,** 195-202.

SIEST, G. 2014. Systems medicine, stratified medicine, personalized medicine but not precision medicine. *Drug Metabol Drug Interact,* 29**,** 1-2.

SMITH, C. E., TUCKER, K. L., LAI, C. Q., PARNELL, L. D., LEE, Y. C. & ORDOVÁS, J. M. 2010. Apolipoprotein A5 and lipoprotein lipase interact to modulate anthropometric measures in Hispanics of Caribbean origin. *Obesity (Silver Spring),* 18**,** 327-32.

SOPPERT, J., LEHRKE, M., MARX, N., JANKOWSKI, J. & NOELS, H. 2020. Lipoproteins and lipids in cardiovascular disease: from mechanistic insights to therapeutic targeting. *Adv Drug Deliv Rev,* 159**,** 4-33.

SPENCER-JONES, N. J., WANG, X., SNIEDER, H., SPECTOR, T. D., CARTER, N. D. & O'DELL, S. D. 2005. Protein tyrosine phosphatase-1B gene PTPN1: selection of tagging single nucleotide polymorphisms and association with body fat, insulin sensitivity, and the metabolic syndrome in a normal female population. *Diabetes,* 54**,** 3296-304.

SRINIVASAN, S., YEE, S. W. & GIACOMINI, K. M. 2018. Pharmacogenetics of Antidiabetic Drugs. *Adv Pharmacol,* 83**,** 361-389.

STUMVOLL, M., GOLDSTEIN, B. J. & VAN HAEFTEN, T. W. 2005. Type 2 diabetes: principles of pathogenesis and therapy. *Lancet,* 365**,** 1333-46.

SUBRAMANIAN, M., WOJTUSCISZYN, A., FAVRE, L., BOUGHORBEL, S., SHAN, J., LETAIEF, K. B., PITTELOUD, N. & CHOUCHANE, L. 2020.

Precision medicine in the era of artificial intelligence: implications in chronic disease management. *J Transl Med,* 18**,** 472.

SWETS, J. A. 1988. Measuring the accuracy of diagnostic systems. *Science,* 240**,** 1285-93.

SZCZERBINSKI, L. & FLOREZ, J. C. 2023. Precision medicine of obesity as an integral part of type 2 diabetes management - past, present, and future. *Lancet Diabetes Endocrinol,* 11**,** 861-878.

TALAYERO, B. G. & SACKS, F. M. 2011. The role of triglycerides in atherosclerosis. *Curr Cardiol Rep,* 13**,** 544-52.

TALMUD, P. J., HINGORANI, A. D., COOPER, J. A., MARMOT, M. G., BRUNNER, E. J., KUMARI, M., KIVIMÄKI, M. & HUMPHRIES, S. E. 2010. Utility of genetic and non-genetic risk factors in prediction of type 2 diabetes: Whitehall II prospective cohort study. *Bmj,* 340**,** b4838.

TAO, T. T., LIN, X. H., TANG, S. J., GUI, W. W., ZHU, W. F. & LI, H. 2022. Association of genetic variants in the Sirt1 and Nrf2 genes with the risk of metabolic syndrome in a Chinese Han population. *BMC Endocr Disord,* 22**,** 84.

TASKINEN, M. R. 1987. Lipoprotein lipase in diabetes. *Diabetes Metab Rev,* 3**,** 551-70.

TAYLOR, J. Y., KRAJA, A. T., DE LAS FUENTES, L., STANFILL, A. G., CLARK, A. & CASHION, A. 2013. An overview of the genomics of metabolic syndrome. *J Nurs Scholarsh,* 45**,** 52-9.

TEKOLA-AYELE, F., DOUMATEY, A. P., SHRINER, D., BENTLEY, A. R., CHEN, G., ZHOU, J., FASANMADE, O., JOHNSON, T., OLI, J., OKAFOR, G., EGHAN, B. A., JR., AGYENIM-BOATENG, K., ADEBAMOWO, C., AMOAH, A., ACHEAMPONG, J., ADEYEMO, A. & ROTIMI, C. N. 2015. Genome-wide association study identifies African-ancestry specific variants for metabolic syndrome. *Mol Genet Metab,* 116**,** 305-13.

THIND, H., FAVA, J. L., GUTHRIE, K. M., STROUD, L., GOPALAKRISHNAN, G., SILLICE, M., GIDRON, N. & BOCK, B. C. 2018. Yoga as a Complementary Therapy for Adults with Type 2 Diabetes: Design and Rationale of the Healthy, Active, and in Control (HA1C) Study. *Int J Yoga Therap,* 28**,** 123-132.

THOMPSON, R. & MCNAMEE, M. J. 2017. Consent, ethics and genetic biobanks: the case of the Athlome project. *BMC Genomics,* 18**,** 830.

TRUONG, S., TRAN, N. Q., MA, P. T., HOANG, C. K., LE, B. H., DINH, T., TRAN, L., TRAN, T. V., GIA LE, L. H., VU, H. A., MAI, T. P. & DO, M. D. 2022. Association of ADIPOQ Single-Nucleotide Polymorphisms with the Two Clinical Phenotypes Type 2 Diabetes Mellitus and Metabolic Syndrome in a Kinh Vietnamese Population. *Diabetes Metab Syndr Obes,* 15**,** 307-319.

TRUSHEIM, M. R. & BERNDT, E. R. 2015. The clinical benefits, ethics, and economics of stratified medicine and companion diagnostics. *Drug Discov Today,* 20**,** 1439-50.

TRUSHEIM, M. R., BURGESS, B., HU, S. X., LONG, T., AVERBUCH, S. D., FLYNN, A. A., LIEFTUCHT, A., MAZUMDER, A., MILLOY, J., SHAW, P. M., SWANK, D., WANG, J., BERNDT, E. R., GOODSAID, F. & PALMER, M. C. 2011. Quantifying factors for the success of stratified medicine. *Nat Rev Drug Discov,* 10**,** 817-33.

TU, J., GAN, Y., LIANG, T., HU, Q., WANG, Q., REN, T., SUN, Q., WAN, H. & WANG, P. 2018. Graphene FET Array Biosensor Based on ssDNA Aptamer for Ultrasensitive Hg(2+) Detection in Environmental Pollutants. *Front Chem,* 6**,** 333.

TU, X., TU, J., WEN, X., WANG, J. & ZHANG, D. 2005. A study of lipoprotein lipase gene intron 8 polymorphisms in Chinese Han race essential hypertension patients. *Int J Cardiol,* 99**,** 263-7.

TUMIN, D., BREWER, K. L., CUMMINGS, D. M., KEENE, K. L. & CAMPBELL, K. M. 2022. Estimating clinical research project duration from idea to publication. *J Investig Med,* 70**,** 108-109.

UNGER, R. H. & ZHOU, Y. T. 2001. Lipotoxicity of beta-cells in obesity and in other causes of fatty acid spillover. *Diabetes,* 50 Suppl 1**,** S118-21.

UNNIKRISHNAN, R., PRADEEPA, R., JOSHI, S. R. & MOHAN, V. 2017. Type 2 Diabetes: Demystifying the Global Epidemic. *Diabetes,* 66**,** 1432-1442.

VARGHESE, M. P., BALAKRISHNAN, R. & PAILOOR, S. 2018. Association between a guided meditation practice, sleep and psychological well-being in type 2 diabetes mellitus patients. *J Complement Integr Med,* 15.

VELAZQUEZ-ROMAN, J., ANGULO-ZAMUDIO, U. A., LEÓN-SICAIROS, N., MEDINA-SERRANO, J., DELIRA-BUSTILLOS, N., VILLAMIL-RAMÍREZ, H., CANIZALES-QUINTEROS, S., MACÍAS-KAUFFER, L., CAMPOS-ROMERO, A., ALCÁNTAR-FERNÁNDEZ, J. & CANIZALEZ-ROMAN, A. 2021. Association of FTO, ABCA1, ADRB3, and PPARG variants with obesity, type 2 diabetes, and metabolic syndrome in a Northwest Mexican adult population. *J Diabetes Complications,* 35**,** 108025.

VERLINDEN, M., NYS, H., ECTORS, N. & HUYS, I. 2016. Qualitative study on custodianship of human biological material and data stored in biobanks. *BMC Med Ethics,* 17**,** 15.

VERSCHUREN, J. J., TROMPET, S., WESSELS, J. A., GUCHELAAR, H. J., DE MAAT, M. P., SIMOONS, M. L. & JUKEMA, J. W. 2012. A systematic review on pharmacogenetics in cardiovascular disease: is it ready for clinical application? *Eur Heart J,* 33**,** 165-75.

VIRANI, A. H. & LONGSTAFF, H. 2015. Ethical considerations in biobanks: how a public health ethics perspective sheds new light on old controversies. *J Genet Couns,* 24**,** 428-32.

VISHRAM, J. K., HANSEN, T. W., TORP-PEDERSEN, C., MADSBAD, S., JØRGENSEN, T., FENGER, M., LYNGBÆK, S. & JEPPESEN, J. 2016. Relationship Between Two Common Lipoprotein Lipase Variants and the Metabolic Syndrome and Its Individual Components. *Metab Syndr Relat Disord,* 14**,** 442-448.

VOIGHT, B. F., SCOTT, L. J., STEINTHORSDOTTIR, V., MORRIS, A. P., DINA, C., WELCH, R. P., ZEGGINI, E., HUTH, C., AULCHENKO, Y. S., THORLEIFSSON, G., MCCULLOCH, L. J., FERREIRA, T., GRALLERT, H., AMIN, N., WU, G., WILLER, C. J., RAYCHAUDHURI, S., MCCARROLL, S. A., LANGENBERG, C., HOFMANN, O. M., DUPUIS, J., QI, L., SEGRÈ, A. V., VAN HOEK, M., NAVARRO, P., ARDLIE, K., BALKAU, B., BENEDIKTSSON, R., BENNETT, A. J., BLAGIEVA, R., BOERWINKLE, E., BONNYCASTLE, L. L.,

BENGTSSON BOSTRÖM, K., BRAVENBOER, B., BUMPSTEAD, S., BURTT, N. P., CHARPENTIER, G., CHINES, P. S., CORNELIS, M., COUPER, D. J., CRAWFORD, G., DONEY, A. S., ELLIOTT, K. S., ELLIOTT, A. L., ERDOS, M. R., FOX, C. S., FRANKLIN, C. S., GANSER, M., GIEGER, C., GRARUP, N., GREEN, T., GRIFFIN, S., GROVES, C. J., GUIDUCCI, C., HADJADJ, S., HASSANALI, N., HERDER, C., ISOMAA, B., JACKSON, A. U., JOHNSON, P. R., JØRGENSEN, T., KAO, W. H., KLOPP, N., KONG, A., KRAFT, P., KUUSISTO, J., LAURITZEN, T., LI, M., LIEVERSE, A., LINDGREN, C. M., LYSSENKO, V., MARRE, M., MEITINGER, T., MIDTHJELL, K., MORKEN, M. A., NARISU, N., NILSSON, P., OWEN, K. R., PAYNE, F., PERRY, J. R., PETERSEN, A. K., PLATOU, C., PROENÇA, C., PROKOPENKO, I., RATHMANN, W., RAYNER, N. W., ROBERTSON, N. R., ROCHELEAU, G., RODEN, M., SAMPSON, M. J., SAXENA, R., SHIELDS, B. M., SHRADER, P., SIGURDSSON, G., SPARSØ, T., STRASSBURGER, K., STRINGHAM, H. M., SUN, Q., SWIFT, A. J., THORAND, B., et al. 2010. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet,* 42**,** 579-89.

VRABLIK, M., DLOUHA, D., TODOROVOVA, V., STEFLER, D. & HUBACEK, J. A. 2021. Genetics of Cardiovascular Disease: How Far Are We from Personalized CVD Risk Prediction and Management? *Int J Mol Sci,* 22.

WALDEN, R. & TOMLINSON, B. 2011. Cardiovascular Disease. *In:* BENZIE, I. F. F. & WACHTEL-GALOR, S. (eds.) *Herbal Medicine: Biomolecular and Clinical Aspects.* Boca Raton (FL): CRC Press/Taylor & Francis

WALSH, R., LAHROUCHI, N., TADROS, R., KYNDT, F., GLINGE, C., POSTEMA, P. G., AMIN, A. S., NANNENBERG, E. A., WARE, J. S., WHIFFIN, N., MAZZAROTTO, F., ŠKORIĆ-MILOSAVLJEVIĆ, D., KRIJGER, C., ARBELO, E., BABUTY, D., BARAJAS-MARTINEZ, H., BECKMANN, B. M., BÉZIEAU, S., BOS, J. M., BRECKPOT, J., CAMPUZANO, O., CASTELLETTI, S., CELEN, C., CLAUSS, S., CORVELEYN, A., CROTTI, L., DAGRADI, F., DE ASMUNDIS, C., DENJOY, I., DITTMANN, S., ELLINOR, P. T., ORTUÑO, C. G., GIUSTETTO, C., GOURRAUD, J. B., HAZEKI, D., HORIE, M., ISHIKAWA, T., ITOH, H., KANEKO, Y., KANTERS, J. K., KIMOTO, H., KOTTA, M. C., KRAPELS, I. P. C., KURABAYASHI, M., LAZARTE, J., LEENHARDT, A., LOEYS, B. L., LUNDIN, C., MAKIYAMA, T., MANSOURATI, J., MARTINS, R. P., MAZZANTI, A., MÖRNER, S., NAPOLITANO, C., OHKUBO, K., PAPADAKIS, M., RUDIC, B., MOLINA, M. S., SACHER, F., SAHIN, H., SARQUELLA-BRUGADA, G., SEBASTIANO, R., SHARMA, S., SHEPPARD, M. N., SHIMAMOTO, K., SHOEMAKER, M. B., STALLMEYER, B., STEINFURT, J., TANAKA, Y., TESTER, D. J., USUDA, K., VAN DER ZWAAG, P. A., VAN DOOREN, S., VAN LAER, L., WINBO, A., WINKEL, B. G., YAMAGATA, K., ZUMHAGEN, S., VOLDERS, P. G. A., LUBITZ, S. A., ANTZELEVITCH, C., PLATONOV, P. G., ODENING, K. E., RODEN, D. M., ROBERTS, J. D., SKINNER, J. R., TFELT-HANSEN, J., VAN DEN BERG, M. P., OLESEN, M. S., LAMBIASE, P. D., BORGGREFE, M., HAYASHI, K., RYDBERG, A.,

NAKAJIMA, T., YOSHINAGA, M., SAENEN, J. B., KÄÄB, S., BRUGADA, P., ROBYNS, T., GIACHINO, D. F., et al. 2021. Enhancing rare variant interpretation in inherited arrhythmias through quantitative analysis of consortium disease cohorts and population controls. *Genet Med,* 23**,** 47-58.

WANG, C. S., HARTSUCK, J. & MCCONATHY, W. J. 1992. Structure and functional properties of lipoprotein lipase. *Biochim Biophys Acta,* 1123**,** 1-17.

WANG, J., DU, S., WANG, J., ZHU, M., WEN, X. & YANG, W. 2017. Association of the lipoprotein lipase gene Ser447Ter polymorphism with hypertension and blood pressure variation: evidence from an updated meta-analysis. *Clin Exp Hypertens,* 39**,** 655-664.

WEINHOLD, B. 2006. Epigenetics: the science of change. *Environ Health Perspect,* 114**,** A160-7.

WILLEMS, E. L., WAN, J. Y., NORDEN-KRICHMAR, T. M., EDWARDS, K. L. & SANTORICO, S. A. 2020. Transethnic meta-analysis of metabolic syndrome in a multiethnic study. *Genet Epidemiol,* 44**,** 16-25.

WILSON, K., FRY, G. L., CHAPPELL, D. A., SIGMUND, C. D. & MEDH, J. D. 2001. Macrophage-specific expression of human lipoprotein lipase accelerates atherosclerosis in transgenic apolipoprotein e knockout mice but not in C57BL/6 mice. *Arterioscler Thromb Vasc Biol,* 21**,** 1809-15.

WU, Y., CHENG, S., GU, H., YANG, K., XU, Z., MENG, X., WANG, Y., JIANG, Y., LI, H., ZHOU, Y. & WANG, Y. 2023. Variants within the LPL gene confer susceptility to diabetic kidney disease and rapid decline in kidney function in Chinese patients with type 2 diabetes. *Diabetes Obes Metab,* 25**,** 3012-3019.

XIA, J., XU, T., QING, J., WANG, L. & TANG, J. 2021. Detection of Single Nucleotide Polymorphisms by Fluorescence Embedded Dye SYBR Green I Based on Graphene Oxide. *Front Chem,* 9**,** 631959.

XIE, C., WANG, Z. C., LIU, X. F. & YANG, M. S. 2010. The common biological basis for common complex diseases: evidence from lipoprotein lipase gene. *Eur J Hum Genet,* 18**,** 3-7.

YAMADA, Y., MATSUO, H., WARITA, S., WATANABE, S., KATO, K., OGURI, M., YOKOI, K., METOKI, N., YOSHIDA, H., SATOH, K., ICHIHARA, S., AOYAGI, Y., YASUNAGA, A., PARK, H., TANAKA, M. & NOZAWA, Y. 2007. Prediction of genetic risk for dyslipidemia. *Genomics,* 90**,** 551-8.

YANG, L., CHEN, Y., YU, C. & SHEN, B. 2016. Biobanks and Their Clinical Application and Informatics Challenges. *Adv Exp Med Biol,* 939**,** 241-257.

YANG, W., HUANG, J., GE, D., YAO, C., DUAN, X., SHEN, Y., QIANG, B. & GU, D. 2004. Lipoprotein lipase gene is in linkage with blood pressure phenotypes in Chinese pedigrees. *Hum Genet,* 115**,** 8-12.

YANG, W., HUANG, J., YAO, C., SU, S., LIU, D., GE, D. & GU, D. 2005. Linkage and linkage disequilibrium analysis of the lipoprotein lipase gene with lipid profiles in Chinese hypertensive families. *Clin Sci (Lond),* 108**,** 137-42.

YANG, W. J., HUANG, J. F., YAO, C. L., FAN, Z. J., GE, D. L., GAN, W. Q., HUANG, G. Y., HUI, R. T., SHEN, Y., QIANG, B. Q. & GU, D. F. 2003.

Evidence for linkage and association of the markers near the LPL gene with hypertension in Chinese families. *J Med Genet,* 40**,** e57.

YE, S., DHILLON, S., KE, X., COLLINS, A. R. & DAY, I. N. 2001. An efficient procedure for genotyping single nucleotide polymorphisms. *Nucleic Acids Res,* 29**,** E88-8.

YUSUF, S., HAWKEN, S., OUNPUU, S., DANS, T., AVEZUM, A., LANAS, F., MCQUEEN, M., BUDAJ, A., PAIS, P., VARIGOS, J. & LISHENG, L. 2004. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet,* 364**,** 937-52.

ZABANEH, D. & BALDING, D. J. 2010. A genome-wide association study of the metabolic syndrome in Indian Asian men. *PLoS One,* 5**,** e11961.

ZEE, R. Y., COOK, N. R., CHENG, S., ERLICH, H. A., LINDPAINTNER, K. & RIDKER, P. M. 2006. Polymorphism in the beta2-adrenergic receptor and lipoprotein lipase genes as risk determinants for idiopathic venous thromboembolism: a multilocus, population-based, prospective genetic analysis. *Circulation,* 113**,** 2193-200.

ZHENG, Y., LEY, S. H. & HU, F. B. 2018. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. *Nat Rev Endocrinol,* 14**,** 88-98.

ZHOU, K., PEDERSEN, H. K., DAWED, A. Y. & PEARSON, E. R. 2016. Pharmacogenomics in diabetes mellitus: insights into drug action and drug discovery. *Nat Rev Endocrinol,* 12**,** 337-46.

ZHU, Y., ZHANG, D., ZHOU, D., LI, Z., LI, Z., FANG, L., YANG, M., SHAN, Z., LI, H., CHEN, J., ZHOU, X., YE, W., YU, S., LI, H., CAI, L., LIU, C., ZHANG, J., WANG, L., LAI, Y., RUAN, L., SUN, Z., ZHANG, S., WANG, H., LIU, Y., XU, Y., LING, J., XU, C., ZHANG, Y., LV, D., YUAN, Z., ZHANG, J., ZHANG, Y., SHI, Y. & LAI, M. 2017. Susceptibility loci for metabolic syndrome and metabolic components identified in Han Chinese: a multi-stage genome-wide association study. *J Cell Mol Med,* 21**,** 1106-1116.

ZUKIN, E., CULVER, J. O., LIU, Y., YANG, Y., RICKER, C. N., HODAN, R., STURGEON, D., KINGHAM, K., CHUN, N. M., ROWE-TEETER, C., SINGH, K., ZELL, J. A., LADABAUM, U., MCDONNELL, K. J., FORD, J. M., PARMIGIANI, G., BRAUN, D., KURIAN, A. W., GRUBER, S. B. & IDOS, G. E. 2023. Clinical implications of conflicting variant interpretations in the cancer genetics clinic. *Genet Med,* 25**,** 100837.

1. Metabolic Syndrome Criteria

1A. Metabolic Syndrome Criteria Summary. Summary of diagnostic criteria for metabolic syndrome proposed by major international organisations. Criteria include measures of abdominal/central obesity, lipid abnormalities, blood pressure, fasting glucose, and other metabolic risk factors, with variations in thresholds across definitions. Adapted from Kassi et al., 2011.

Criteria for metabolic syndrome (MetS) definitions in adults

World Health Organization criteria (1998)
Insulin resistance is defined as type 2 diabetes mellitus (DM) or impaired fasting glucose (IFG) (> 100 mg/dl) or impaired glucose tolerance (IGT), plus two of the following:
- Abdominal obesity (waist-to-hip ratio > 0.9 in men or > 0.85 in women, or body mass index (BMI) > 30 kg/m^2.
- Triglycerides 150 mg/dl or greater, and/or high-density lipoprotein (HDL)-cholesterol < 40 mg/dl in men and < 50 mg/dl in women.
- Blood pressure (BP) 140/90 mmHg or greater.
- Microalbuminuria (urinary albumin secretion rate 20 µg/min or greater, or albumin-to-creatinine ratio 30 mg/g or greater).

European Group for the Study of Insulin Resistance criteria (1999)
Insulin resistance defined as insulin levels > 75th percentile of non-diabetic patients, plus two of the following:
- Waist circumference 94 cm or greater in men, 80 cm or greater in women.
- Triglycerides 150 mg/dl or greater and/or HDL-cholesterol < 39 mg/dl in men or women.
- BP 140/90 mmHg or greater or taking antihypertensive drugs.
- Fasting glucose 110 mg/dl or greater.

National Cholesterol Education Program Adult Treatment Panel III (NCEP:ATPIII) criteria (2001)
Any three or more of the following:
- Waist circumference > 102 cm in men, > 88 cm in women.
- Triglycerides 150 mg/dl or greater.
- HDL-cholesterol < 40 mg/dl in men and < 50 mg/dl in women.
- BP 130/85 mmHg or greater.
- Fasting glucose 110 mg/dl* or greater.

* In 2003, the American Diabetes Association (ADA) changed the criteria for IFG tolerance from 110 mg/dl to 100 mg/dl.

American Association of Clinical Endocrinology criteria (2003)
IGT plus two or more of the following:
- BMI 25 kg/m^2 or greater.
- Triglycerides 150 mg/dl or greater and/or HDL-cholesterol < 40 mg/dl in men and < 50 mg/dl in women.
- BP 130/85 mmHg or greater.

American Heart Association/National Heart, Lung, and Blood Institute (AHA/NHLBI) criteria (2004)
Any three of the following:
- Waist circumference 102 cm or greater in men, 88 cm or greater in women.
- Triglycerides 150 mg/dl or greater.
- HDL-cholesterol < 40 mg/dl in men and < 50 mg/dl in women.
- BP 130/85 mmHg or greater.
- Fasting glucose 100 mg/dl or greater.

International Diabetes Federation (IDF) criteria (2005)
Central obesity (defined as waist circumference but can be assumed if BMI > 30 kg/m^2) with ethnicity-specific values,* plus two of the following:
- Triglycerides 150 mg/dl or greater.
- HDL-cholesterol < 40 mg/dl in men and < 50 mg/dl in women.
- BP 130/85 mmHg or greater.
- Fasting glucose 100 mg/dl or greater.

*To meet the criteria, waist circumference must be: for Europeans, > 94 cm in men and > 80 cm in women; and for South Asians, Chinese, and Japanese, > 90 cm in men and > 80 cm in women. For ethnic South and Central Americans, South Asian data are used, and for sub-Saharan Africans and Eastern Mediterranean and Middle East (Arab) populations, European data are used.

Consensus definition (incorporating IDF and AHA/NHLBI definitions) (2009)
Any three of the following:
- Elevated waist circumference (according to population and country-specific definitions).
- Triglycerides 150 mg/dl or greater.
- HDL-cholesterol < 40 mg/dl in men and < 50 mg/dl in women.
- BP 130/85 mmHg or greater.
- Fasting glucose 100 mg/dl or greater.

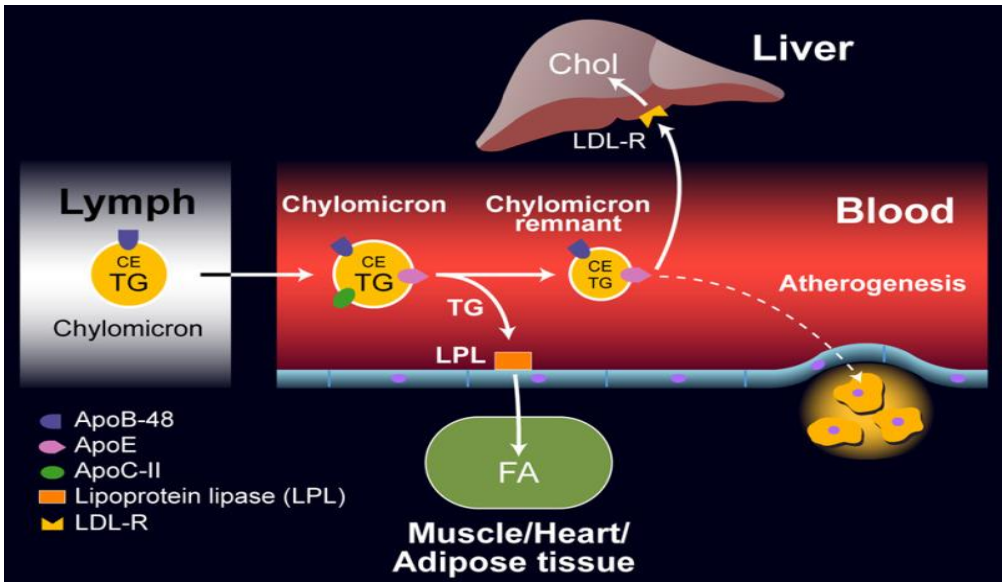
1B. Metabolic syndrome criteria comparison. Comparison of diagnostic criteria for metabolic syndrome from various organizations, outlining the specific criteria for central obesity, blood glucose, high triglycerides (TG), low high-density lipoprotein (HDL), and high blood pressure (BP), as well as a number of criteria required for a diagnosis. Adapted from Fahed, et al., 2022.

Clinical Measure	Criteria					Diagnosis
	Central Obesity	Blood Glc	High TG	Low HDL	High BP	
AHA/ NHLBI (2009) [4]	• WC >40″ (men) or >35″ (women)	• IFG or • on high blood Glc txt or • T2DM dx	• ≥150 mg/dL or • on TG txt	• <40 mg/dL (men) or <50 mg/dL (women) or • on HDL txt	• ≥130 mmHg systolic and/or • ≥85 mmHg diastolic or • on HTN txt	≥3 criteria
IDF (2005) [5,6]	• WC >37″ (men) or >32″ (women) or • BMI >30 kg/m^2					≥3 criteria one of which should be central obesity
ATPIII (2001) [7]	WC >40″ (men) or >35″ (women)		• ≥150 mg/dL	• <40 mg/dL (men) or <50 mg/dL (women)	• ≥130 mmHg systolic and ≥85 mmHg diastolic or • on HTN txt	≥3 criteria
EGIR (1999) [8]	• WC >37″ (men) or >32″ (women)	• IFG or • IGT		• <39 mg/dL (men and women)	• ≥140 mmHg systolic and ≥90 mmHg diastolic or • on HTN txt	≥3 criteria one of which should be IR *
WHO (1998) [1]	• Waist/hip ratio > 0.9 (men) or > 0.85 (women) or • BMI > 30 kg/m^2	• IFG orx • IIGT or • T2DM dx		• <35 mg/dL (men) or <39 mg/dL (women)	• ≥140 mmHg systolic and ≥90 mmHg diastolic	≥3 criteria one of which should be IR **

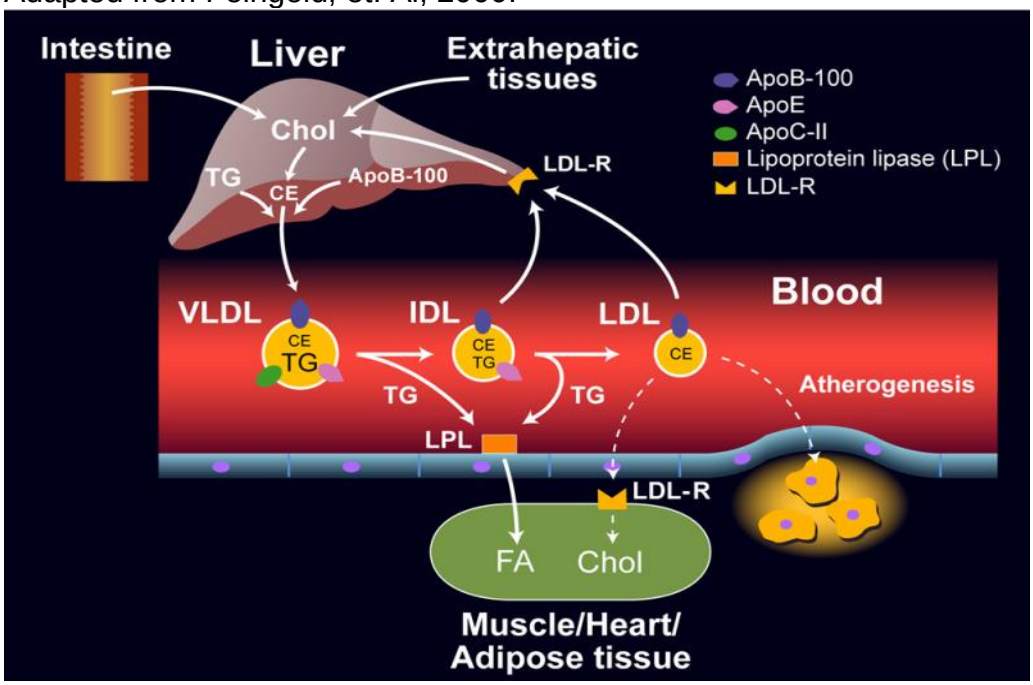
Note that IFG is defined as ≥110 mg/dL in 2001 but this was momdified in 2004 to be ≥100 mg/dL, IGT is defined as 2 h glucose >140 mg/dL. * EGIR IR is defined as plasma insulin levels >75th percentile. ** WHO IR is defined as presence of IR or IFG or IGT. Abbreviations: AHA: American Heart Association, ATPIII: National Cholesterol Education Program Adult Treatment Panel III; dx: diagnosis; EGIR: European group for study of insulin resistance; Glc: glucose; HDL: high density lipoprotein; HTN: hypertension; IR: insulin resistance; IDF: International Diabetes Federation; IGT: impaired glucose tolerance; IFG: impaired fasting glucose; NHLBI: National Heart, Lung, and Blood Institute; TG: triglyceride; txt: treatment; WC: waist circumference; WHO: World Health Organization.

2. Lipid Metabolism Pathway

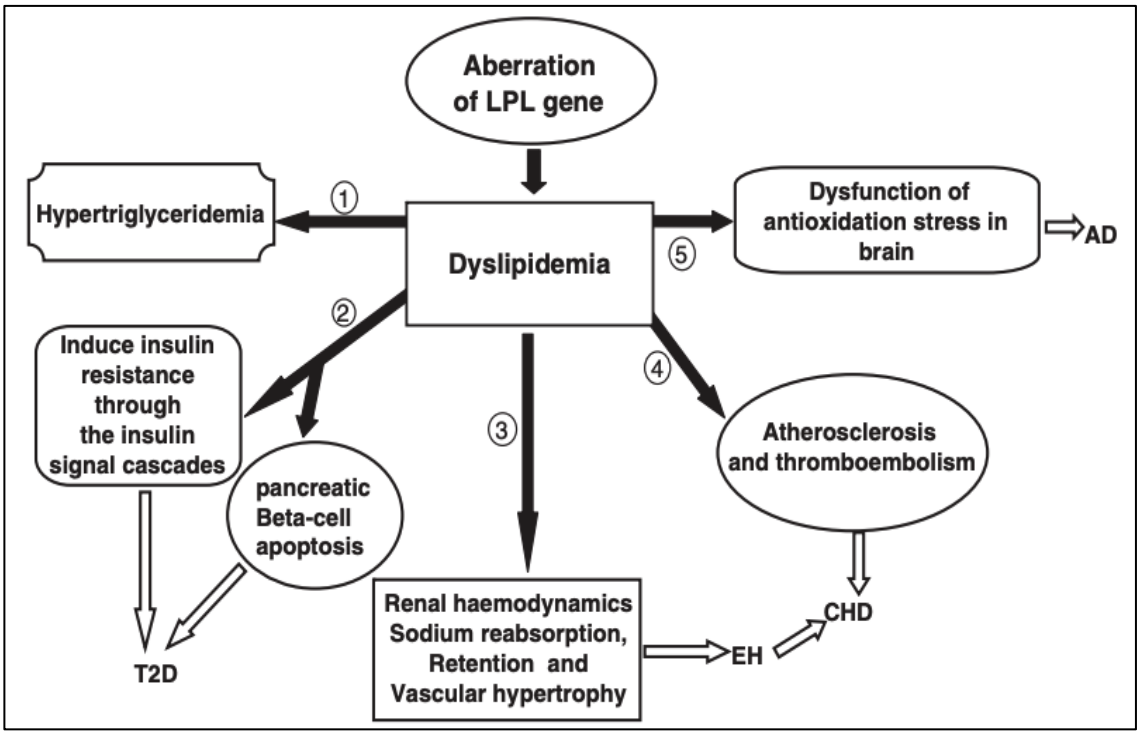
2A. Exogenous lipoprotein pathway. Chylomicrons, synthesized in the intestine from dietary lipids, are secreted into the lymph and then the bloodstream. Lipoprotein lipase (LPL) on the capillary surface hydrolyzes triglycerides (TG) from the chylomicrons, releasing fatty acids (FA) for use by muscle, heart, and adipose tissue. The resulting cholesterol ester (CE)-rich chylomicron remnants are then taken up by the liver via the LDL receptor (LDL-R). Adapted from Feingold, et al., 2000.



2B. Endogenous lipoprotein pathway. The liver synthesizes very-low-density lipoprotein (VLDL) particles containing triglycerides (TG) and secretes them into the bloodstream. VLDL particles deliver TG to extrahepatic tissues via lipoprotein lipase (LPL). As VLDL particles lose TG, they become intermediate-density lipoprotein (IDL) and then low-density lipoprotein (LDL), which is rich in cholesterol esters (CE). LDL is taken up by extrahepatic tissues and the liver via the LDL receptor (LDL-R). Adapted from Feingold, et. Al, 2000.



3. Schematic on the relations among *LPL* aberration, T2D, CHD. Aberration in lipoprotein lipase (LPL) gene may lead to dyslipidemia, being the shared intermediate process on the development of the presented diseases. Hyperglyceridemia, a feature of dyslipidemia, is involved in regulating the B-cell function, pancreatic B-cell apoptosis and the insulin signal cascades, which aggravates insulin resistance, which then induces Type 2 diabetes (T2D). Hyperinsulinemia induced by elevated triglycerides (TG), may be partly responsible for essential hypertension (EH) by its functions on renal hemodynamics, sodium re-absorption, retention and vascular hypertrophy. Elevated TG is associated with atherosclerosis and thromboembolism - both of these are high-risk factors for the development of coronary heart disease (CHD). Dyslipidemia causing dysfunction of antioxidation stress in the brain may cause Alzheimer's disease (AD). Adapted from Xie, et al., 2010.



4. *LPL* SNPs in this Study.

Screenshots of the LPL SNPs in the study from the National Institutes of Health National Library of Medicine National Center for Biotechnology Information, detailing the variant as a single nucleotide variant (SNV), and pertinent details as minor allele frequency (MAF) and Human Genome Variation Society (HGVS) nomenclature.

rs268 *[Homo sapiens]*

1.

Variant type:	SNV
Alleles:	A>G [Show Flanks]
Chromosome:	8:19956018 (GRCh38)
	8:19813529 (GRCh37)
Canonical SPDI:	NC_000008.11:19956017:A:G
Gene:	LPL (Varview)
Functional Consequence:	missense_variant,coding_sequence_variant
Clinical significance:	
conflicting-interpretations-of-pathogenicity,uncertain-significance,pathogenic,benign,risk-factor	
Validated:	by frequency,by alfa,by cluster
MAF:	G=0.016273/3560 (ALFA)
	G=0.000035/1 (TOMMO)
	G=0.004685/23 (1000Genomes)
	...more
HGVS:	NC_000008.11:g.19956018A>G, NC_000008.10:g.19813529A>G, NG_008855.2:g.59302A>G, NM_000237.3:c.953A>G, NM_000237.2:c.953A>G, NP_000228.1:p.Asn318Ser

rs11542065 *[Homo sapiens]*

2.

Variant type:	SNV
Alleles:	C>G,T [Show Flanks]
Chromosome:	8:19948304 (GRCh38) 8:19805815 (GRCh37)
Canonical SPDI:	NC_000008.11:19948303:C:G,NC_000008.11:19948303:C:T
Gene:	LPL (Varview)
Functional Consequence:	missense_variant,synonymous_variant,coding_sequence_variant
Clinical significance:	conflicting-interpretations-of-pathogenicity,likely-benign
Validated:	by frequency,by alfa,by cluster
MAF:	G=0.00018/16 (ALFA) T=0./0 (TWINSUK) T=0.00026/1 (ALSPAC) ...more
HGVS:	NC_000008.11:g.19948304C>G, NC_000008.11:g.19948304C>T, NC_000008.10:g.19805815C>G, NC_000008.10:g.19805815C>T, NG_008855.2:g.51588C>G, NG_008855.2:g.51588C>T, NM_000237.3:c.213C>G, NM_000237.3:c.213C>T, NM_000237.2:c.213C>G, NM_000237.2:c.213C>T, NP_000228.1:p.His71Gln

☐ rs116403115 *[Homo sapiens]*

3.

Variant type:	SNV
Alleles:	T>C,G [Show Flanks]
Chromosome:	8:19962117 (GRCh38) 8:19819628 (GRCh37)
Canonical SPDI:	NC_000008.11:19962116:T:C,NC_000008.11:19962116:T:G
Gene:	LPL (Varview)
Functional Consequence:	coding_sequence_variant,missense_variant
Clinical significance:	uncertain-significance,conflicting-interpretations-of-pathogenicity
Validated:	by frequency,by alfa,by cluster
MAF:	G=0.000535/104 (ALFA) G=0.000156/1 (1000Genomes) G=0.000267/21 (PAGE_STUDY) ...more
HGVS:	NC_000008.11:g.19962117T>C, NC_000008.11:g.19962117T>G, NC_000008.10:g.19819628T>C, NC_000008.10:g.19819628T>G, NG_008855.2:g.65401T>C, NG_008855.2:g.65401T>G, NM_000237.3:c.1325T>C, NM_000237.3:c.1325T>G, NM_000237.2:c.1325T>C, NM_000237.2:c.1325T>G, NP_000228.1:p.Val442Ala, NP_000228.1:p.Val442Gly

rs118204057 *[Homo sapiens]*

4.

Variant type:	SNV
Alleles:	G>A,C [Show Flanks]
Chromosome:	8:19954222 (GRCh38) 8:19811733 (GRCh37)
Canonical SPDI:	NC_000008.11:19954221:G:A,NC_000008.11:19954221:G:C
Gene:	LPL (Varview)
Functional Consequence:	coding_sequence_variant,missense_variant
Clinical significance:	pathogenic-likely-pathogenic,conflicting-interpretations-of-pathogenicity,pathogenic
Validated:	by frequency,by alfa,by cluster
MAF:	A=0.000271/54 (ALFA) A=0.000132/16 (ExAC) A=0.000142/2 (TOMMO)
	...more
HGVS:	NC_000008.11:g.19954222G>A, NC_000008.11:g.19954222G>C, NC_000008.10:g.19811733G>A, NC_000008.10:g.19811733G>C, NG_008855.2:g.57506G>A, NG_008855.2:g.57506G>C, NM_000237.3:c.644G>A, NM_000237.3:c.644G>C, NM_000237.2:c.644G>A, NM_000237.2:c.644G>C, NP_000228.1:p.Gly215Glu, NP_000228.1:p.Gly215Ala

rs118204061 *[Homo sapiens]*

5.

Variant type:	SNV
Alleles:	T>C [Show Flanks]
Chromosome:	8:19954240 (GRCh38) 8:19811751 (GRCh37)
Canonical SPDI:	NC_000008.11:19954239:T:C
Gene:	LPL (Varview)
Functional Consequence:	coding_sequence_variant,missense_variant
Clinical significance:	pathogenic-likely-pathogenic,conflicting-interpretations-of-pathogenicity,pathogenic
Validated:	by frequency,by alfa,by cluster
MAF:	C=0.000062/7 (ALFA) C=0./0 (PAGE_STUDY) C=0.000008/1 (ExAC)
	...more
HGVS:	NC_000008.11:g.19954240T>C, NC_000008.10:g.19811751T>C, NG_008855.2:g.57524T>C, NM_000237.3:c.662T>C, NM_000237.2:c.662T>C, NP_000228.1:p.Ile221Thr

rs144466625 *[Homo sapiens]*

6.

Variant type:	SNV
Alleles:	G>A [Show Flanks]
Chromosome:	8:19956063 (GRCh38) 8:19813574 (GRCh37)
Canonical SPDI:	NC_000008.11:19956062:G:A
Gene:	LPL (Varview)
Functional Consequence:	coding_sequence_variant,missense_variant
Clinical significance:	conflicting-interpretations-of-pathogenicity,pathogenic
Validated:	by frequency,by alfa,by cluster
MAF:	A=0.000174/4 (ALFA) A=0.000048/12 (GnomAD_exomes) A=0.000049/6 (ExAC) ...more
HGVS:	NC_000008.11:g.19956063G>A, NC_000008.10:g.19813574G>A, NG_008855.2:g.59347G>A, NM_000237.3:c.998G>A, NM_000237.2:c.998G>A, NP_000228.1:p.Arg333His

rs547644955 *[Homo sapiens]*

7.

Variant type:	DELINS
Alleles:	T>-,TT [Show Flanks]
Chromosome:	8:19948166 (GRCh38) 8:19805677 (GRCh37)
Canonical SPDI:	NC_000008.11:19948165:TTTTT:TTTT,NC_000008.11:19948165:TTTTT:TTTTTT
Gene:	LPL (Varview)
Functional Consequence:	intron_variant
Clinical significance:	conflicting-interpretations-of-pathogenicity
Validated:	by frequency,by alfa,by cluster
MAF:	TTTT=0.00204/47 (ALFA) -=0.000939/236 (GnomAD_exomes) -=0.001195/145 (ExAC) ...more
HGVS:	NC_000008.11:g.19948170del, NC_000008.11:g.19948170dup, NC_000008.10:g.19805681del, NC_000008.10:g.19805681dup, NG_008855.2:g.51454del, NG_008855.2:g.51454dup

5. Reported Metabolic Syndrome (MetS) Single Nucleotide Polymorphisms (SNPs) from publications in Genome-wide Association Studies (GWAS) Central.

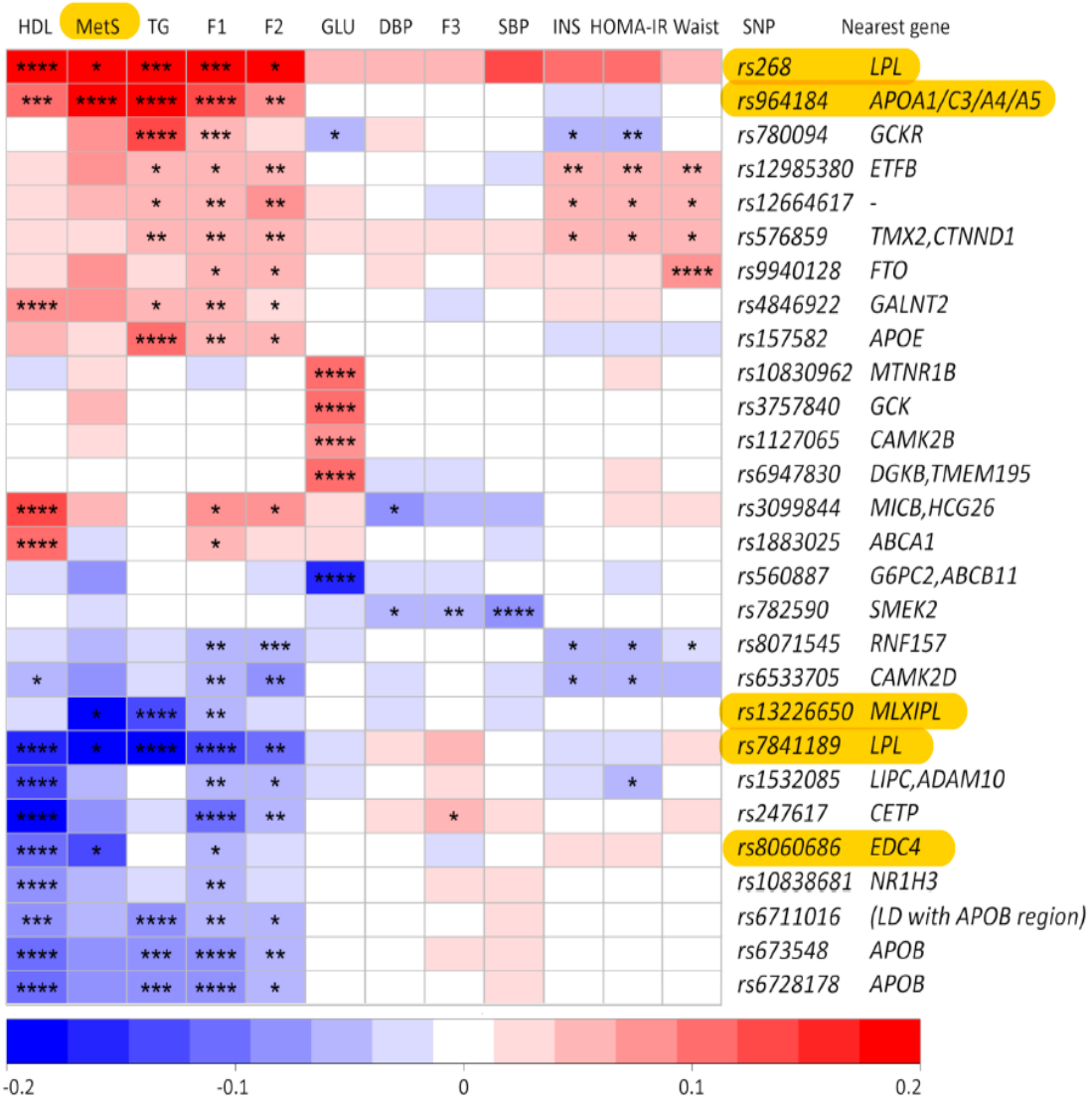
Screenshots as reported from specified publications presented.

Adapted from Kraja, et al., 2011.

A summary of STAMPEED β-meta-analyses of the most significant results

Gene	Variant	Trait	Ch	Position	Meta-β	Meta-SE	Meta *P*	Homog. *P* value	Coded	Genotype
GCKR	rs780093	TG-BP	2	27596107	0.18	0.03	3.0E-10	0.70	A	A/G
GCKR	rs780093	WC-TG	2	27596107	0.19	0.03	1.9E-12	0.64	A	A/G
C2orf16	rs1919128	WC-TG	2	27655263	−0.18	0.03	2.0E-09	0.82	A	A/G
ZNF512	rs13022873	WC-TG	2	27669014	−0.17	0.03	5.0E-09	0.47	A	A/C
CCDC121	rs3749147	WC-TG	2	27705422	−0.18	0.03	1.4E-09	0.79	C	C/T
ABCB11	rs569805	HDLC-GLUC	2	169491126	0.16	0.03	8.5E-08	0.46	A	A/T
TFAP2B	rs2206277	WC-GLUC	6	50906485	0.17	0.03	1.3E-07	0.75	A	A/G
(LOC100129150)	rs9987289	HDLC-TG	8	9220768	0.25	0.04	1.1E-08	0.54	A	A/G
(LOC100129150)	rs9987289	HDLC-WC	8	9220768	0.24	0.04	3.7E-08	0.58	A	A/G
LPL	rs295	MetS	8	19860518	0.17	0.03	1.7E-09	0.47	A	A/C
LPL	rs301	HDLC-WC	8	19861214	−0.22	0.03	3.2E-11	0.58	C	C/T
LPL	rs13702	HDLC-TG	8	19868772	0.29	0.03	1.0E-16	0.67	A	A/G
LPL	rs15285	TG-BP	8	19868947	−0.27	0.04	1.3E-10	0.65	A	A/G
(LPL)	rs2197089	TG-GLUC	8	19870653	0.18	0.03	1.6E-09	1.00	C	C/T
(LPL)	rs1441756	BP-HDLC	8	19912666	−0.18	0.03	2.7E-08	0.43	G	G/T
(TRIB1)	rs2954026	HDLC-TG	8	126553708	−0.16	0.03	7.9E-09	0.46	G	G/T
(TRIB1)	rs2954033	TG-BP	8	126562928	0.17	0.03	8.5E-09	0.55	A	A/G
(LOC100128354)	rs1387153	BP-GLUC	11	92313476	−0.19	0.03	8.1E-09	0.48	C	C/T
(LOC100128354)	rs1387153	HDLC-GLUC	11	92313476	−0.21	0.03	2.4E-09	0.49	C	C/T
(LOC100128354)	rs10830956	TG-GLUC	11	92320661	−0.20	0.03	4.8E-11	0.67	C	C/T
BUD13	rs11825181	TG-BP	11	116131468	0.32	0.05	3.0E-09	0.98	A	A/G
BUD13	rs11820589	TG-GLUC	11	116139072	0.32	0.06	5.5E-09	0.83	A	A/G
BUD13	rs10790162	HDLC-TG	11	116144314	0.38	0.05	2.8E-15	0.44	A	A/G
BUD13	rs10790162	MetS	11	116144314	0.25	0.04	5.4E-09	0.44	A	A/G
BUD13	rs10790162	WC-TG	11	116144314	0.39	0.05	6.6E-16	0.79	A	A/G
ZNF259	rs11823543	TG-BP	11	116154345	0.35	0.06	2.5E-09	1.00	A	A/G
ZNF259	rs12286037	TG-GLUC	11	116157417	−0.32	0.06	1.1E-08	0.86	C	C/T
ZNF259	rs2075290	HDLC-TG	11	116158506	0.39	0.05	1.5E-14	0.39	C	C/T
ZNF259	rs2075290	MetS	11	116158506	0.26	0.04	2.1E-09	0.64	C	C/T
ZNF259	rs2075290	WC-TG	11	116158506	0.41	0.05	1.1E-16	0.94	C	C/T
APOA5	rs2266788	HDLC-TG	11	116165896	0.39	0.05	4.6E-13	0.36	C	C/T
APOA5	rs2266788	MetS	11	116165896	0.26	0.04	1.9E-09	0.66	C	C/T
APOA5	rs2266788	TG-BP	11	116165896	0.37	0.07	3.5E-08	0.18	C	C/T
APOA5	rs2266788	WC-TG	11	116165896	0.41	0.05	2.2E-16	0.92	A	A/G
(LIPC)	rs10468017	HDLC-WC	15	56465804	0.16	0.03	5.5E-08	0.47	C	C/T
(LIPC)	rs2043085	HDLC-GLUC	15	56468246	−0.17	0.03	1.3E-08	0.83	A	A/G
(CETP)	rs173539	HDLC-TG	16	55545545	0.26	0.03	4.5E-16	0.61	C	C/T
(CETP)	rs173539	HDLC-WC	16	55545545	0.29	0.03	1.0E-16	0.65	C	C/T
(CETP)	rs173539	MetS	16	55545545	0.16	0.03	9.1E-09	0.41	C	C/T
(CETP)	rs3764261	BP-HDLC	16	55550825	0.29	0.04	3.3E-13	0.43	G	G/T
CETP	rs9939224	HDLC-GLUC	16	55560233	−0.31	0.05	6.9E-12	0.46	G	G/T
LOC100129500	rs439401	HDLC-TG	19	50106291	0.24	0.04	1.0E-08	0.44	C	C/T

Adapted from Kristiansson, et al., 2012.



Adapted from Shim, et al, 2014.

Table 2. Significant SNPs associated with metabolic syndrome in the discovery phase of GWAS

Chromosome	SNP	BP	Nearby gene	Minor allele	p-value	
					Unadjusted	Bonferroni-adjusted
12	rs11066280	111302166	*C12orf51*	T	1.38E-07	0.043
12	rs2074356	111129784	*C12orf51*	T	4.25E-07	0.133
12	rs12229654	109898844	*MYL2*	G	3.00E-06	0.937

SNP, single-nucleotide polymorphism; GWAS, genome-wide association study; BP, base pair.

Adapted from Lin , et al., 2017.

Table 2: Odds ratio analysis with odds ratios after adjustment for covariates between the MetS and two SNPs (including *APOA5* rs662799 and *COLEC12* rs16944558) with genome-wide significance.

Gene	SNP	Chr	A1	A2	Additive model			Dominant model			Recessive model		
					OR	95% CI	*P*	OR	95% CI	*P*	OR	95% CI	*P*
APOA5	rs662799	11	G	A	1.25	1.14-1.37	3.7×10^{-6}	1.40	1.27-1.56	**1.2×10^{-10}**	1.35	1.13-1.62	0.0012
COLEC12	rs16944558	18	T	C	1.20	1.12-1.30	1.2×10^{-6}	1.40	1.25-1.57	**1.3×10^{-8}**	1.18	1.04-1.34	0.0101

Adapted from Zhu, et al. (2017).

Table 1 SNPs associated with metabolic syndrome with genome-wide significance in combined analyses

Phenotype	SNP	Chr	Position (bp)	Gene	Alleles*	MAF	Stages	*N*	OR (95% CI)	*P*-value	*P*-heterogeneity
MetS	rs651821	11	116 662 579	*APOA5*	C/T	0.28	Discovery	1742	1.30 (1.10, 1.49)	6.1×10^{-06}	–
							Replication I	1580	1.31 (1.05, 1.56)	2.8×10^{-04}	0.18
							Replication II	2494	1.27 (1.09, 1.46)	1.9×10^{-04}	0.06
							Replication III	6113	1.27 (1.16, 1.37)	2.3×10^{-08}	–
							Combined	11 929	1.28 (1.20, 1.36)	4.2×10^{-17}	0.99
MetS	rs671	12	112 241 766	*ALDH2*	A/G	0.29	Discovery	1741	0.68 (0.59, 0.79)	1.0×10^{-05}	–
							Replication I	1581	0.80 (0.63, 0.96)	2.9×10^{-02}	0.96
							Replication II	2359	0.80 (0.67, 0.93)	5.5×10^{-03}	0.56
							Replication III	6759	0.70 (0.64, 0.75)	1.3×10^{-19}	–
							Combined	12 440	0.71 (0.67, 0.76)	5.4×10^{-28}	0.29

Adapted from Lee, et al., 2018.

SNPs	Nearest Gene	Chr	CA/ NCA	MAF	Position	Discovery			Replication			Meta-Analysis		
						P_{MulA}	OR	P_{OR}	P_{MulA}	OR	P_{OR}	P_{MulA}	OR	P_{OR}
rs11065756	CCDC63	12	A/G	0.17	109823177	9.58×10^{-9}	0.88	<0.01	0.01	0.97	0.72	2.43×10^{-9}	0.93	0.04
rs10849915	CCDC63	12	G/A	0.18	109818005	1.82×10^{-8}	0.88	<0.01	0.02	0.96	0.69	7.86×10^{-9}	0.92	0.03
rs17482310	LPL	8	T/C	0.21	19910554	9.55×10^{-18}	0.87	<0.01	3.91×10^{-7}	0.74	<0.01	6.68×10^{-10}	0.80	2.68×10^{-5}
rs3782889	MYL2	12	C/T	0.17	109835038	4.19×10^{-9}	0.87	<0.01	0.01	0.96	0.64	1.48×10^{-9}	0.91	0.02

Table 1. Novel common SNPs[1)] for metabolic syndrome achieving genome-wide significance in Korean population. [1)]SNPs not previously reported in GWAS of any other MetS traits. Chr, chromosome; MAF, minor allele frequency; OR, odds ratio; P_{MulA}, P for multivariate analysis.

Adapted from Moon, et al., 2018.

Table 2 Three SNP pairs that were significant only in the multi-SNP–multi-trait analysis

SNP pair	CHR	Position (hg19)	Gene	Annotation	Single-multi *P* value			Multi-multi *P* value (A Bonferroni-adjusted *P* value of discovery stage = 1.45×10^{-7})		
					Discovery	Replication	Meta *P* value	Discovery	Replication	Meta *P* value
rs7107152	11	117,056,080	*SIDT2*	Intronic	2.59E-01	8.43E-01	5.51E-01	1.40E-08	2.32E-03	8.17E-10
rs1242229	11	117,062,370	*SIDT2*	Intronic	5.39E-03	7.68E-02	3.64E-03			
rs10892876	11	122,540,281	*UBASH3B*	Intronic	2.93E-01	1.27E-01	1.60E-01	3.34E-12	3.82E-03	4.21E-13
rs12290043	11	122,540,528	*UBASH3B*	Intronic	1.62E-01	1.46E-01	1.12E-01			
rs886126	12	111,679,214	*CUX2*	Intronic	3.51E-01	1.79E-01	2.37E-01	5.09E-13	3.35E-03	5.97E-14
rs2078851	12	111,690,579	*CUX2*	Intronic	7.31E-05	6.63E-02	6.42E-05			

Single-multi P value P value from a single-SNP–multi-trait association analysis, *multi-multi P value P* value from a multi-SNP–multi-trait association analysis, *CHR* chromosome, *Meta P value P* value from meta-analysis

Adapted from Kong, et al., 2019.

Table 2 Metabolic syndrome (MetS) associated loci in females. Discovery stage was GWAS for MetS in each sex-stratified group of KARE study. Overall association results (P_{meta}) were obtained from meta-analyses combining Discover (KARE) and Replication (Rural1816, Rural3667 and HEXA) stages

CHR	SNP	BP (GRCh37)	Candidate gene	EA	EAF	Female											Male			Combining female & male		
						Discovery		Replication						Overall								
						KARE		Rural1816		Rural3667		HEXA										
						OR	P_{KARE}	OR	$P_{Rural1816}$	OR	$P_{Rural3667}$	OR	P_{HEXA}	OR	P_{meta}	N	OR	P_{meta}	N	OR	P_{meta}	N
1	rs2209363	187,163,851	*LINC01036*	C	0.18	0.67	5.32E-04	–	–	0.72	1.27E-02	–	–	0.69	2.11E-05	2496	1.08	3.64E-01	1870	0.88	4.24E-02	4366
2	rs768072	160,233,383	*BAZ2B*	T	0.25	0.72	7.66E-04	–	–	0.72	1.08E-02	–	–	0.72	2.41E-05	2596	0.96	5.72E-01	1922	0.82	2.34E-04	4518
2	rs284544	217,309,111	*SMARCAL1*	A	0.21	1.51	1.65E-04	–	–	1.36	1.77E-02	–	–	1.44	1.06E-05	2572	0.97	7.25E-01	1907	1.15	1.01E-02	4479
2	rs284541	217,368,839	*RPL37A*	T	0.20	1.51	2.56E-04	–	–	1.34	2.90E-02	–	–	1.43	2.66E-05	2531	0.92	3.28E-01	1879	1.12	4.59E-02	4410
2	rs2012243	217,412,087	*LOC101928180*	A	0.21	1.50	1.77E-04	–	–	1.30	3.98E-02	–	–	1.41	2.77E-05	2599	0.93	3.59E-01	1924	1.11	6.45E-02	4523
6	rs10947646	36,881,535	*C6orf89*	G	0.02	0.33	2.66E-04	–	–	–	–	0.35	4.97E-02	0.34	3.46E-05	3196	1.48	9.18E-02	2353	0.78	1.28E-01	5549
8	rs2283113	17,880,243	*PCM1*	A	0.47	1.32	1.42E-03	–	–	1.31	1.15E-02	–	–	1.31	4.77E-05	2524	1.07	3.10E-01	1869	1.18	3.61E-04	4393
9	rs16923249	5,592,145	*PDCD1LG2-RIC1*	A	0.03	0.44	4.26E-04	–	–	–	–	0.39	2.89E-02	0.43	3.46E-05	3214	1.02	9.29E-01	2360	0.73	1.28E-02	5574
13	rs9568558	51,810,953	*FAM124A*	G	0.26	1.33	4.77E-03	–	–	1.39	1.11E-02	1.24	1.15E-01	1.32	4.98E-05	3436	0.99	8.60E-01	2466	1.13	9.69E-03	5902
13	rs9516416	95,103,694	*DCT*	C	0.13	0.68	2.96E-03	0.68	3.71E-01	–	–	0.55	2.16E-03	0.64	2.04E-05	3166	1.00	9.64E-01	2325	0.82	4.17E-03	5491
13	rs6492111	109,055,846	*TNFSF13B-MYO16*	C	0.02	0.35	3.83E-03	0.96	9.66E-01	–	–	0.31	1.74E-03	0.35	3.47E-05	3176	0.75	2.91E-01	2325	0.54	7.51E-04	5501
16	rs4072617	20,178,590	*GPR139*	A	0.19	1.36	3.83E-03	1.41	4.70E-01	1.36	3.02E-02	1.54	2.91E-03	1.41	2.84E-06	3954	0.97	6.36E-01	2857	1.15	4.09E-03	6811
19	rs8107274	37,285,393	*ZNF790*	C	0.03	4.56	5.27E-04	–	–	2.26	1.34E-02	–	–	2.91	4.92E-05	2579	0.85	4.67E-01	1908	1.40	2.94E-02	4487
21	rs2827976	24,600,783	*LOC105372747*	G	0.19	0.69	5.47E-04	0.61	1.50E-01	–	–	0.72	3.74E-02	0.70	1.99E-05	3205	0.97	6.87E-01	2357	0.83	8.28E-04	5562

Information for the SNP ID and chromosomal position is based on NCBI genome build 37/hg19

The '-' sign indicates data not available

CHR chromosome, *BP* Physical position (base-pair), *EA* effect allele, *EAF* effect allele frequency, *OR* Odds Ratio, N sample size of meta-analysis combining cases and controls

Adapted from Lind, 2019.

TABLE 2. GENOME-WIDE ASSOCIATION STUDY OF METABOLIC SYNDROME (BINARY) DISCLOSED 93 INDEPENDENT LOCI WITH $P<5\times10^{-8}$

SNP	*Chr*	*Position*	*Effect allele*	*Other allele*	*EAF*	*Beta*	*SE*	P	*Nearest gene*	*Previous MetS trait*
rs964184	11	116648917	G	C	0.132408	0.252628	0.0091752	6.90e-167	ZPR1	HDL, TG
rs247617	16	56990716	A	C	0.324218	−0.196432	0.0071461	2.43e-166	CETP	HDL, TG
rs3844510	8	19861361	C	A	0.26399	−0.191134	0.0076266	1.30e-138	LPL	HDL, TG
rs72836561	17	41926126	T	C	0.0322648	0.23953	0.0174139	4.74e-43	CD300LG*	HDL
rs2980888	8	126507308	T	C	0.299847	0.0959522	0.0070574	4.23e-42	TRIB1*	TG
rs116843064	19	8429323	A	G	0.0191837	−0.337476	0.0263335	1.34e-37	ANGPTL4*	HDL,TG
rs2281721	1	230297136	C	T	0.38621	0.0820747	0.0066785	1.03e-34	GALNT2	HDL,TG, fat mass
rs483082	19	45416178	T	G	0.238624	0.0899299	0.0075775	1.73e-32	APOC1	HDL, TG, WC, diabetes
rs673548	2	21237544	A	G	0.204049	−0.0977899	0.0082417	1.79e-32	APOB	HDL, TG
rs1535	11	61597972	G	A	0.347774	0.0792125	0.0068083	2.74e-31	FADS2*	HDL, TG, Glucose
rs261290	15	58678720	T	C	0.347343	−0.0779563	0.0069174	1.86e-29	LIPC	HDL, TG
rs56094641	16	53806453	G	A	0.402905	0.0736765	0.0066394	1.30e-28	FTO	WC
rs2138161	2	227095159	T	C	0.351852	−0.0752413	0.0068822	8.03e-28	NEU2*	Diabetes, hypertension, fat mass
rs998584	6	43757896	A	C	0.481524	0.0706852	0.0065509	3.83e-27	VEGFA*	Fat mass
rs28597716	8	19936687	G	A	0.189687	−0.152184	0.0086723	1.04e-24	SLC18A1*	Diabetes
rs12056034	7	72878645	G	A	0.124753	−0.0968365	0.0100916	8.33e-22	BAZ1B*	HDL, TG
rs66922415	18	57848651	G	A	0.234519	0.0678461	0.0076254	5.72e-19	MC4R*	WC
rs2925979	16	81534790	T	C	0.301767	0.0616096	0.007063	2.71e-18	CMIP*	Diabetes, hypertension, HDL
rs7124681	11	47529947	A	C	0.408583	0.055786	0.0066152	3.37e-17	CELF1*	WC, HDL, hypertension,
rs11206374	1	40048009	A	G	0.225766	0.065077	0.0077364	4.04e-17	PABPC4*	WC, HDL, TG, diabetes
NA	20	44553722	T	TA	0.187918	0.0694639	0.0082966	5.64e-17	PCIF1*	HDL, TG
rs61789601	3	135954979	T	C	0.202909	−0.0683717	0.008209	8.16e-17	PCCB*	WC, HDL, TG
rs11429307	5	55857025	GT	G	0.190881	0.0669661	0.0082284	4.01e-16	C5orf67*	WC, HDL, TG, diabetes, hypertension
rs11754773	6	34577257	G	A	0.0942279	0.089294	0.0110053	4.91e-16	C6orf106*	WC
rs1260326	2	27730940	T	C	0.392938	0.0536971	0.0066555	7.14e-16	GCKR	TG, glucose
rs139974673	15	44027885	C	T	0.0251253	0.160596	0.0200473	1.14e-15	CATSPER2P1*	HDL, TG, BMI
rs9987289	8	9183358	A	G	0.0919748	0.0873883	0.0110761	3.03e-15	LOC157273	HDL, glucose
rs1800961	20	43042364	T	C	0.0314816	0.138848	0.0180355	1.38e-14	HNF4A*	HDL, diabetes
rs10822155	10	65071215	A	C	0.41581	−0.0511185	0.0066509	1.52e-14	JMJD1C*	TG, blood pressure
rs12599637	16	69593355	C	T	0.403682	−0.0509657	0.0067725	5.26e-14	MIR1538/NFAT5*	—
rs638714	1	62906489	T	G	0.347785	−0.0514954	0.0069156	9.60e-14	USP1*	TG
rs632057	6	139834012	T	G	0.372799	0.0486791	0.0067304	4.73e-13	LINC01625*	TG, HDL, HbA1c, trunk fat
rs10260148	7	130430969	T	C	0.27901	0.0521287	0.0072552	6.72e-13	KLF14*	0
rs114165349	1	27021913	C	G	0.0235174	0.147253	0.0208744	1.74e-12	ARID1A*	TG, HDL, hypertension, trunk fat
rs10187501	2	165532454	G	A	0.346097	−0.0465591	0.0069069	1.57e-11	COBLL1*	TG, HDL, diabetes, WC
rs1532127	19	47571938	G	A	0.314222	−0.0472097	0.0070678	2.40e-11	ZC3H4*	Diabetes, WC
rs7239575	18	21120035	C	T	0.49327	−0.043234	0.0065326	3.64e-11	NPC1*	WC
rs10049088	3	156797648	T	C	0.386891	−0.0443217	0.0067433	4.94e-11	LINC00880*	WC
rs1009360	2	65276049	C	T	0.415241	−0.0435887	0.0066372	5.12e-11	CEP68*	TG, blood pressure
rs7660883	4	87982876	G	C	0.377013	−0.0443878	0.0067925	6.37e-11	AFF1*	TG, HDL
rs11075253	16	15148646	A	C	0.297121	−0.0468788	0.0071929	7.15e-11	PDXDC1*	WC, TG
rs1534696	7	26397239	C	A	0.460436	0.0424362	0.0065434	8.85e-11	SNX10*	TG, HDL, diabetes, WC, hypertension

(*coniinued*)

TABLE 2. Continued

SNP	*Chr*	*Position*	*Effect allele*	*Other allele*	*EAF*	*Beta*	*SE*	P	*Nearest gene*	*Previous MetS trait*
rs7563362	2	620297	A	G	0.142333	−0.0613566	0.0095051	1.08e-10	TMEM18*	WC
rs9378248	6	31326289	A	G	0.339668	0.0440684	0.0068853	1.55e-10	HLA-B	Diabetes, trunk fat mass
rs563296	10	99772404	G	A	0.44029	−0.0419399	0.0065928	2.00e-10	CRTAC1*	WC, diabetes
rs1023193	15	41855736	T	G	0.305574	−0.045397	0.0071595	2.29e-10	TYRO3*	BMI, blood pressure
rs13107325	4	103188709	T	C	0.0748556	0.0760904	0.0121877	4.29e-10	SLC39A8*	HDL.WC, blood pressure
NA	12	124503803	CAA	C	0.436617	−0.0414889	0.0066476	4.34e-10	ZNF664*	HDL.WC,TG
rs779518917	17	76398124	GTGT	G	0.403203	0.0415352	0.0066831	5.13e-10	PGS1*	HDL, blood pressure
rs35661464	11	64828842	T	C	0.252033	0.0463799	0.0074696	5.33e-10	NAALADL1*	0
rs764311894	20	51204733	CTTT	C	0.169241	−0.0569809	0.0091796	5.39e-10	LINC01524*	—
rs748326686	4	67851522	CT	C	0.37669	0.042439	0.0068632	6.27e-10	LOC101927237*	—
rs577721086	6	127440047	C	T	0.0509599	0.0900937	0.0147076	9.03e-10	RSPO3*	HDL.WC,TG
rs537069642	3	50119387	CTT	C	0.423828	−0.0418519	0.0068962	1.29e-09	RBM6*	HDL, diabetes, WC, hypertension
rs76376137	6	34173330	G	T	0.050549	0.0820329	0.0146096	1.30e-09	HMGA1*	WC
rs10913469	1	177913519	C	T	0.207314	0.0482641	0.0079955	1.58e-09	SEC16B*	WC
rs56133711	11	27723334	A	G	0.262839	0.0443302	0.0073785	1.88e-09	BDNF*	WC
rs3814883	16	29994922	T	C	0.48348	0.0392918	0.0065507	2.00e-09	TAOK2*	WC, hypertension
rs56282717	7	150657095	A	G	0.244425	−0.0460964	0.0076943	2.09e-09	KCNH2	WC, hypertension
rs12945575	17	40713071	T	C	0.25004	0.044779	0.0075111	2.50e-09	COASY*	WC
rs2306363	11	65405600	T	G	0.206589	−0.0482761	0.0081366	2.97e-09	SIPA1*	Hypertension
rs781006834	7	17941865	G	GT	0.344659	0.0411062	0.0069362	3.10e-09	SNX13*	HDL
rs11751347	6	161092438	T	C	0.102714	0.0624875	0.0106566	4.53e-09	LPA*	WC, TG
rs73123462	4	36077604	T	C	0.0155967	0.149549	0.0255208	4.63e-09	ARAP2*	Diabetes
rs62107261	2	422144	C	T	0.0482567	−0.089776	0.0156627	5.05e-09	LINC01874*	WC
rs3808439	8	116563675	A	G	0.447923	0.0383457	0.0065731	5.42e-09	TRPS1*	WC, HDL
rs56959712	12	123188475	T	G	0.210819	−0.0470487	0.0080746	5.65e-09	HCAR2*	HDL, fat mass
rs56001710	7	25983400	A	T	0.417154	0.0388139	0.0068186	6.85e-09	MIR148A*	—
rs779470261	3	131646163	C	CTCTAA	0.26329	0.0431103	0.0074441	6.99e-09	CPNE4*	WC
rs5789783	11	13347748	T	TA	0.402641	−0.0399511	0.0069148	7.57e-09	ARNTL*	WC, hypertension
rs3949781	9	92178472	T	A	0.461633	−0.0380199	0.0065968	8.25e-09	GADD45G*	0
rs11871285	17	65840809	T	G	0.193205	0.0473409	0.0082208	8.48e-09	BPTF*	WC
rs6545703	2	59032635	G	T	0.415927	−0.0382425	0.0066477	8.78e-09	LINC01122*	WC
rs58175144	1	150839698	CA	C	0.301172	0.0413753	0.0072197	9.99e-09	ARNT*	HDL, trunk fat mass
rs7188873	16	24727064	A	G	0.376011	−0.03863	0.0067729	1.17e-08	TNRC6A*	WC, blood pressure
rs5021727	6	32578633	G	A	0.461566	0.0371336	0.0065596	1.51e-08	HLA-DRB1*	Diabetes
rs1105654	18	47147746	G	A	0.345091	0.038799	0.0068934	1.82e-08	LIPG*	HDL
rs12752223	1	93837133	T	C	0.396475	−0.0374831	0.0066755	1.97e-08	DR1*	HDL
rs140628616	11	56201709	A	AAT	0.16585	−0.0528712	0.0094553	2.25e-08	OR5R1*	0
rs9971210	10	21879084	G	C	0.489597	0.0365804	0.0065457	2.29e-08	MLLT10*	WC, blood pressure
rs10954772	8	30863938	T	C	0.313991	0.0391984	0.0070493	2.69e-08	PURG*	WC, hypertension
rs12472667	2	171629063	G	C	0.371305	0.0374895	0.0067634	2.97e-08	ERICH2*	WC
rs1143015	17	7485206	A	G	0.157971	0.0491401	0.0088808	3.14e-08	CD68*	Weight
rs11789603	9	107647019	T	C	0.107873	−0.058891	0.0106912	3.62e-08	ABCA1	HDL

(*continued*)

TABLE 2. Continued

SNP	*Chr*	*Position*	*Effect allele*	*Other allele*	*EAF*	*Beta*	*SE*	P	*Nearest gene*	*Previous MetS trait*
rs9332817	11	118365210	C	G	0.0265006	−0.115795	0.0210291	3.66e-08	KMT2A*	WC
rs535286942	10	94797201	AT	A	0.374794	−0.037682	0.0068466	3.72e-08	EXOC6*	Diabetes
rs8121509	20	62712053	C	T	0.451827	−0.0360691	0.0065647	3.92e-08	OPRL1*	WC, hypertension
rs143321598	3	39116681	CT	C	0.488259	−0.0382859	0.0069834	4.20e-08	WDR48*	0
rs10945840	6	162974904	C	G	0.332614	−0.0382931	0.0069905	4.30e-08	PARK2*	—
rs7138803	12	50247468	A	G	0.368675	0.0369574	0.0067465	4.30e-08	BCDIN3D*	WC
rs11655056	17	47364107	C	T	0.46087	0.0359929	0.0065729	4.35e-08	MIR6129*	0
rs4921913	8	18272377	C	T	0.221605	0.0426441	0.0077967	4.51e-08	NAT2*	TG
rs2307111	5	75003678	C	T	0.393101	−0.0366715	0.006708	4.58e-08	POC5*	WC

The loci given with an *asterisk* after the name of the nearest gene are novel findings for MetS as binary trait. The beta represents the log odds. EAF, effect allele frequency. Build37, hg19 is used for chromosomal positions. Previously published associations with MetS components are given in the right-hand-side column using data from Phenoscanner. Details on those findings are given in Supplementary Table S2—denotes no associations versus any trait. Zero denotes no association with a MetS trait, but associations with other traits.

Adapted from Oh, et al., 2020.

Table 2. Significant variants associated with metabolic syndrome.

Chr	SNP|	Position	Gene	M	Discovery set			Replication set		
					MAF (case / control)	OR (95% CI)	P	MAF (case / control)	OR (95% CI)	P
11	rs662799	116663707	APOA5	G	0.345 / 0.288	1.346 (1.227–1.476)	2.85×10^{-10}	0.334 / 0.290	1.268 (1.083–1.485)	3.19×10^{-3}

Chr, chromosome; rs number, SNP ID in dbSNP database; M, minor allele; MAF, minor allele frequency; OR, odds ratio; CI, confidence interval, respectively

Adapted from Willems, et al., 2020.

Supplemental Table D.1: Summary of suggestive and genome-wide significant SNPs

Chr	Base Pair	rsID	Phenotype	FE P-value	RE2 P-value	TransMeta P-value	MR-MEGA P-value	Cochran's Q P-value	Functional Info	Nearest gene
2	27598097	rs4665972	Triglycerides	4.82E-07	9.04E-08	5.66E-07	1.72E-06	0.0106	intronic	SNX17
2	27730940	rs1260326	Triglycerides	3.39E-07	4.86E-08	2.91E-07	9.83E-07	0.0070	exonic	GCKR
2	27741237	rs780094	Triglycerides	6.60E-07	3.87E-07	5.87E-07	2.98E-06	0.0498	intronic	GCKR
2	65819883	rs1115848	Systolic BP	6.03E-07	8.84E-07	5.78E-07	1.45E-06	0.4796	intergenic	SPRED2;MIR4778
2	65820608	rs11687213	Systolic BP	6.03E-07	8.84E-07	5.78E-07	1.45E-06	0.4796	intergenic	SPRED2;MIR4778
2	65824325	rs12614551	Systolic BP	7.88E-07	1.16E-06	6.33E-07	2.31E-06	0.5475	intergenic	SPRED2;MIR4778
2	65824380	rs12614575	Systolic BP	7.88E-07	1.16E-06	6.33E-07	2.31E-06	0.5475	intergenic	SPRED2;MIR4778
2	161536779	rs4504007	Weight	2.76E-06	3.16E-06	3.01E-07	1.49E-05	0.0223	intergenic	RBMS1;TANK
2	161580892	rs113055309	Weight	1.01E-06	1.43E-06	2.93E-07	4.60E-06	0.0439	intergenic	RBMS1;TANK
2	161650240	rs35013036	Weight	9.79E-07	1.44E-06	3.08E-07	5.71E-06	0.1876	intergenic	RBMS1;TANK
2	161743385	rs1615586	Weight	3.04E-06	4.44E-06	5.72E-07	1.18E-05	0.0807	intergenic	RBMS1;TANK
2	161746908	rs1404359	Weight	3.77E-06	4.35E-06	5.68E-07	9.82E-06	0.0509	intergenic	RBMS1;TANK
2	161755027	rs1710654	Weight	2.95E-06	3.66E-06	2.95E-07	7.05E-06	0.0717	intergenic	RBMS1;TANK
3	2001175	rs12631510	HDL	6.40E-06	7.26E-07	2.80E-07	3.06E-07	0.0021	intergenic	CNTN6;CNTN4
3	2004251	rs17005939	HDL	6.98E-06	5.23E-07	2.80E-07	1.86E-07	0.0011	intergenic	CNTN6;CNTN4
3	105643849	rs6765145	Waist	5.46E-07	7.98E-07	5.75E-07	2.41E-06	0.7921	intergenic	CBLB;LINC00882
4	2707032	rs4690015	Diastolic BP	0.00227145	9.34E-07	0.00245669	0.0013334	1.10E-05	intronic	FAM193A
5	157023304	rs1895338	Triglycerides	0.0204752	5.82E-07	0.014237215	0.00017668	5.22E-07	intergenic	ADAM19;SOX30
8	126485294	rs2954027	Triglycerides	1.63E-06	2.38E-06	2.97E-07	1.64E-06	0.2682	intergenic	TRIB1;LINC00861
8	126486409	rs17321515	Triglycerides	3.18E-06	4.64E-06	6.30E-07	3.53E-06	0.2725	intergenic	TRIB1;LINC00861
8	126488235	rs2980868	Triglycerides	1.54E-06	2.26E-06	5.91E-07	2.83E-06	0.4005	intergenic	TRIB1;LINC00861
8	126488250	rs2980869	Triglycerides	1.54E-06	2.26E-06	5.91E-07	2.83E-06	0.4005	intergenic	TRIB1;LINC00861
8	126491733	rs2954031	Triglycerides	1.13E-06	1.56E-06	1.36E-07	7.09E-07	0.1824	intergenic	TRIB1;LINC00861
8	126495818	rs10808546	Triglycerides	2.06E-06	2.97E-06	2.94E-07	2.47E-06	0.2286	intergenic	TRIB1;LINC00861
9	114376753	rs2418173	Diastolic BP	3.12E-07	4.54E-07	2.92E-07	1.87E-06	0.8098	upstream	LRRC37A5P
9	114377336	rs10817195	Diastolic BP	9.38E-07	1.38E-06	1.10E-06	5.33E-06	0.8902	intergenic	LRRC37A5P; DNAJC25-GNG10
11	89224718	rs2289123	Triglycerides	0.0127646	1.67E-05	0.006225193	7.62E-07	1.50E-05	UTR5	NOX4
13	66731477	rs9599076	Waist	4.80E-08	7.10E-08	6.47E-08	3.29E-07	0.8503	intergenic	MIR548X2; MIR4704

Chr	Base Pair	rsID	Phenotype	FE P-value	RE2 P-value	TransMeta P-value	MR-MEGA P-value	Cochran's Q P-value	Functional Info	Nearest gene
13	66732565	rs9592449	Waist	9.95E-07	1.46E-06	1.11E-06	6.29E-06	0.8952	intergenic	MIR548X2; MIR4704
13	90409398	rs317962	Triglycerides	6.99E-07	1.02E-06	6.03E-07	1.10E-06	0.3117	intergenic	LINC00353; LINC00559
15	29964742	rs4522365	Triglycerides	2.58E-06	7.13E-08	1.65E-07	1.20E-06	0.0013	intergenic	FAM189A1; LOC100130111
15	64276143	rs8038345	Triglycerides	4.46E-07	6.51E-07	2.89E-07	2.71E-06	0.4122	intronic	DAPK2
15	64284719	rs28478668	Triglycerides	4.83E-07	7.08E-07	5.68E-07	2.96E-06	0.5707	intronic	DAPK2
15	64285189	rs11633956	Triglycerides	3.64E-07	5.30E-07	2.87E-07	2.21E-06	0.6302	intronic	DAPK2
15	64285659	rs34867794	Triglycerides	3.61E-07	5.26E-07	2.87E-07	2.19E-06	0.6361	intronic	DAPK2
15	64286221	rs28544905	Triglycerides	3.61E-07	5.26E-07	2.87E-07	2.19E-06	0.6361	intronic	DAPK2
15	64286236	rs28459332	Triglycerides	3.61E-07	5.26E-07	2.87E-07	2.19E-06	0.6361	intronic	DAPK2
15	64286836	rs11631973	Triglycerides	3.61E-07	5.26E-07	2.87E-07	2.19E-06	0.6361	intronic	DAPK2
15	64287495	rs28444644	Triglycerides	5.22E-07	7.64E-07	5.71E-07	3.11E-06	0.6526	intronic	DAPK2
15	64290136	rs7167478	Triglycerides	1.23E-06	1.70E-06	3.16E-07	2.79E-06	0.2719	intronic	DAPK2
15	64290385	rs55963180	Triglycerides	3.53E-07	5.15E-07	1.47E-07	1.48E-06	0.4344	intronic	DAPK2
15	64291219	rs8024045	Triglycerides	3.61E-07	5.26E-07	2.87E-07	2.19E-06	0.6361	intronic	DAPK2
15	64297369	rs11633496	Triglycerides	6.19E-07	9.08E-07	5.75E-07	3.44E-06	0.6399	intronic	DAPK2
15	64297435	rs11633611	Triglycerides	6.19E-07	9.08E-07	5.75E-07	3.44E-06	0.6399	intronic	DAPK2
15	64313764	rs11635284	Triglycerides	1.00E-06	1.47E-06	5.92E-07	5.35E-06	0.4791	intronic	DAPK2
15	64333606	rs7173139	Triglycerides	6.69E-07	9.80E-07	3.07E-07	3.10E-06	0.4899	intronic	DAPK2
15	64334978	rs881232	Triglycerides	5.67E-07	8.27E-07	5.68E-07	3.09E-06	0.5751	intronic	DAPK2
15	64334992	rs968654	Triglycerides	3.84E-07	5.58E-07	2.83E-07	2.08E-06	0.5306	intronic	DAPK2
15	64335225	rs1868444	Triglycerides	5.31E-07	7.76E-07	3.03E-07	2.90E-06	0.5876	intronic	DAPK2
15	64335240	rs1868443	Triglycerides	3.84E-07	5.58E-07	2.83E-07	2.08E-06	0.5306	intronic	DAPK2
16	56987015	rs12446515	HDL	2.44E-07	3.56E-07	1.43E-07	9.03E-07	0.6364	intergenic	HERPUD1;CETP
16	56987369	rs56156922	HDL	1.73E-07	2.51E-07	1.37E-07	6.35E-07	0.6118	intergenic	HERPUD1;CETP
16	56987765	rs56228609	HDL	3.04E-07	4.42E-07	2.80E-07	1.00E-06	0.6451	intergenic	HERPUD1;CETP
16	56988044	rs173539	HDL	7.04E-06	5.39E-06	5.71E-07	8.32E-07	0.0357	intergenic	HERPUD1;CETP
16	56989590	rs247616	HDL	5.54E-08	8.17E-08	6.27E-08	3.82E-07	0.7879	intergenic	HERPUD1;CETP
16	56990716	rs247617	HDL	5.54E-08	8.17E-08	6.27E-08	3.82E-07	0.7879	intergenic	HERPUD1;CETP
16	56991363	rs183130	HDL	5.54E-08	8.17E-08	6.27E-08	3.82E-07	0.7879	intergenic	HERPUD1;CETP

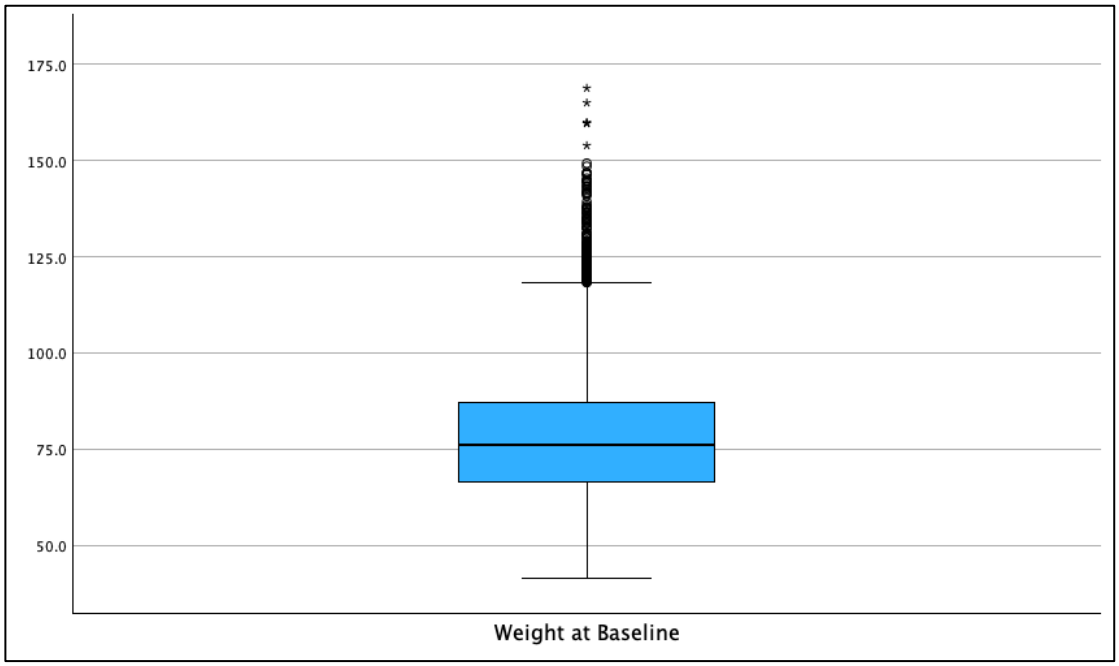
Chr	Base Pair	rsID	Phenotype	FE P-value	RE2 P-value	TransMeta P-value	MR-MEGA P-value	Cochran's Q P-value	Functional Info	Nearest gene
16	56993161	rs12149545	HDL	5.51E-07	8.05E-07	5.65E-07	1.62E-06	0.5903	intergenic	HERPUD1;CETP
16	56993324	rs3764261	HDL	5.73E-06	6.84E-06	6.21E-07	1.57E-06	0.0738	intergenic	HERPUD1;CETP
16	56994528	rs17231506	HDL	3.13E-07	4.55E-07	2.81E-07	1.18E-06	0.5802	intergenic	HERPUD1;CETP
16	56998918	rs12720926	HDL	1.66E-08	2.33E-08	1.37E-08	1.07E-07	0.7494	intronic	CETP
16	56999328	rs11508026	HDL	1.79E-08	2.56E-08	1.39E-08	1.14E-07	0.7455	intronic	CETP
16	57001216	rs4784741	HDL	9.74E-08	1.44E-07	6.62E-08	2.46E-07	0.4294	intronic	CETP
16	57001438	rs12444012	HDL	9.74E-08	1.44E-07	6.62E-08	2.46E-07	0.4294	intronic	CETP
16	57004889	rs7205804	HDL	4.85E-08	7.16E-08	2.90E-08	7.41E-08	0.2832	intronic	CETP
16	57005301	rs1532625	HDL	8.22E-08	1.21E-07	6.23E-08	8.58E-08	0.2118	intronic	CETP
16	57005479	rs1532624	HDL	8.22E-08	1.21E-07	6.23E-08	8.58E-08	0.2118	intronic	CETP
18	74352797	rs9951751	Systolic BP	0.00143975	2.39E-06	3.21E-06	1.17E-07	2.65E-05	intergenic	LINC01927; LINC01879
19	38039675	rs11665759	Triglycerides	0.0640212	0.00018267	0.071594838	8.29E-07	1.71E-05	upstream	ZNF571-AS1
19	38040879	rs73031322	Triglycerides	0.0478103	0.00016422	0.053617891	9.76E-07	2.60E-05	ncRNA intronic	ZNF571-AS1
19	38043022	rs73031326	Triglycerides	0.0478103	0.00016422	0.053617891	9.76E-07	2.60E-05	ncRNA intronic	ZNF571-AS1
19	38046331	rs111694872	Triglycerides	0.0478103	0.00016422	0.053617891	9.76E-07	2.60E-05	ncRNA intronic	ZNF571-AS1
19	38062195	rs2045911	Triglycerides	0.0478103	0.00016422	0.053617891	9.76E-07	2.60E-05	ncRNA intronic	ZNF571-AS1
19	38073146	rs73033117	Triglycerides	0.0640212	0.00018267	0.071594838	8.29E-07	1.71E-05	ncRNA intronic	ZNF571-AS1
19	38074152	rs11083427	Triglycerides	0.0478103	0.00016422	0.053617891	9.76E-07	2.60E-05	ncRNA intronic	ZNF571-AS1
19	38080535	rs73033129	Triglycerides	0.0478103	0.00016422	0.053617891	9.76E-07	2.60E-05	intronic	ZNF540;ZNF571
19	38082385	rs12162238	Triglycerides	0.0478103	0.00016422	0.053617891	9.76E-07	2.60E-05	intronic	ZNF540;ZNF571
19	38083967	rs11083428	Triglycerides	0.0478103	0.00016422	0.053617891	9.76E-07	2.60E-05	intronic	ZNF540;ZNF571

6. Normality Tests for Continuous Variables

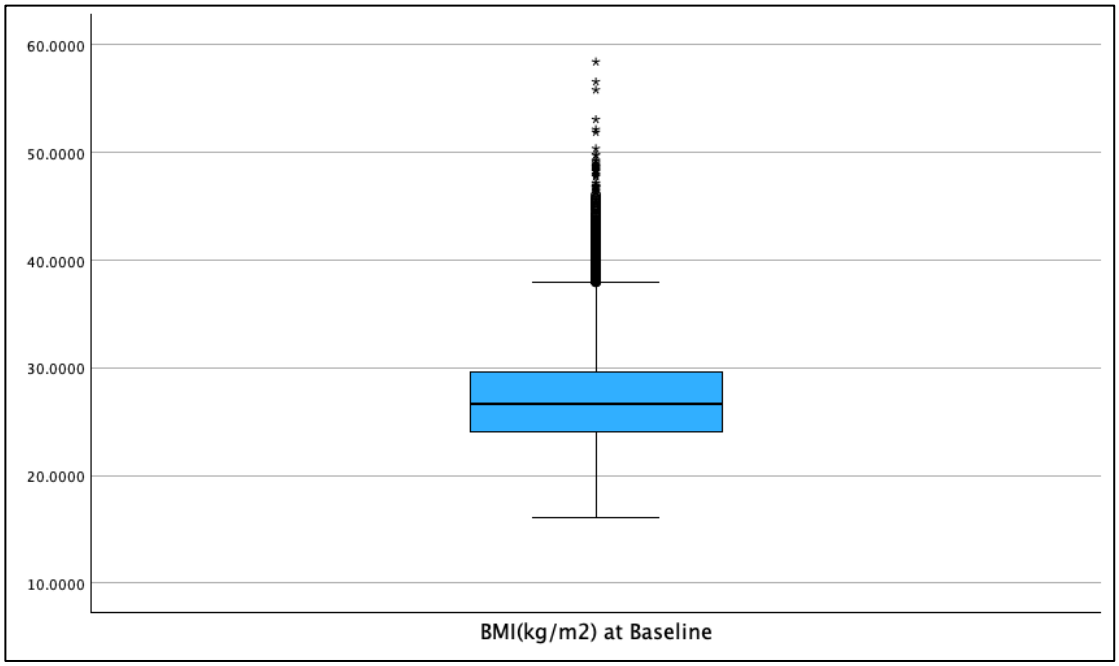
Boxplots presented for primary continuous variables in the study, as derived from SPSS version 29; notes at the end of Section 6.

6A. Age at Recruitment Boxplot

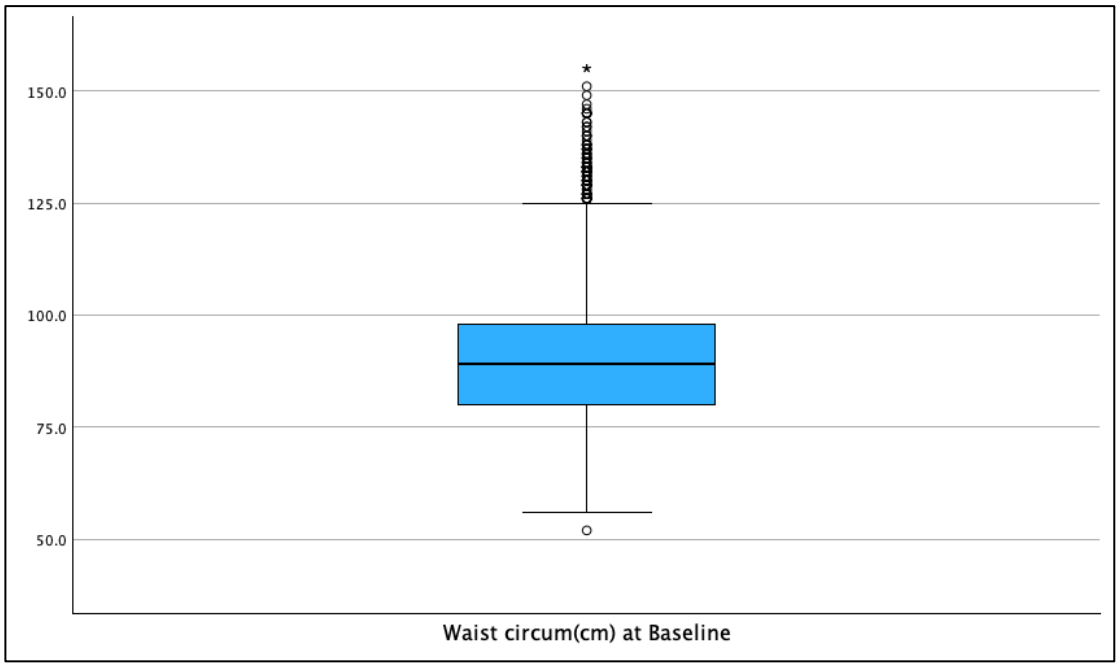
6B. Weight at Baseline Boxplot



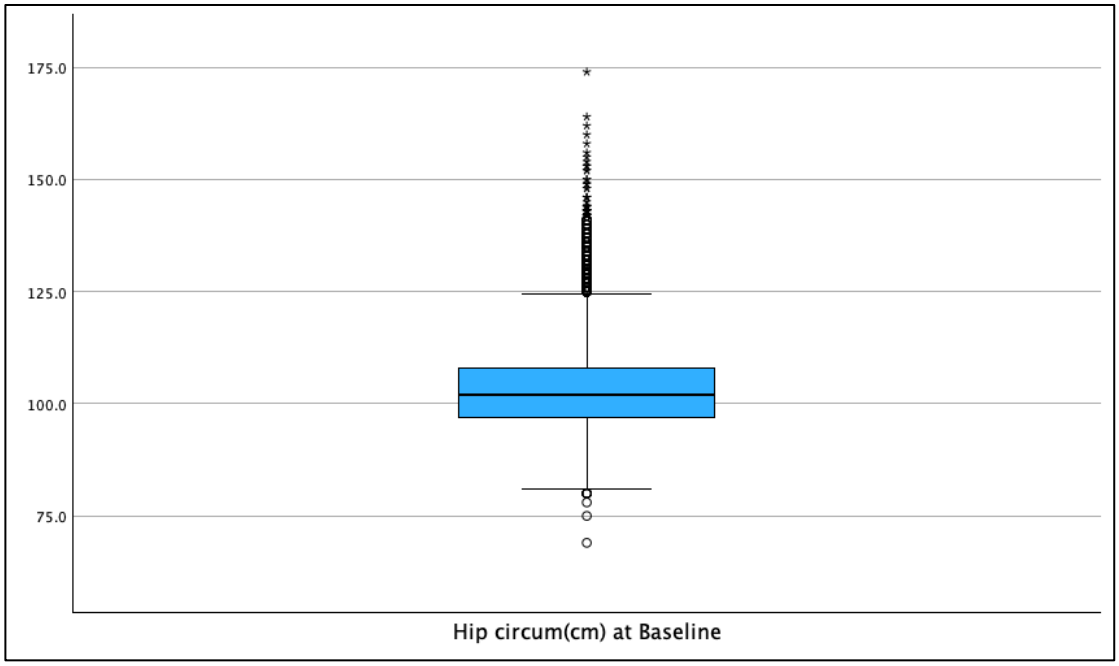
6C. BMI Boxplot



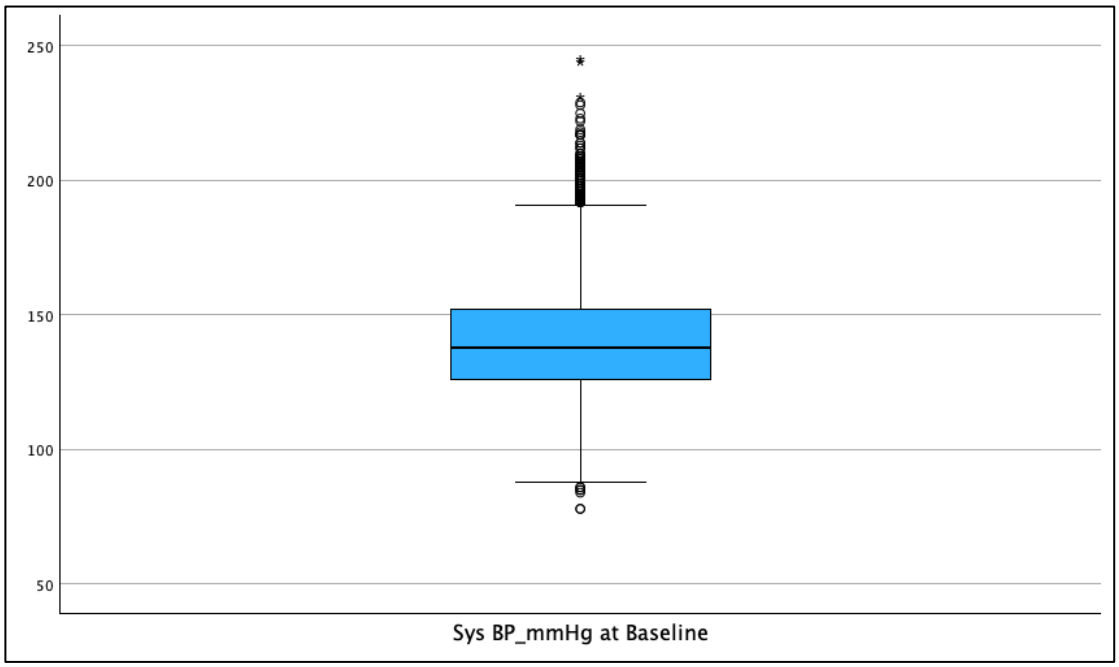
6D. Waist circumference Boxplot



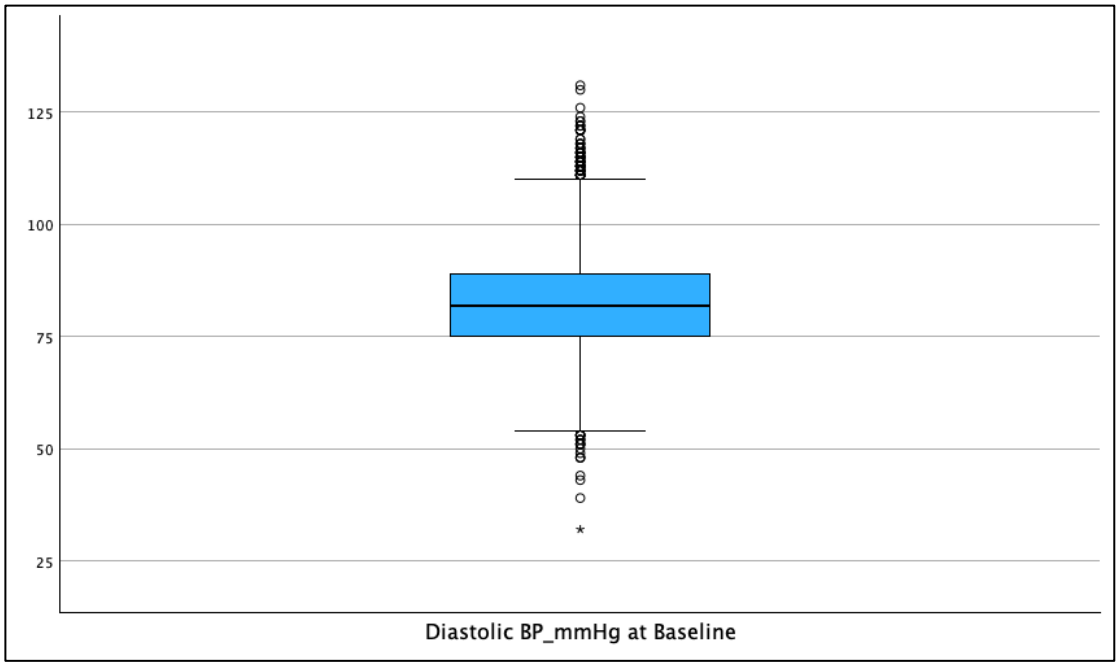
6E. Hip circumference Boxplot



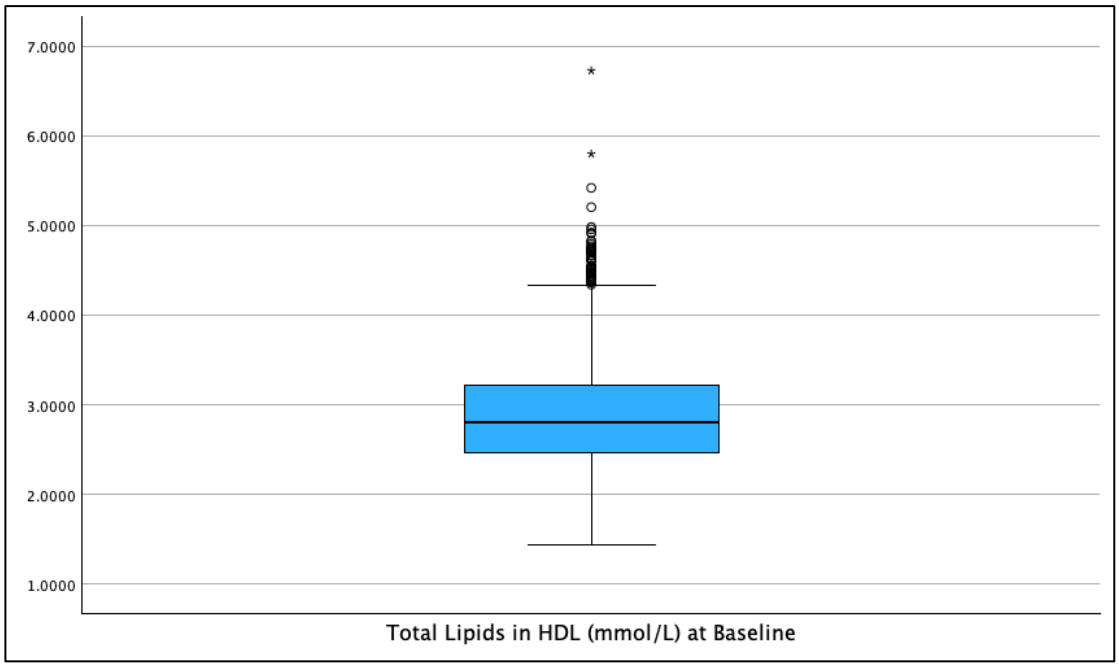
6F. Systolic BP Boxplot



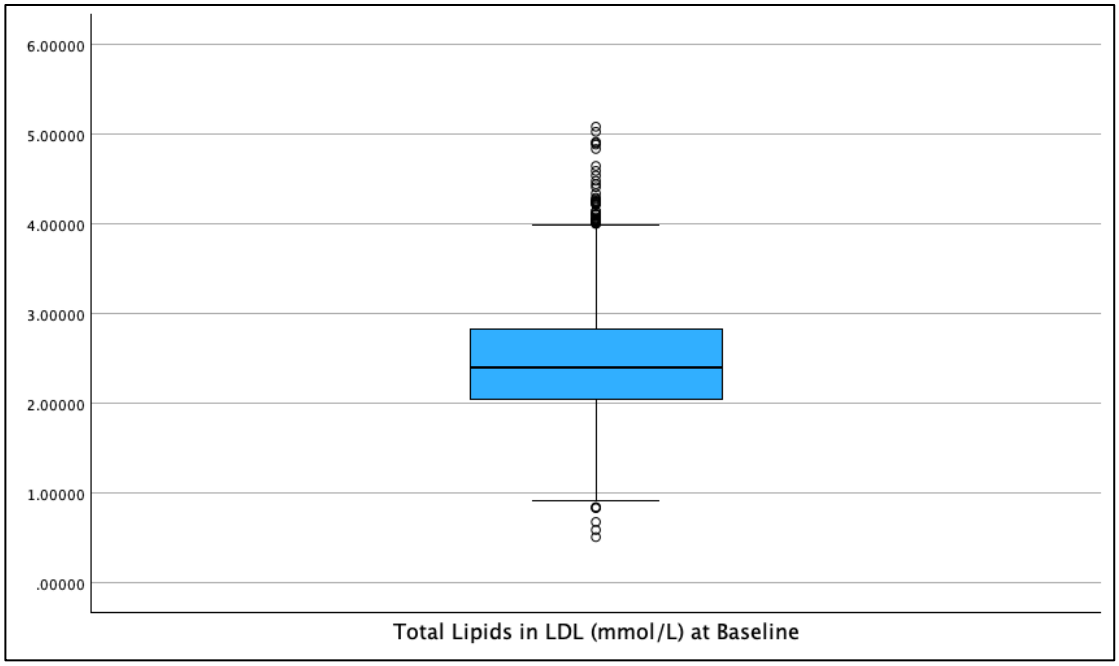
6G. Diastolic BP Boxplot



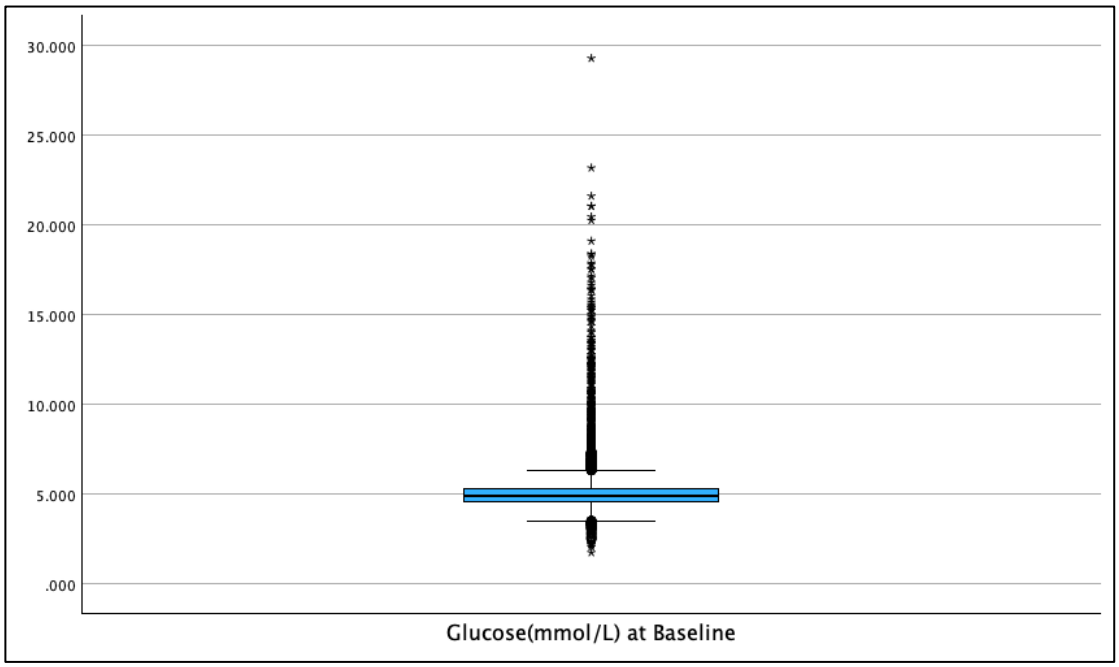
6H. Total Lipids in HDL Boxplot



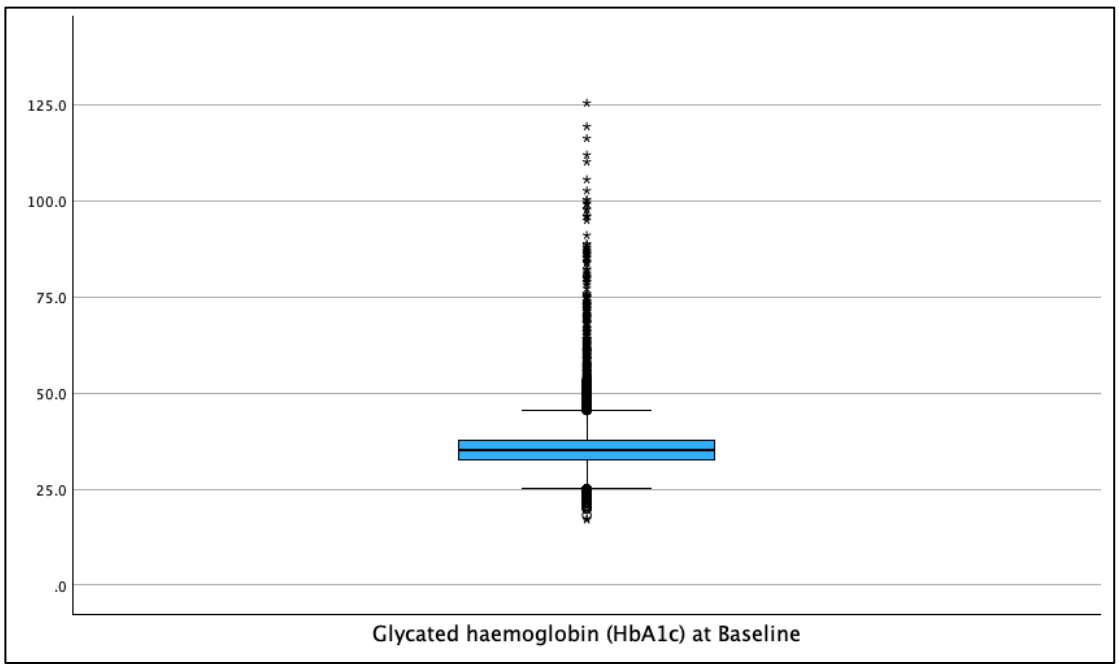
6I. Total Lipids in LDL Boxplot



6J. Glucose Boxplot



6K. HbA1c Boxplot



Notes:

In each plot (6A to 6K), the central line inside the box indicates the median, the lower and upper edges of the box represent the 25th and 75th percentiles (interquartile range, IQR), and the whiskers extend to the most extreme values within 1.5 × IQR from the box. Data points beyond the whiskers (except 6A for age) were plotted individually as circles (mild outliers; >1.5 × IQR but ≤3 × IQR from the quartiles) or asterisks (extreme outliers; >3 × IQR from the quartiles). A large number of outliers were present in most plots and were included in the normally distributed data for complete data representation and insight provision.

7. UK Biobank Material Transfer Agreement

biobank uk
Enabling scientific discoveries that improve human health

Material Transfer Agreement

Dear Mrs Fojas

Applicant: Staffordshire University
Application Reference Number: 77577

UK Biobank Limited ("**UK Biobank**") is pleased to approve your Application to use the UK Biobank resource. UK Biobank's approval of this Application is valid for 90 days during which time the Applicant must execute this Material Transfer Agreement ("**MTA**") and pay the Access Charges. These are the final steps before access is granted. If these steps are not taken by the Applicant within 90 days, the Applicant will need to re-apply for access.

Parties

This is an agreement between UK Biobank and the Applicant (each a "party", together the "parties"). The Applicant PI is not a party to the MTA, however, UK Biobank requires the Applicant PI to sign the MTA to acknowledge that the provisions of this MTA have been "read and understood" so that they are fully aware of the Applicant's obligations to both UK Biobank and to UK Biobank's Participants. The Applicant shall be responsible for the conduct of the Applicant PI and any and all Applicant Researchers involved in this Approved Research Project but shall not be responsible for the conduct of any Collaborator Institution(s), Lead Collaborator(s) or Collaborator Researcher(s).

Structure of agreement

This MTA shall become effective on the Effective Date. If you have agreed a previous version of the MTA for this Application/Approved Research Project, the previous version shall automatically terminate on the Effective Date and be replaced by this MTA.

This MTA is conditional upon UK Biobank receiving from the Applicant within thirty (30) days of the Effective Date, cleared funds covering the Access Charges and any applicable VAT.

Standard terms and Annexes

The content of UK Biobank's standard MTA, and the conditions contained within it, are non-negotiable.

This MTA incorporates the attached Applicant Terms and Conditions (including any documents and/or the materials that are referred to in them), the contents of the Application Form (where applicable) and the attached Annexes:

- Annex 1 (Data Processing Description);
- Annex 2 (Security Measures);
- Annex 3 (Applicant Annual Project Report Template); and
- Annex 4 (Approved Research Project – which summarises the Materials that will be made available to the Applicant).

Definitions used in this MTA can be found on pages 15-17.

Payment

The Access Charges which are payable are set out in the payment section of your application. This allows you to generate an invoice on which VAT will be included (if appropriate, and as such VAT will not be included if the Applicant is based outside the UK). A summary of these Access Charges is also set out in Annex 4.

Payment should be submitted via bank transfer or Sage Pay, in cleared funds and in British pounds sterling (GBP) to:

Bank: Barclays Bank PLC
Account name: UK Biobank Limited
Account number: 33069427
Sort code: 20-24-41
IBAN: GB78 BARC 2024 4133 0694 27

Yours sincerely

For and on behalf of UK Biobank
Jonathan Sellors
General Counsel & Company Secretary

Applicant Terms and Conditions

1. Supply of Materials by UK Biobank

1.1 UK Biobank agrees to supply the Materials to the Applicant in the timeframe and manner set out in this MTA, subject to the provisions of this MTA.

1.2 UK Biobank warrants to the Applicant that for the purposes of this MTA:

1.2.1 it is entitled to supply the Materials to the Applicant;

1.2.2 consent to take part in UK Biobank has been obtained from the Participants and further, consent under the Human Tissue Act 2004, has been obtained from the relevant Participants; and

1.2.3 the use of the Materials for the Approved Research Project falls within UK Biobank's generic Research Tissue Bank (RTB) approval from the NHS North West REC, available here.

1.3 The Applicant agrees that the Materials are provided on an "as is" basis without any warranty of satisfactory quality or fitness for a particular purpose or use, or that use of the Materials shall not infringe the rights of any third party. Except as expressly stated in this MTA, all warranties, terms and conditions, whether express or implied by statute, common law or otherwise, are excluded to the fullest extent permitted by law.

2. Usage of Materials by the Applicant

2.1 The Applicant agrees that the Materials shall only be used:

2.1.1 in accordance with the terms and conditions of this MTA;

2.1.2 to conduct the Approved Research Project for the Permitted Purpose only; and

2.1.3 by the Applicant institution and on an individual level within the Applicant, by the Applicant PI, the Applicant Researchers and by Affiliates and Third Party Processors (appointed by the Applicant).

2.2 The Applicant shall not share, sub-license, disclose, transfer, sell, gift or supply the Materials to any other person or unauthorised third party.

2.3 Without prejudice to the other provisions of this MTA, any actual or anticipatory breach of any provision of clause 2.1 or 2.2 shall entitle UK Biobank to terminate this Agreement with immediate effect, and require the immediate return or destruction of any Materials provided by UK Biobank.

2.4 The Applicant shall and shall procure that the Applicant PI, the Applicant Researchers and any Affiliate and any Third Party Processors are made aware of, and shall comply with, the terms and conditions of this MTA and the Data Protection Legislation. Any act or omission of the Applicant PI or any Applicant Researcher or any Affiliate or any Third Party Processor shall be deemed to be an act of the relevant Applicant for which the relevant Applicant is fully responsible and liable.

2.5 This MTA confers on the Applicant only those rights that are expressly granted to the Applicant. For the avoidance of doubt, nothing in this MTA shall prevent UK Biobank from supplying the same Materials (or other data and/or samples in the UK Biobank resource) to another third party, in line with the access procedures (available on UK Biobank's website here as may be updated by UK Biobank from time to time) or for UK Biobank's other operational purposes.

2.6 In relation to the Materials supplied to the Applicant:

2.6.1 UK Biobank is the owner of the Materials, and UK Biobank is the owner of the Intellectual Property Rights in the Materials; and

2.6.2 UK Biobank hereby grants to the Applicant a revocable, worldwide, royalty-free, non-exclusive, non-transferable licence (but not any ownership rights) during the Term to use the Materials for the Permitted Purpose, subject to the terms and conditions of this MTA.

3. Generation of data by the Applicant

Generation of data by or on behalf of the Applicant during the Approved Research Project

3.1 The data generated by the Applicant in the performance of the Approved Research Project shall be deemed to fall into the following categories:

3.1.1 **Results Data**: data and methodology (for example, the SAS/R/Stata scripts) which underlie the Findings and which would enable another competent researcher to generate the Findings;

3.1.2 **Findings**: the findings generated by the Applicant as a result of the Approved Research Project; and/or

3.1.3 **Other Data**: all other data generated by the Applicant which is not in one of the above two categories.

Ownership of generated data

3.2 Except as provided in clause 3.3, the Applicant shall own the IPRs in their Findings, the Results Data and the Other Data. The Applicant hereby grants a perpetual, irrevocable, worldwide, fully paid up, royalty free, fully sub-licensable non-exclusive licence to UK Biobank to use, reproduce, distribute, publish, store and otherwise disseminate the Findings, the Results Data and the Other Data.

3.3 Nothing in this MTA shall operate to assign to the Applicant any IPRs in the Materials. To the extent that the Findings, the Results Data or the Other Data incorporate any Materials, the IPRs in those Materials shall remain the property of UK Biobank and shall not belong to the Applicant.

3.4 The Applicant warrants to UK Biobank that UK Biobank's receipt of and use of the Applicant's Findings and Results Data shall not infringe the rights, including any IPRs, of any third party.

Rights to inventions/developments made by the Applicant

3.5 Subject always to the restriction in clause 3.7, UK Biobank confirms that it shall have no rights or licence to the IPRs in relation to any inventions made by the Applicant as a result of using the Materials, Results Data, Findings or Other Data ("**Applicant-Generated Inventions**").

3.6 However, the Applicant acknowledges that the UK Biobank resource has been (a) produced using a combination of the goodwill and contribution of 500,000 UK participants (b) charitable and public funding (from in particular Wellcome and the Medical Research Council) (c) the use of public resources (such as UK health-record data) and (d) established with the express purpose of promoting the conduct of health-related research which is in the public interest. UK Biobank also acknowledges the contribution which is being made to enhance the resource by the Applicant (in the form of the generation and availability to other researchers of, inter alia, the Findings and Results Data of the Approved Research Project).

3.7 In terms of specific obligations, taking into account the acknowledgements in clause 3.6 above, the Applicant agrees (and this clause is a material provision of this MTA) that it shall not and shall not attempt to:

3.7.1 file any patents with claims directed to; or

3.7.2 otherwise seek to claim or enforce any IPRs in;

the genotype-phenotype data within the Materials or in the genotype-phenotype data which has been generated by (or on behalf of) the Applicant in the course of the Approved Research Project (whether such genotype-phenotype data is in the form of Results Data, Findings or Other Data). Without limiting the above, the parties agree that this clause 3.7 shall not prohibit the Applicant from patenting, or enforcing IPRs in drugs, therapeutics, diagnostics, other technology or methods of treatment provided this does not limit UK Biobank's ability to allow approved researchers to use the data generated by the Applicant (as defined in clause 3.1), including any biomarker data identified by the Applicant, through its use of the UK Biobank resource.

Limitation on rights granted

3.8 UK Biobank expressly excludes (directly or indirectly) (i) any right of the Applicant to sub-licence any of the rights granted to the Applicant to the Materials under this MTA and/or (ii) any right of the Applicant to publish or distribute any of the Materials, except for the sole purpose of including a commensurate amount of supporting data (which shall not include any Participant Level Data) in the Applicant's publication of its Findings (which may include commensurate publication of certain of the Results Data, as the same may be reasonably required by the relevant publisher).

3.9 For the avoidance of doubt, the rights granted under this MTA to the Applicant to use the Materials are for the Permitted Purpose only.

4. Confirmations from the Applicant

General

4.1 The Applicant hereby confirms to UK Biobank that all work performed by it using the Materials shall be carried out in compliance with all applicable laws, regulations, guidelines and approvals, including without limitation the Human Tissue Act 2004, the Data Protection Legislation and any approvals required from a Research Ethics Committee (or the applicable equivalent in the jurisdiction where the Approved Research Project is to be conducted).

Security

4.2 The Applicant shall retain the Materials in a secure network system, at such standard which would be reasonably expected for the storage of valuable and proprietary sensitive/confidential data. Further, the Applicant shall be obliged to implement the appropriate technical and organisational measures as set out in Annex 2 (Security Measures) to protect the Materials from the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to the Materials (a "**Data Security Incident**"). By signing this MTA, the Applicant PI confirms that the Security Measures set out in Annex 2 are in place in order to protect the Materials.

4.3 The Applicant shall notify UK Biobank without undue delay (and in any event no later than 24 hours) after becoming aware of a reasonably suspected "near miss" or actual Data Security Incident which affects the Materials. Such notification must be sent by email to DPO@ukbiobank.ac.uk with a copy to access@ukbiobank.ac.uk.

4.3.1 The Applicant shall not delay such notification on the basis that the information is incomplete or the relevant investigation is ongoing. Further, the Applicant shall not make any external announcement, notifications to a supervisory authority or regulator about any such Data Security Incident without the express prior written consent of UK Biobank, unless required by law to do so.

4.3.2 Both parties shall cooperate and provide reasonable assistance to each other to facilitate the handling of the Data Security Incident.

Withdrawal of consent by participants

4.4 The Applicant confirms that it shall deal promptly and appropriately (in accordance with the Participants option to withdraw as set out on the UK Biobank website here) with any "no further use" withdrawals by Participants which UK Biobank notifies to the Applicant.

Identification of participants

4.5 The Applicant is expressly prohibited from (or attempting to):

4.5.1 developing, linking or re-engineering the Materials supplied to it so as to identify (directly or indirectly) any Participant;

4.5.2 identifying any Participant from the Materials provided by UK Biobank; or

4.5.3 contacting any Participant.

4.6 In the event that the Applicant inadvertently identifies any Participant then it shall notify UK Biobank immediately setting out (in reasonable detail) the circumstances by which it happened. Such notification must be sent by email to DPO@ukbiobank.ac.uk with a copy to access@ukbiobank.ac.uk.

4.7 Other than for the purposes of clause 4.6, the Applicant shall not:

4.7.1 share the identification of that Participant with any other person; or

4.7.2 attempt to contact the Participant themselves.

4.8 Without prejudice to the other provisions of this MTA, any actual or anticipatory breach of any provision of clauses 4.1, 4.2 and 4.4 to 4.7 inclusive shall entitle UK Biobank to terminate this Agreement with immediate effect and require the immediate return or destruction of any Materials provided by UK Biobank.

5. Publication of lay summary, submission of annual report and return and publication of Findings

Publication of summary on UK Biobank's website

5.1 After the Applicant has received the Materials for the Approved Research Project, UK Biobank shall be entitled to publish on its website:

5.1.1 the lay summary of the Approved Research Project contained in the Application (with the exception of any material that has been agreed by UK Biobank would be kept confidential); and

5.1.2 summary details of the Applicant.

Annual Project Report

5.2 During the Term, the Applicant shall provide UK Biobank with:

5.2.1 a report (with a summary section) setting out in reasonable detail the progress of the Approved Research Project in the form attached as Annex 3 (or in such other format as required by UK Biobank from time to time) on an annual basis (from the Effective Date) which shall include the Findings the Applicant has made which in its reasonable view may be:

(a) published or pending publication;

(b) disclosed in a published patent; or

(c) otherwise of significance (in the context of medical research); and

5.2.2 a summary (and a copy of the application if requested) of any patents whose claims cover, or are intended to cover, an Applicant-Generated Invention within two (2) months of their publication.

5.3 In relation to the Annual Project Report, UK Biobank:

5.3.1 shall have the ability to make the summary section of the Annual Project Report public, subject to the Applicant (as referred to in clause 5.6 below) retaining a reasonable period of confidentiality on items where patent rights still need to be filed; and

5.3.2 shall have the opportunity to ask the Applicant any reasonable questions arising from the Annual Project Report and the Applicant shall respond to such questions in a timely manner.

5.4 In the event that the Annual Project Report is not received by UK Biobank in the timeframe, manner and form prescribed, then the Applicant's rights under this MTA shall be suspended and the Applicant will not be able to obtain access to updated or additional Materials until such time as the Annual Project Report has been duly and compliantly provided. If the Annual Project Report is still outstanding, notwithstanding a reminder from UK Biobank, 3 months after the relevant anniversary of the Effective Date, then UK Biobank has the right to terminate this MTA by giving the Applicant written notice of termination and/or prevent the Applicant Institution (or Applicant PI) from applying for or accessing any further Materials from UK Biobank.

Publication of Findings

5.5 The Applicant shall use All Reasonable Endeavours to publish the Findings (and provide UK Biobank with a link thereto) within six (6) months after the Completion Date for the Approved Research Project:

5.5.1 in an academic journal; or

5.5.2 on an open-source publication site.

5.6 UK Biobank acknowledge and agree that the Applicant may keep such Findings confidential for a reasonable time in accordance with its reasonable research and development practices. For the avoidance of doubt, the Applicant is entitled to retain confidentiality regarding any Finding over which patent protection is being sought (and the patent has not yet been published).

5.7 If such Findings are made publicly available, UK Biobank requires that the Results Data underlying such Findings shall be promptly returned or otherwise made available to UK Biobank[1]. UK Biobank also

[1] For the avoidance of doubt, the intention of this provision is not to require the return of irrelevant or extraneous data sets but rather to make summary information available to other researchers (in a comparable form to that which academic journals often require), in particular so that it is not necessary for

requires that the Results Data are returned in a format which is appropriate and comprehensible (particularly for other researchers) along with any documentation which would be reasonably necessary to enable another researcher to interpret and understand the Results Data.

5.8 Within six (6) months after the publication of the Findings, the relevant Applicant shall provide to UK Biobank the Results Data in such form and format as set out in clause 5.7 above (alternatively UK Biobank and the relevant Applicant may agree that the relevant Applicant retains the Results Data on the basis that they are made publicly available to other Researchers and/or publicly available generally).

5.9 UK Biobank shall consider reasonably any written requests (containing an appropriate explanation) for an extension of the time limits set out in this clause.

5.10 The Applicant shall use All Reasonable Endeavours to publish a commensurate level of Findings in relation to the Approved Research Project within the first three (3) years of the Term (and in any subsequent extensions). Where this is not possible, the Applicant shall provide UK Biobank with a reasonable explanation as to why it is not possible and an estimation of when a publication can be expected.

Notification to UK Biobank

5.11 Unless otherwise stated in Annex 4, the Applicant is not required to obtain UK Biobank's approval to any report of its Findings. The Applicant shall nevertheless provide a copy of any report of its Findings and any press release to UK Biobank at least two (2) weeks before their expected date of first public presentation or publication in any format (e.g. paper journal, on-line report, meeting abstract). The Applicant shall upload such documents to AMS in the first instance. If this is not possible, the Applicant shall email such documents to access@ukbiobank.ac.uk.

5.12 However, and notwithstanding the provisions of clause 5.11 above, the Applicant is required to promptly notify UK Biobank in advance (in writing) if any report of its Findings is reasonably likely to provoke controversy or otherwise attract significant public attention. In such circumstances, UK Biobank reserves the right to make such recommendations, reservations or suggestions on the report as it sees fit (and which it may make public) for consideration by the Applicant.

Credit to UK Biobank

5.13 UK Biobank requires that any and all publications of Findings using UK Biobank data include the following credit, which should be incorporated within the "Acknowledgements" of such publication:

"This research has been conducted using the UK Biobank Resource under application number [ ]."

5.14 This acknowledgement to UK Biobank should, when possible, be linked to reference search tools (such as PubMed and MEDLINE and/or DOI reference).

6. Charges

6.1 The Applicant agrees to pay the Access Charges as set out in the payment section of the Applicant's Application to UK Biobank via bank transfer or Sage Pay, in cleared funds and in British pounds sterling (GBP). The Access Charges are stated to be exclusive of VAT. The Applicant shall pay any applicable VAT in addition to the Access Charges.

6.2 When paying the Access Charges, the Applicant shall quote the invoice number and/or Application Reference Number as the payment reference, and also send a remittance note to creditcontrol@ukbiobank.ac.uk.

6.3 The rights granted to the Applicant by UK Biobank under this MTA are conditional on the Access Charges (and applicable VAT) being paid and so, for the avoidance of doubt, no Materials shall be provided to the Applicant until or unless the Access Charges (and applicable VAT) are received in full. The Applicant shall pay the Access Charges (and applicable VAT) no later than thirty (30) days from the Effective Date.

6.4 If payment of the Access Charges has not been made within ninety (90) days of receipt of this MTA by the Applicant, the Applicant shall be required to re-apply for access to the UK Biobank resource and Materials.

---

a researcher (reviewing the Findings) to have to re-create certain derived variable or related metrics. Also, for clarity, Applicants shall have no obligation to provide to UK Biobank or publish, and do not grant UK Biobank any rights in or to, any genotype-phenotype data obtained or generated outside of the Approved Research Project.

7. Annual Confirmation and Audit

7.1 During the Term, UK Biobank requires the Applicant PI to confirm on an annual basis that the Approved Research Project remains compliant with the provisions of the MTA (and the Annexes). Specifically, the Applicant PI shall provide UK Biobank with such confirmation as part of the Annual Project Report in the form attached at Annex 3. In the event that the Annual Project Report is not received by UK Biobank in the timeframe, manner and form prescribed, UK Biobank reserves the rights set out in clause 5.4 above.

7.2 In circumstances where UK Biobank reasonably believes that a Data Security Incident or other serious incident has occurred then, on notice to the Applicant, in order to confirm or investigate compliance with the provisions of this MTA, UK Biobank may itself or via appropriate third parties:

7.2.1 choose to undertake an audit (either in person or remotely) in order to review the security, storage or other arrangements for the Materials; and

7.2.2 request such additional information about the Approved Research Project and/or its progress as UK Biobank may, from time to time, reasonably require.

7.3 UK Biobank shall bear the costs of such audits unless a material default within the procedures and processes of the relevant Applicant is discovered, in which case the relevant Applicant shall be obliged to reimburse the reasonable costs of UK Biobank and any relevant third parties.

7.4 UK Biobank confirms that its audit rights shall be exercisable no more than once a year and on the provision of reasonable notice (which may be immediate in the event of a Data Security Incident or other serious incident) to the Applicant. As far as practically possible, UK Biobank agrees to coordinate any site visits and audits with the other relevant parties.

8. Confidentiality

8.1 Subject to the exceptions in clause 8.2, UK Biobank shall keep confidential any information disclosed to it in writing by the Applicant that is marked confidential ("**Applicant's Confidential Information**") and shall not disclose such information to any person.

8.2 UK Biobank may disclose the Applicant's Confidential Information where expressly permitted by this MTA or when:

8.2.1 it is required to be disclosed by law, by any governmental or other regulatory authority, by a court or other authority of competent jurisdiction; or

8.2.2 it can be shown by UK Biobank (to the Applicant's reasonable satisfaction) to have been known by UK Biobank before disclosure to it by such Applicant; or

8.2.3 it was lawfully disclosed to UK Biobank by a third party who did not impose any restrictions on its disclosure; or

8.2.4 the information was in (or enters into) the public domain other than by reason of a breach of this clause by UK Biobank; or

8.2.5 UK Biobank and the Applicant agree, acting reasonably, that such information is trivial or obvious, or they agree in writing that such disclosure may be permitted.

9. Data Protection

Relationship of the parties[2]

9.1 The parties acknowledge that UK Biobank and the Applicant are independent controllers with respect to the Participant Level Data that is processed in accordance with this MTA, and that the Applicant shall

[2] This clause 9 addresses the requirements of the prevailing data protection legislation in the UK: principally the Data Protection Act 2018 (https://www.legislation.gov.uk/ukpga/2018/12/contents) and the UK GDPR and related guidance from the relevant regulators, particularly the ICO (https://ico.org.uk/). This clause also addresses the impact, from UK Biobank's perspective, of the United Kingdom leaving the European Union.
In relation to identifiable data two factors remain the same as under the original MTA:

- UK Biobank has and will continue to go to significant lengths to de-identify the data it releases to researchers, by removing direct and indirect identifiers, such that (even taking into account publicly available information) it should not be possible for a researcher to re-identify a participant;
- Further, the Applicant is expressly prohibited from actual (and making any attempt at) re-identification any Participant in accordance with clause 4.5 of the MTA.

Nevertheless, UK Biobank considers that it is appropriate for UK Biobank to require researchers to treat the UK Biobank data as if it is personal data, which requires the Applicant to agree to the provisions of this clause. The Applicant will be considered to be a separate and independent data controller (and not a data processor) under Data Protection Legislation. Please see the FAQs on the UK Biobank website (which shall be updated by UK Biobank from time to time)

process the Participant Level Data strictly for the Permitted Purpose. In no event shall the parties process the Participant Level Data as joint controllers.

9.2 Each party shall be individually and separately responsible for complying with the obligations that apply to it as a controller under Data Protection Legislation[3].

Cooperation

9.3 In the event that the Applicant, Applicant PI or any Applicant Researcher receives any correspondence, enquiry or complaint from a Participant, regulator or other third party ("**Correspondence**") in connection with the processing of the Participant Level Data, it shall promptly inform UK Biobank giving full details of the same. In all circumstances, the Applicant, Applicant PI or any Applicant Researcher shall: (i) obtain UK Biobank's written approval before responding to the Correspondence, including approval of the contents of any response; and (ii) subject to Data Protection Legislation, permit UK Biobank to respond directly to the Correspondence.

Where the Applicant is located outside of the UK

9.4 Where UK Biobank transfers Participant Level Data to an Applicant outside the UK in a territory that has not been specified as ensuring an adequate level of protection in accordance with Data Protection Legislation, the parties agree that the C2C Model Clauses shall be incorporated into this MTA by reference from the Effective Date as follows:

9.4.1 UK Biobank shall be the data exporter;

9.4.2 the Applicant shall be the data importer;

9.4.3 where the C2C Model Clauses being relied upon are those approved by the European Commission: (i) under the "II Obligations of the data importer" section of the C2C Model Clauses option h (iii) (the data processing principles set forth in Annex A) shall be deemed to have been selected; (ii) the provisions of Annex 1 shall be deemed to be set out in Annex B to the C2C Model Clauses; and (iv) the optional illustrative commercial clauses shall be deemed to have been deleted; and

9.4.4 if there is any conflict between the MTA and the C2C Model Clauses, the C2C Model Clauses shall prevail.

The parties agree to use All Reasonable Endeavours to put in place any additional or supplementary measures that may be required in order to give effect to the C2C Model Clauses.

International transfers by the Applicant

9.5 The Applicant shall not process any Participant Level Data (nor permit any Participant Level Data to be processed) in a territory outside of the UK (or where clause 9.4 applies, where processing occurs in a subsequent territory) unless it has taken such measures as are necessary to ensure the transfer is in compliance with Data Protection Legislation.

10. Limitation of Liability

10.1 The parties agree that:

10.1.1 subject to clauses 10.2, 10.3 and 10.4, UK Biobank's maximum aggregate Liability under this MTA and/or in relation to the Approved Research Project shall be limited to the Access Charges paid or payable by the Applicant to UK Biobank (whether or not invoiced to the Applicant) in relation to the Approved Research Project; and

10.1.2 subject to clauses 10.2, 10.3 and 10.5, the Applicant's maximum aggregate Liability under this MTA and/or in relation to the Approved Research Project shall be limited to the Access Charges paid or payable by the Applicant to UK Biobank (whether or not invoiced to the Applicant) in relation to the Approved Research Project.

10.2 Notwithstanding clause 10.1 above, UK Biobank shall have no Liability to the Applicant and the Applicant shall have no Liability to UK Biobank for any:

10.2.1 loss of profit (whether direct, indirect or consequential);

[3] https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/controllers-and-processors/what-does-it-mean-if-you-are-a-controller/#1

10.2.2 loss of use, loss of revenue, loss of production or loss of business (in each case whether direct, indirect or consequential);

10.2.3 loss of goodwill, loss of reputation or loss of opportunity (in each case whether direct, indirect or consequential);

10.2.4 loss of anticipated savings or loss of margin (in each case whether direct, indirect or consequential);

10.2.5 loss of use or value of any data or software (in each case whether direct, indirect or consequential); or

10.2.6 indirect or consequential loss.

10.3 Nothing in this MTA shall operate to exclude or limit any Liability which cannot legally be limited including but not limited to liability for:

10.3.1 death or personal injury caused by negligence;

10.3.2 for its fraud or fraudulent misrepresentation; and

10.3.3 for any matter for which it is not permitted by law to exclude or limit, or to attempt to exclude or limit, its Liability.

10.4 For the avoidance of doubt, UK Biobank shall have no responsibility or Liability (including but without limitation any product-related Liability) for any finding, product, test or treatment developed directly or indirectly by the Applicant using the Materials.

10.5 Nothing in this MTA shall operate to exclude or limit the Applicant's Liability to UK Biobank for any loss, damage, costs or expenses arising from:

10.5.1 the Applicant's failure to comply with clause 9 (Data Protection) and clauses 14.5 to 14.10 inclusive (Third Party Processors);

10.5.2 any breach of clause 2.2 or any circumstance in which the Applicant sub-licenses, distributes or otherwise shares the Materials (including any IPRs) with any unauthorised person or third party;

10.5.3 any circumstance set out in clauses 4.5 and 4.7; and

10.5.4 any Data Security Incident which is caused by the Applicant.

11. Term

11.1 The term of this MTA shall commence on the Effective Date and shall end on the Completion Date unless terminated sooner in accordance with clause 12 or in accordance with law.

11.2 The Term of this MTA may be extended by the Applicant (and with the agreement of UK Biobank) during the final year of the Approved Research Project in the following one (1) year increments:

11.2.1 for a minimum of period of one (1) year;

11.2.2 for a period of two (2) years; or

11.2.3 for a maximum period of three (3) years;

on application to UK Biobank setting out (in reasonable detail) the reasons for the extension request and subject to the payment of the relevant further Access Charges.

11.3 For the avoidance of doubt, the extensions set out in clause 11.2 above can be applied cumulatively (subject to applicable Access Charges) so that, for example, an extension of 3 years may be granted to take the Approved Research Project duration from 3 years to 6 years, and this may then be extended by a further 3 years to 9 years and so on.

12. Termination and consequences of termination

12.1 UK Biobank shall be entitled to terminate this MTA immediately by written notice to the Applicant if the Applicant:

12.1.1 commits any breach of a material provision of this MTA or a material breach of this MTA, and, in the case of a breach capable of remedy, fails to remedy the same within 10 days after receipt of a written notice giving particulars of the breach and requiring it to be remedied; or

12.1.2 ceases, is likely to cease, or threatens to cease carrying on business or suffers an Insolvency Event, or is subject to a serious, adverse regulatory finding.

12.2 Upon expiry of the MTA pursuant to clause 11.1 above or termination of this MTA by UK Biobank pursuant to clause 12.1 or in accordance with law:

12.2.1 the grant of rights and all licences to the Applicant under this MTA shall be automatically terminated; and

12.2.2 the Applicant shall destroy the Materials or otherwise render them permanently inaccessible. For the avoidance of doubt, the Applicant shall not be required to destroy Results Data or Other Data subject to the provisions of this MTA being complied with.

12.3 Without prejudice to the foregoing and to any other rights or remedies that UK Biobank may have, UK Biobank may take the following steps if there is a breach that entitles UK Biobank to terminate this MTA under clause 12.1:

12.3.1 it may prohibit the Applicant PI, Applicant Researchers and any other researchers from the Applicant Institution from accessing any further Materials from within the UK Biobank resource for an indefinite period of time; and/or

12.3.2 it may elect to inform the relevant personnel within the defaulting Applicant Institution, funders of the defaulting Applicant PI and/or governing or other relevant regulatory bodies.

12.4 Notwithstanding termination of this MTA for any reason, the provisions of clauses 2, 3, 4, 5, 7, 8, 9, 10, 12, 13, 14, 16 and 17 shall continue in force in accordance with their respective terms.

12.5 Termination or expiry of this MTA shall not affect the rights and obligations of the parties accrued at the date or termination or expiry.

13. Notices

13.1 Notices required under this MTA shall be in writing and shall be:

13.1.1 sent by email to the addresses set out below; or

13.1.2 (in the event of failure to deliver an email) by post to the registered address of UK Biobank or the Applicant.

13.2 Any notice shall be deemed to be received:

13.2.1 if sent by email, upon receipt at the recipient's email server, (or, if this time falls outside business hours in the place of receipt, when business hours resume); or

13.2.2 if sent by post, on the date of delivery if a business day in the place of receipt (or, if not a business day, on the first business day thereafter).

13.3 Notices to UK Biobank shall be sent to the access team at access@ukbiobank.ac.uk. Notices to the Applicant shall be sent by email to the relevant Applicant and the Applicant PI.

14. Affiliates, assignment and sub-contracting

Affiliates

14.1 The rights granted to the Applicant under this MTA for the Approved Research Project include the Affiliates of the Applicant, subject to the Applicant:

14.1.1 providing updated details of each Affiliate in the Annual Project Report submitted to UK Biobank on an annual basis in accordance with clause 7.1 of the MTA;

14.1.2 remaining fully liable and responsible to UK Biobank for all acts, defaults and omissions of each of its Affiliates as if they were the Applicant's own; and

14.1.3 ensuring that each of its Affiliates comply with the terms and conditions of this MTA.

Assignment

14.2 Neither UK Biobank nor the Applicant shall be entitled to assign this MTA or any of its rights or obligations hereunder without first having received the written approval of the other party, such approval not to be unreasonably withheld or delayed.

Subcontracting

14.3 Other than in the circumstances set out in clause 14.5, the Applicant shall not sub-contract the performance of any of its obligations under the MTA or any part thereof without having first obtained the prior written consent of UK Biobank, such consent not to be unreasonably withheld.

14.4 In the event that consent is granted under clause 14.3, the relevant Applicant shall be responsible for the acts, defaults and omissions of its sub-contractors as if they were the Applicant's own, and any consent given shall not relieve such relevant Applicant of any of its obligations under this MTA.

Third Party Processors

14.5 UK Biobank acknowledges and agrees that the Applicant may subcontract to third party processors to process the Materials strictly for the Permitted Purpose and only in relation to discrete elements of data computation and analysis (such processors being, "**Third Party Processors**"). The Applicant must comply with, and only engage Third Party Processors strictly in accordance with the terms set out in clauses 14.6 to 14.10 inclusive.

14.6 The Applicant warrants that the Third Party Processor is not a Collaborator and shall only be engaged for the purposes of discrete elements of data computation and analysis in relation to the Permitted Purpose (the "**Processor Task**").

14.7 Prior to engaging a Third Party Processor, the Applicant shall conduct and document the following assessment:

14.7.1 whether the Third Party Processor is necessary for the progress of the research aims of the Approved Research Project;

14.7.2 whether the Third Party Processor is a suitable recipient for the data in terms of both its provenance on past data security and past data usage / activities (for example Cambridge Analytica would not qualify); and

14.7.3 whether the Third Party Processor is able to provide sufficient assurance(s) that it shall process the Materials in a manner that will meet the requirements of Data Protection Legislation.

14.8 The Applicant shall:

14.8.1 remain fully responsible to UK Biobank for all acts, defaults and omissions of the Third Party Processor as if they were the Applicant's own;

14.8.2 provide only such Materials to the Third Party Processor as is strictly necessary for the Third Party Processor to perform the Processor Task;

14.8.3 provide details of each Third Party Processor and the Processor Task in the Annual Project Report submitted to UK Biobank on an annual basis in accordance with clause 7.1 of the MTA; and

14.8.4 only engage the Third Party Processor on the basis that a written agreement with the Third Party Processor is executed prior to any data transfer or processing of Materials taking place. Such agreement must include inter alia:

(a) a clear definition and scope of the Processor Task, including an agreement only to process the data in accordance with the Applicant's documented instructions;

(b) to authorise the Third Party Processor only to undertake the Processor Task and not to perform any other act, unless expressly authorised to do so;

(c) to store, process and use the Materials to the security standards set out in the MTA (as a minimum) and implements appropriate technical and organisational security measures to protect the Materials against a Data Security Incident;

(d) to delete (or render permanently inaccessible) the Materials (and any data generated as a result of the Processor Task) once the Processor Task has been completed;

(e) to confirm that the Third Party Processor has no rights (directly or indirectly) in either any Materials (or data derived therefrom) or from anything which the Applicant has created or done as part of the Approved Research Project (which is covered by the MTA between UK Biobank and the Applicant);

(f) to confirm that the Third Party Processor is bound by the provisions which are equivalent to the relevant provisions in the MTA, including, but not limited to: a) not to transfer the Materials (or data derived therefrom) to any third party and b) not to make any attempt to re-identify any Participant;

(g) that the Third Party Processor provides sufficient assurance(s) that it shall process the Materials in a manner that will meet the requirement of Data Protection Legislation; and

(h) that the Applicant has an unfettered unilateral right to terminate its agreement with the Third Party Processor immediately if a material problem arises (including a breach by the Third Party Processor of any of the above provisions).

14.9 The Applicant must keep the activities of the Third Party Processor under reasonable review in order to ensure compliance with clauses 14.5 to 14.10 inclusive.

14.10 In the event that UK Biobank raises any concern regarding the identity of the Third Party Processor or the activities of a Third Party Processor, the Applicant shall investigate and report on the matter promptly. UK Biobank may require, if reasonably necessary (and subject to a dialogue with the Applicant), the Applicant to:

14.10.1 to audit the Third Party Processor; and / or

14.10.2 terminate the agreement with the Third Party Processor.

15. Force majeure

15.1 If a party is prevented from, hindered or delayed in performing any of its obligations under this MTA by reason of a Force Majeure Event, such party shall promptly notify the other of the date of its commencement and the effects of the Force Majeure Event on its ability to perform its obligations under this MTA. If mutually agreed by the parties, then the obligations of the party so affected shall thereupon be suspended for so long as the Force Majeure Event may continue.

15.2 The party affected by a Force Majeure Event shall not be liable for any failure to perform or delay in performing such of its obligations as are prevented, hindered or delayed by the Force Majeure Event provided that such party shall use every reasonable effort to minimise the effects thereof and shall resume performance as soon as possible after the removal of such Force Majeure Event. If the period of non-performance exceeds 90 days from the start of the Force Majeure Event then the non-affected party shall have the option, by written notice to the other party, to terminate this MTA by giving thirty (30) days' written notice to the other party.

15.3 The provisions of this clause 15 shall not affect any other right which any party may have to terminate this MTA.

16. Dispute resolution

16.1 If a Dispute arises, the parties shall follow the procedure set out in this clause 16.

16.2 Either party may give the other party written notice of a Dispute, setting out its nature and full particulars ("**Notice of Dispute**"), together with relevant supporting documents. Within five (5) business days of service of the Notice of Dispute, a UK Biobank representative and a representative from the Applicant shall attempt in good faith to resolve such Dispute.

16.3 If for any reason the respective representatives of the parties are unable to resolve the Dispute within ten (10) business days of the Notice of Dispute, then any of the parties involved in the respective Dispute may refer it for discussion by UK Biobank's Principal Investigator and appropriate senior officer(s) of the Applicant. These senior representatives of the parties (or their respective nominees) shall seek to arrange a meeting or telephone or videoconference call promptly with a view to resolving the Dispute.

16.4 If, following escalation of any Dispute as set out in clause 16.3, UK Biobank's Principal Investigator and appropriate senior officer(s) of the Applicant are for any reason unable to resolve the Dispute within

thirty (30) business days of it being escalated to them, then the parties agree to enter into mediation in good faith the settle the Dispute in accordance with the Centre for Effective Dispute Resolution (CEDR) Model Mediation Procedure. Unless otherwise agreed between the parties within 20 business days of service of the Notice of Dispute, the mediator shall be nominated by CEDR. To initiate the mediation, a party must serve notice in writing to the other party to the Dispute, referring the Dispute to mediation.

16.5 For avoidance of doubt, Disputes with respect to scientific or technical issues or business decisions, and not legal issues, shall remain with senior representatives to be resolved.

16.6 If the Dispute is not settled by mediation within 10 business days of commencement of mediation or within such further period as the parties may agree in writing, either party may issue court proceedings in accordance with clause 17.10 of this MTA.

16.7 Nothing in this clause 16 shall serve to prevent any of the parties from seeking interim/injunctive relief to protect its rights and interests in any court of England and Wales; provided that such relief shall not prevent or stay any mediation.

17. General

17.1 The parties agree that the Applicant may change the Applicant PI at any time, and from time to time, by written notice to UK Biobank.

17.2 This MTA governs and constitutes the entire agreement between the parties and supersedes, replaces and extinguishes all previous agreements, promises, assurances, warranties, representations and understandings between them (whether oral or written) relating to the subject matter hereof. Further, each party acknowledges and agrees that it does not rely on, and shall have no remedy in respect of, any statement, promise, assurance, statement, warranty, undertaking or representation made (whether innocently or negligently) by the other party or any other person except as expressly set out in this MTA in respect of which its sole remedy shall be for breach of contract.

17.3 If there is any conflict between the provisions of this MTA and any of the Annexes, then the provisions of the relevant Annex shall apply.

17.4 A waiver, delay or forbearance by any party, whether express or implied, in enforcing or exercising any of its rights or remedies hereunder shall not constitute a waiver of such right or remedy, unless set forth in a writing signed by the waiving party.

17.5 No provision of this MTA is intended to be enforceable by any person who is not a party to this MTA and nor are any rights granted to any third party under statute or otherwise.

17.6 Nothing in this MTA shall create a partnership, joint venture or relationship of agency among the parties.

17.7 All variations to this MTA must be agreed, set out in writing and signed on behalf of the parties before they take effect.

17.8 If any provision or part-provision of this MTA is or becomes invalid, illegal or unenforceable, it shall be deemed deleted, but that shall not affect the validity and enforceability of the rest of this MTA.

17.9 If any provision or part-provision of this MTA is deemed deleted under clause 17.8, the parties shall negotiate in good faith to agree a replacement provision that, to the greatest extent possible, achieves the intended commercial result of the original provision.

17.10 This MTA and any dispute or claim (including non-contractual disputes or claims) arising out of or in connection with it or its subject matter or formation shall be governed by and construed in accordance with the laws of England and Wales. Subject to clause 16 above, the parties irrevocably agree that the English courts shall have exclusive jurisdiction over any suit, action, proceedings or dispute arising out of, or in connection with this MTA or its subject matter or formation.

This MTA is executed by duly authorised representatives of the parties.

For and on behalf of UK Biobank:		**For and on behalf of the Applicant Institution:**	
Signature:	DocuSigned by:	**Signature:**	DocuSigned by: Dr Tim Horne
Print name:	Jonathan Sellors	**Print name:**	Dr T J Horne
Position:	General Counsel & Company Secretary	**Position:**	Director of Research
Date:	5/4/2022	**Date:**	14/2/2022

I am the Applicant Principal Investigator of this Approved Research Project and by signing below I confirm that I have read and understood the provisions of this MTA.

Signature: DocuSigned by:

Print name: Esphie Grace Fojas

Position: Mrs Researcher

Date: 14/2/2022

Definitions

Access Charges: the charges payable by the Applicant (which may include VAT) to access the Materials and, where applicable, to allow a Collaborator access to the Materials as summarised in Annex 4 and detailed in the payment section of the Application on AMS.

Affiliate: any company or other entity that is directly or indirectly Controlling, Controlled by or under common Control with an Applicant (which includes if such Applicant is a company, a subsidiary or parent or holding company of such Applicant, or a subsidiary of such parent or holding company) for so long as such Control exists.

All Reasonable Endeavours: in respect of a party obliged to use "All Reasonable Endeavours", the pursuance of a reasonable course of action to achieve the stated outcome which may require reasonable expenditure, but does not require the party to pursue every available course of action to achieve the outcome or act outside its own operational or commercial interests.

AMS: the online Access Management System the Applicant uses to apply for and manage its access to the UK Biobank resource.

Applicant *or* ***Applicant Institution:*** the organization (e.g. University, company or other identifiable legal entity) making the Application for access in respect of the Approved Research Project and by which an Applicant PI is employed or otherwise contractually attached.

Applicant's Confidential Information: as defined in clause 8.1 of this MTA.

Applicant-Generated Inventions: as defined in clause 3.5 of this MTA.

Application: the application by the Applicant PI and their Institution to UK Biobank for access to the Materials for use in relation to the Approved Research Project.

Applicant Principal Investigator or ***Applicant PI:*** the principal investigator of the Approved Research Project.

Applicant Researcher: a researcher at the Applicant who is working with an Applicant PI on the Approved Research Project.

Approved Research Project: the research project approved by UK Biobank (specifically including any conditions or stipulations made by UK Biobank) and as set out in Annex 4.

C2C Model Clauses: the model clauses for the transfer of personal data to controllers established in third countries approved by the European Commission, the approved version, of which, in force at present is that set out in the European Commission's Decision 2004/915/EC of 27 December 2004 (available at https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32004D0915), as such model clauses may be amended or superseded by the Secretary of State or standard data protection clauses specified in a document issued (and not withdrawn) by the UK Information Commissioner;

Collaborator or ***Collaborator Institution:*** the organization (e.g. University, company or other identifiable legal entity) which employs the Lead Collaborator who is collaborating with the Applicant PI on the Approved Research Project.

Collaborator Researcher: a researcher who is working with the Lead Collaborator at a Collaborator Institution on the Approved Research Project.

Completion Date: the date or dates contained within Annex 4 for the end-date of the Approved Research Project, including any extensions.

Control: means the direct or indirect ownership of at least fifty percent (50%) of the outstanding shares or other voting rights of the subject entity having the power to vote on or direct the affairs of the entity (or such lesser percentage which is the maximum allowed to be owned by a foreign company in a particular jurisdiction), and "Controlling" and "Controlled" shall be construed accordingly.

controller, processor, data subject, personal data, processing *(and* ***process)*** *and* ***special categories of personal data:*** have the meanings given in Data Protection Legislation;

Data Protection Legislation: means all laws applicable (in whole or in part) to a party's processing of personal data under or in connection with this MTA, and including, as applicable: (i) the GDPR as it forms part of UK law by virtue of section 3 of the European Union (Withdrawal) Act 2018 (the "**UK GDPR**"); (ii) the UK Data Protection Act 2018; (iii) the Privacy and Electronic Communications (EC Directive) Regulations 2003 as they continue to have effect by virtue of section 2 of the European Union (Withdrawal) Act 2018; and (iv) any other laws in force in the UK from

time to time applicable (in whole or in part) to the processing of personal data, in each case as amended or superseded from time to time.

Data Security Incident: as defined in clause 4.2 of the MTA.

Dispute: any dispute, controversy, proceeding or claim (including any legal disputes) between UK Biobank, on the one hand, and the Applicant, on the other hand, arising out of or in connection with this MTA or the performance, validity or enforceability of it.

Effective Date: the date on which this MTA is executed by an authorised signatory of UK Biobank, having already been signed by the Applicant Institution and signed as "read and understood" by the relevant Applicant PI.

Findings: as defined in clause 3.1.2 of this MTA and shall mean literally what is found, in terms of conclusions and results, by the Applicant as a result of the Approved Research Project. For clarity, Findings do not include Applicant-Generated Inventions and nor do they include findings which result from data which is not UK Biobank Materials.

Force Majeure Event: any cause which arises from or is attributable to acts, events, omissions or accidents beyond the reasonable control of the affected party including without limitation act of God, war, riot, civil commotion, non-performance by sub-contractors or suppliers, compliance with any law or governmental order, rule, regulation or direction, accident, breakdown of plant or machinery, supply failure, epidemic, pandemic, fire, flood or storm.

Insolvency Event: means where a person is unable to pay its debts within the meaning of the Insolvency Act 1986 section 123 (without the need for a determination by a court), has an administrator, receiver, administrative receiver or manager appointed over the whole or any part of its assets, enters into any composition with creditors generally, or has an order made or resolution passed for it to be wound up (unless as part of any scheme for solvent amalgamation or solvent reconstruction) or undergoes any similar or equivalent process in any jurisdiction or undergoes any other arrangement which affects the rights of creditors;

Intellectual Property Rights or IPRs: all present and future intellectual property rights including but not limited to patents, trade and service marks, design rights, copyright, database rights, trade secrets and know-how, in all cases whether registered or not or registerable, and including all registrations and applications for registrations of any of these and rights to apply for the same as well as any renewals, extensions, continuations, combinations or divisions thereof, and all rights and forms of protection of a similar nature or having equivalent or similar effect to any of these anywhere in the world.

Lead Collaborator: the lead investigator at a Collaborator Institution.

Liability: liability arising out of or in connection with this MTA, whether in contract, tort, misrepresentation, restitution, under statute or otherwise, including but not limited to arising from a breach of, or a failure to perform or defect or delay in performance of, any of a party's obligations under this MTA, in each case howsoever caused, including if caused by negligence.

Materials: the data as set out in Annex 4 supplied by UK Biobank to the Applicant under or in connection with this MTA including any Participant Level Data.

MTA: this Material Transfer Agreement, the Applicant Terms and Conditions (including any documents and/or materials that are referred to in them), the Annexes and where applicable the contents of the Applicant's Application Form.

Notice of Dispute: as defined in clause 16.2 of the MTA.

Other Data: as defined in clause 3.1.3 of the MTA.

Participant(s): the individuals who participate in UK Biobank.

Participant Level Data: the personal data as described in Annex 1 contained within the Materials and any applicable generated data (as described in clause 3.1 of the MTA).

Permitted Purpose: to conduct the Approved Research Project in accordance with the approved project scope and the timeframe as set out in the Annex 4, subject to the provisions of this MTA.

Results Data: as defined in clause 3.1.1 of this MTA.

Term: as defined in clause 11.1 of this MTA.

Third Party Processors: as defined in clause 14.5 of this MTA.

VAT: value added tax chargeable under the Value Added Tax Act 1994 (and all amendments and updates thereto) or any similar replacement or additional tax.

Annex 1
Data Processing Description

This Annex 1 forms part of this MTA and describes the types of Participant Level Data disclosed by UK Biobank to the Applicant, the Applicant PI and Applicant Researchers to process strictly for the Permitted Purpose described in this MTA (or as otherwise agreed in writing by the parties).[4]

Data subjects	The Participants
Categories of data	The Participant Level Data to be processed concern the following categories of personal data: • EIDs – the encoded and unique pseudonymised identifiers, which are specific to the Approved Research Project; and • data derived from baseline questionnaire responses and interviews which do not contain special category data, such as birthplace, early life and education, employment history, marital status and number of children.
Special categories of data	The Participant Level Data to be processed concern the following special categories of data: The UK Biobank resource contains health, genetic and biometric data. All special categories of data contained in the Materials is de-identified (the direct and indirect identifiers are removed). The types of special category of data may include: • measures of the Participant's phenotype, such as height, weight and blood pressure (approximately 2,000 phenotypes per Participant, as further detailed here http://biobank.ndph.ox.ac.uk/showcase/schema.cgi?id=1) • measures of the Participant's genome, this includes genotype, exome sequence and whole sequence data; • biomarkers created by assay of the Participant's samples, which include common biomarkers (such as cholesterol), infectious disease markers, proteomic and metabolomic markers; • imaging data (on up to 100,000 Participants) as the result of MRI scans of the head, the heart and the body, plus ultrasound and DEXA; • data derived from health record linkages including hospital records, primary care records, death and cancer registries or any other sources of clinical data; and • other special category data derived from baseline/online questionnaire responses and interviews, such as past illness / disease history, dietary, cognitive and physical measures.
Purpose of the transfer	The transfer is made to allow the Applicant to conduct the Permitted Purpose.
Recipients	The Participant Level Data transferred may be disclosed only to the following recipients or categories of recipients: • authorised personnel within the Applicant, namely the Applicant Principal Investigator and Applicant Researchers; • Third Party Processors subject to the relevant provisions of the MTA; • Affiliates subject to the relevant provisions of the MTA; • law enforcement agencies acting under Data Protection Legislation; • the relevant data protection authority acting under Data Protection Legislation; and • auditors (UK Biobank or appropriate third parties).
Processing activities	The Participant Level Data will be subject to the following basic processing activities: • access and use of Participant Level Data within the research analysis platform for the Permitted Purpose; • where approved by UK Biobank the transmission to, making available to and storage on the Applicant's systems/network servers, excluding any WGS (whole genome sequence) or WES (whole exome sequence) files which must not be transmitted or downloaded from the research analysis platform; • research operations, including a Processor Task by a Third Party Processor; and • risk management, compliance, legal and audit functions.
UK Biobank's lawful basis for sharing personal data	Personal data: • Legitimate interests (Article 6(1)(f) UK GDPR) Special categories of data: • Scientific research purpose (Article 9(2)(j) UK GDPR)
UK Biobank's DPO contact details:	DPO@ukbiobank.ac.uk
Applicant's DPO (or other person responsible for data protection) contact details:	d Ian Blachford, Chief Operating Officer dataprotection@staffs.ac.uk

[4] For further information about this MTA's Data Protection clauses and an explanation of UK Biobank's position in relation to Data Protection, please see the FAQs on the UK Biobank website (which shall be updated by UK Biobank from time to time).

Annex 2
Security Measures

UK Biobank has an obligation under the UK GDPR to ensure that its Materials are stored, retrieved and used securely, with appropriate organisational and technical measures in place. UK Biobank must also take reasonable steps to ensure that Materials it shares continue to be protected with adequate security. This Annex 2 forms part of the MTA and represents a generic level of security standards for data storage, retrieval and usage that the Applicant must comply with. This Annex 2 may be updated by UK Biobank from time-to-time.

The objective of these security measures is to ensure that Materials provided by UK Biobank are secured and treated as though they are personal data, with appropriate measures in place to restrict access only to authorised users and to protect from unauthorised access by internal and external parties.

1. Information security policy

1.1 The Applicant shall implement and maintain a written information security policy that specifies the technical and organisational measures it shall apply to protect the Materials it processes in accordance with this MTA against unauthorised access and/or unlawful processing. The information security policy shall also describe the measures to be taken in the event of an actual or suspected data security breach.

1.2 The Applicant shall appoint a duly skilled individual with responsibility for ensuring the security of the Materials processed by the Applicant in its organisation and for reviewing, maintaining and updating the Applicant's information security policy.

1.3 The Applicant shall ensure authorised individuals who have access to the Materials are aware of their responsibilities for any data they handle, including appropriate training in order to fulfil their roles.

1.4 The information security policy shall also set out that:

1.4.1 information should be stored in an environment suited to its format and sensitivity, to ensure its preservation from physical harm or degradation and its security from unauthorised access;

1.4.2 data storage devices are appropriately protected and access controlled; and

1.4.3 servers, client devices and applications used for storing, accessing and analysing UK Biobank Materials are appropriately maintained with operating systems, firmware, and software within vendor supported versions where exceptions are documented with mitigations described.

2. Access to data

2.1 The Applicant shall implement access controls that restrict access to data it processes to duly authorised individuals and only to the extent necessary for the performance of their duties. Access should be controlled with usernames and appropriately secure passwords, with consideration given to the use of multi-factor authentication methods. Access logging and monitoring should be put in place.

2.2 The Applicant shall ensure that authorised individuals do not share or use the same username, and exceptions must be documented with adequate mitigations described and auditable.

2.3 The Applicant shall identify and appoint a duly skilled administrator with responsibility for granting, changing or voiding data access privileges; access privileges should be periodically reviewed.

2.4 Where an individual who has access to the Materials leaves or has their authorisation removed (e.g. as a result of a change of role) the Applicant PI shall ensure that their status is updated within 24 hours.

3. Storage and transmission of data

3.1 Where practicable data should be encrypted at rest and during transmission using strong encryption techniques; best practices should be followed for key management.

3.2 Use of portable media should be avoided unless reasonably required to process the data. Where needed, data must be encrypted using a strong password or other secret information.

3.3 Any deletion of data should be permanent and deleted data should not be recoverable.

3.4 An information asset register should be maintained so that all UKB data can be removed on request or at the end of the agreement.

Annex 3
Annual Project Report Template

The purpose of this Annex 3 is to provide the Applicant PI with a template of the Annual Project Report Form that will need to be completed and submitted to UK Biobank on an annual basis (the annual anniversary of the Effective Date). For the avoidance of doubt, the Applicant PI is not required to complete this form on execution of the MTA.

UK Biobank reserves the right to update this form from time to time including the manner in which it is submitted. Up-to-date versions of the form and instructions for submission are accessible on UK Biobank's AMS and website.

Applicant Annual Project Report

Every year, the Applicant Principal Investigator (PI) for a UK Biobank Research Project is required to provide some information regarding their project and confirmation that they are complying with the terms of the Material Transfer Agreement (MTA). Please complete and upload this report to the Access Management System (AMS). For help, please see our AMS User Guide on the UK Biobank website. (Failure to provide this report will result in additional data not being available for the project, and continued failure to comply could lead to the project being terminated and future applications declined.)

Research Project Number:	
Date report completed:	

1. **Are all Collaborators currently accessing UK Biobank data for this Research Project named in the Collaborators list in AMS?**

Yes / No	*If no, please ask any unregistered collaborators to submit a registration as soon as possible. After registration approval you can add them to the Collaborator list. Please remove any collaborators who are no longer accessing UK Biobank data.*

2. **Please provide the names of any Affiliates[5] who have access to UK Biobank data for this Research Project. If none, please say so:**

Affiliates:

3. **Please provide the names of any Third Party Processors who process UK Biobank data as a sub-contractor for this Research Project and provide details of the tasks the Third Party Processor conducts on your behalf. If none, please say so:**

Third Party Processors:
Third Party Processor tasks:

4. **Please provide an update (with a summary section) on the progress of your Research Project and your plans for the next 12 months:**

Progress to date:
Plans for the next 12 months:

5. **Please provide details of any research output (e.g. websites, patents, GWAS summary statistics location) and list all publications (including pre-prints and peer-reviewed publications) that derive from this Project since your last Annual Project Report. If none, please say so:**

Research Output:
Publications: (*in the format: first author, year, title, journal, PMID, DOI, web links to papers/patents*)

[5] An Affiliate means any company or other entity that is directly or indirectly Controlling, Controlled by or under common Control with an Applicant (which includes if such Applicant is a company, a subsidiary or parent or holding company of such Applicant, or a subsidiary of such parent or holding company) for so long as such Control exists. Please see the definition of 'Control' in the MTA for further information.

6. **I confirm that:**

Please mark with an 'x':

- I am the Applicant PI of the Research Project identified above; ☐
- I am aware of all research being undertaken for the Research Project identified above; and ☐
- The provisions of the Material Transfer Agreement (including but not limited to the Annexes) are being complied with. ☐

Note: please save as a .pdf document and upload in AMS

Annex 4

Approved Research Project

Approved Application Reference Number: 77577

Approved Application Name: Metabolic Syndrome-Associated Genes and Progression to Type 2 Diabetes and Cardiovascular Disease in Adults

Details of the Approved Research Project (including the Project Scope, Timeframe, Materials (together with any specific conditions), Access Charges and Collaborators can be accessed here:
Approved project details

Change Requests

The following requests can be made via the UK Biobank Access Management System (AMS) and shall be subject to review and approval (and where applicable, additional Access Charges):

- Requests to extend the Project Scope
- Requests to extend the project duration
- Requests to add additional data to the Approved Research Project

UK Biobank will also facilitate the following changes to the Approved Research Project where this is required by the Applicant:

- Add additional Collaborators to the Approved Research Project (subject to the applicable Access Charge)
- Changes to the Applicant PI of the Approved Research Project

8. University of Staffordshire Research Ethics Proportionate Review

STAFFORDSHIRE UNIVERSITY

Health and Social Care

ETHICAL APPROVAL FEEDBACK

Researcher name:	Esphie Grace Fojas
Title of Study:	SU_21_086 Metabolic Syndrome (MetS)-Associated Genes and Progression to Type 2 Diabetes and Cardiovascular Disease in Adults (approved by the UK Biobank)
Status of approval:	Approved

Thank you for addressing the committee's comments. Your research proposal has now been approved by the Ethics Panel and you may commence the implementation phase of your study. You should note that any divergence from the approved procedures and research method will invalidate any insurance and liability cover from the University. You should, therefore, notify the Panel of any significant divergence from this approved proposal.

You should arrange to meet with your supervisor for support during the process of completing your study and writing your dissertation.

When your study is complete, please send the ethics committee an end of study report. A template can be found on the ethics BlackBoard site.

Signed: **Date: 31st January 2022**

Dr Edward Tolhurst

Chair of the Health and Social Care Ethics Panel

RESEARCH ETHICS

Proportionate Review Form

The Proportionate Review process may be used where the proposed research raises only minimal ethical risk. This research must: focus on minimally sensitive topics; entail minimal intrusion or disruption to others; and involve participants who would not be considered vulnerable in the context of the research.

PART A: TO BE COMPLETED BY RESEARCHER

Name of Researcher:	**Esphie Grace Fojas**
School	**Staffordshire University**

Student/Course Details (If Applicable)			
Student ID Number:			**f022210j**
Name of Supervisor(s)/Module Tutor:			**Prof. Roozbeh Naemi**
PhD/MPhil project:	☒		
Taught Postgraduate Project/Assignment:	☐	**Award Title:**	**Professional Doctorate in Healthcare Science**
Undergraduate Project/Assignment:	☐	**Module Title:**	**Doctoral Research Thesis**

Project Title:	**Metabolic Syndrome (MetS)-Associated Genes and Progression to Type 2 Diabetes and Cardiovascular Disease in Adults (approved by the UK Biobank)**
Project Outline:	**Aim:** **This study aims to investigate the genes linked to progression to Type 2 Diabetes (T2D) and Cardiovascular Disease (CVD) in adults with metabolic syndrome (MetS).** **Objectives:** **(1) Identify all MetS-associated genes from genome-wide association studies (GWASs) through literature review process** **(2) Identify individuals with MetS-associated genotype from the UK Biobank** **(3) Characterize MetS amongst the identified individuals** **(4) Identify genes which predispose individuals with MetS for the development of T2D and CVD** **Context of the Investigation / Synopsis:** **Metabolic syndrome (MetS) is the clustering of CVD and T2D risk factors including abdominal obesity, dyslipidemia, hypertension, and hyperglycemia. Although extensive investigations on individual components of MetS have been widely reported, studies on MetS as an entity are notably scarce in general, and genetic studies are limited in particular. With the burgeoning rise in prevalence of CVD and T2D along with associated mortalities, healthcare burden, and economic impact around the world, investigating causative aspects of the disease is crucial. This includes exploring the genetic basis of the disease across different populations and investigating other relevant factors such as sex, socioeconomic status, lifestyle, and medications. With the advent of personalized medicine, an increasingly applied treatment modality for various**

	diseases, this may prove timely and beneficial. Although GWASs have identified MetS-specific associated genes, the incidences of T2D and CVD in the cohorts were ambiguously evaluated, with consideration of confounding factors. In this light, this study aims to investigate the genes linked to progression to T2D and CVD in adults with MetS amongst participants in the UK Biobank. Theoretical Basis: MetS is a major non-communicable health hazard globally and has seen unprecedented rise in the recent decades, associated with the rapidly increasing obesity prevalence. MetS, a constellation of CVD and T2D risk factors, is characterized by abdominal obesity, insulin resistance, hypertension, and hyperlipidemia. T2D has become one of the major causes of premature illness and death mainly through this increased risk of CVD which is responsible for up to 80 per cent of these deaths. (1) The definition of MetS has not been made universal, however, and recognized international authorities such as World Health Organization (WHO), National Cholesterol Education Program (NCEP), International Diabetes Federation (IDF), American Association of Clinical Endocrinologists (AACE), and European Group of Insulin Resistance (EGIR) have issued similar, yet specific guidelines. (2,3) Numerous research studies have widely investigated on the occurrence of individual risk factors which comprise MetS and their associations, such as dyslipidemia and diabetes (4, 5). Several investigations from linkage analysis, candidate gene approach and GWASs have likewise focused on the causative genetic aspects and established heritability of a single MetS component. The heritability of each component individually was found to be between 16-60% with lipids/glucose and obesity at 44% and blood pressure at 20%, while a study done in Italy found the heritability of MetS to be 27% (2). MetS treatment and therapies are also primarily targeted on only one metabolic trait. As such, MetS medication categories include antidiabetics (metformin, thiazolidinediones, SGLT2 inhibitors, glucagon-like peptide-1 agonist), lipid-lowering agents (statins and non-statins), ACE inhibitors, ARBs, and antiplatelet agents (6). There are also combinations of these drugs, primarily antihypertensives and lipid modifying drugs, known as the polypill which is gaining increasing recognition and use (7). Lifestyle modification, primarily in the form of reduced caloric intake and increased physical activity, is still the chief preventative measure or even treatment for less severe MetS or MetS-related conditions because this has been largely shown to improve metabolic outcomes (8,9). Some specific types of diet have likewise been investigated for this purpose, along with the role of dietary fats and oils (10,11). Furthermore, in light of personalised treatment approach, physical exercise prescription like a drug's dose has also been evaluated and suggested (12). Other less common studies and implicated therapies are on dietary supplements and nutritional anti-inflammatories in the prevention and treatment of MetS and T2D (13). Bariatric surgery (BS), also known as metabolic surgery, has also been increasingly practiced over the past decades primarily to curb morbid obesity along with its comorbidities (14). A recent (2019) GWAS on MetS was performed from the UK Biobank wherein 93 independent loci have been identified to be associated with MetS, 80 of which were novel. (15) Seven genes of particular interest were noted: five (WDR48, KLF14, NAADL1, GADD45G, and OR5R1) were not previously associated with any MetS component; and two (SNX10, C5orf67) were associated with all five MetS components in previous studies- these genes may be investigated further in subgroup analyses in this study. Additionally, the

	calculation method for serum glucose, HDL-cholesterol, and triglycerides used to adjust for the differences in participants' fasting time will be adapted. The definition of MetS utilized the NCEP-criteria, where three of the following criteria must be met: (1) blood pressure ≥130/85 mmHg or antihypertensive treatment; (2) serum glucose ≥6.1 mmol/L or antidiabetic treatment; (3) serum triglycerides ≥1.7 mmol/L; (4) waist circumference >102 cm in men and >88 cm in women; (5) HDL-cholesterol 1.0 mmol/L in men and <1.3 mmol/L in women. Owing to the NCEP-criteria definition of MetS as a binary trait, it will also be used as the definition of MetS in this study. In summary, based on current evidence, research studies are focused on MetS individual components and not as a broader entity. There is a huge scarcity of inquiries regarding examining this disease condition as a whole rather than separately. In addition, T2D and CVDs have undoubtedly a significant worldwide impact on mortality, and health and economic burden. Hence a study on the overlap of genetics and T2D and CVD progression warrants significant contribution to the call for further much-needed breakthroughs on this condition. Individuals with conditions known to affect obesity and T2D other than MetS-associated genes will not be included, as well as those who are metabolically or immuno-compromised or with serious illnesses. Confounding factors such as pharmacotherapy, other interventions (e.g. BS), physical activity level, and diet will be considered. Type of study: Retrospective cohort based on UK Biobank data Study Group: Primary Investigator: Esphie Grace Fojas Adviser 1: Roozbeh Naemi, PhD (Staffordhire University) Adviser 2: Samina Naseeb, PhD (Staffordshire University) Inclusion Criteria: - Participants in the UK Biobank (Age: 40-69 years old) - With MetS-associated genotype Exclusion Criteria: - Monogenic forms of obesity, diabetes or hyperlipidemia - Inborn error of metabolism - Type 1 diabetes - Hepatic or renal failure/disease - Serious illness (e.g. cancer) - History of drug and/or alcohol abuse Endpoints: 1) Incidence of T2D 2) Incidence of cardiovascular endpoints e.g. MACE METHODS: A bespoke dataset for this research project will be created based on the instructions for users and coordination with UK Biobank administration. All information will be derived from the UK Biobank, and no other external data collection methods will be used in this study. Data extraction will include:

	(1) Participant demographics which will include age, sex, ethnicity, and anthropometric measurements (e.g. weight, height, BMI, waist-to-hip ratio) among others; smoking status, alcohol use, as well as diabetes status (2) Biomarkers for cardiovascular (e.g. AHA risk components), diabetes, renal, and liver; and BP measurement (3) Individuals with MetS-associated genes and their genotype (4) Physical activity, diet profile, digestive health, and food preferences (5) Ethnicity and socioeconomic status (where available) (6) Medication for treatment of single or combination of MetS components (7) Bariatric surgery details, if any (8) Health-related outcome, particularly first occurrence of diabetes and/or CVD Data preparation: Excel will be used to organise and code the data, as necessary, derived from the UK Biobank. Subgroup data sets may also be prepared. Data screening: SPSS will be used to screen the data. Details to check for assumptions are as follows: 1) Accuracy- Descriptive Statistics and Frequencies 2) Missing data- Missing Value Analysis 3) Outliers- Explore (z-scores, histogram, box plot) 4) Normality- Explore (normality plots with tests) 5) Homogeneity of variance- Graph (scatter-plot) 6) Linearity- Graph (scatter-plot) Statistical power: Since this is a retrospective study, the power of the study may be computed after an established number of research subjects have been confirmed from the UK Biobank, based on previous similar MetS or chronic disease studies in the UK population. This may be used as part of the results of the study. Data analyses proper: Master data sheet will be prepared and finalized. Exploratory sequential mixed method type may be used for Lifestyle: Quantitative equivalent of qualitative diet and physical activity data may be possible based on the type of data retrieved (e.g. physical activity levels can be translated as low, moderate, or high); how these were measured and reported will be analysed to conclude reliability. Quantitative analyses: SPSS will be used for quantitative data analysis. Descriptive statistics 1) Participant demographics 2) Biomarkers and BP measurement 3) MetS gene mutation types 4) Activity levels and diet profile 5) Types of medication Correlation: The use of Pearson or Spearman correlation will be checked to assess the

	association between the following parameters: - MetS mutation - FPG, RBS, triglycerides, HDL-c, BP - Diet and physical activity - Medication types and outcomes - T2D and CVD onset Non-parametric analyses: When assumptions are not met, non-parametric analyses (e.g. Mann-Whitney U, Wilcoxon signed rank, Kruskal-Wallis, Friedman) will be used. Qualitative analyses: *This may be used primarily for the qualitative data which will be derived from the output on interview-based diet questionnaire as well as the physical activity data depending on the type of results retrieved from the UK Biobank. Ethical consideration: As this project will be solely based on the UK Biobank and corresponding approvals have been sought and have now been duly approved, minimal potential ethical concerns are expected. Timescale: Data request / retrieval / organization / preliminary analysis: 6 months Data analysis and interpretation: 6 months Manuscript preparation: 9 months Thesis Defence Expected outcome/s: At least one publication to a high impact journal is expected as an outcome of this study. 9) References: (1) The IDF consensus worldwide definition of the METABOLIC SYNDROME. (2006). International Diabetes Federation. (2) O'Neill, S. and O'Driscoll, L. (2015). Metabolic syndrome: a closer look at the growing epidemic and its associated pathologies. Obesity Reviews, 16, 1-12. (3) Saklayen, M.G. (2018). The Global Epidemic of the Metabolic Syndrome. Current Hypertension Reports, 20: 12. (4) Misra, A and Halcox J. (2015). Type 2 diabetes mellitus, metabolism, and mixed dyslipidemia: how similar, how different, and how to treat? Metabolic Syndrome and Related Disorders. Feb; 13(1): 1-21. (5) Marott S.C, Nordestgaard BG, Tybjaerg-Hansen A, and Benn M. (2016). Components of Metabolic Syndrome and Type 2 Diabetes. The Journal of Clinical Endrocinology and Metabolism, Aug;101(8): 3212-21. (6) Wang, S. S. (2017, October 17). Metabolic Syndrome Medication. Retrieved from https://emedicine.medscape.com/article/165124-medication#showall (7) Rosolova H. Metabolic syndrome and the role of the polypill in the prevention of its complications. (2017). Vinitrni Lekarstvi. Fall 63(10):712-716. (8) De Sousa S.M. and Norman R.J. (2016) Metabolic syndrome, diet and exercise. Best Practice & Research: Clinical Obstetrics & Gynaecology. Nov;37:140-151.

	(9) M Hernandez Ruiz de Eguilaz M.A., Batlle M.A., Martinez de Morentin B., San-Cristobal R., et al. (2016). Alimentary and lifestyle changes as a strategy in the prevention of metabolic syndrome and diabetes mellitus type 2: milestones and perspectives. Anales del Sistema Sanitario de Navarra. May-Aug 39(2); 269-89. (10) Veissi M., Anari R., Amani R., Shahbazian H., and Latifi S.M. (2016). Mediterranean diet and metabolic syndrome prevalence in type 2 diabetes patients in Ahvaz, southwest of Iran. Diabetes and Metabolic Syndrome. Apr-Jun; 10(2 Suppl 1):S26-9. (11) Misra A., Sighal N., Khurana L. (2010). Obesity, the metabolic syndrome, and type 2 diabetes in developing countries: role of dietary fats and oils. Journal of the American College of Nutrition. Jun;29(3 Suppl):289S-301S. (12) Stefani L., and Galanti G. (2017). Physical Exercise Prescription in Metabolic Chronic Disease. Advances in Experimetnal Medicine and Biology. 1005:123-141. (13) Merone L. and McDermott R. (2017). Nutritional anti-inflammatories in the treatment and prevention of type 2 diabetes mellitus and the metabolic syndrome. Diabetes Research and Clinical Practice. May;127:238-253. (14) Genser L., Casella Mariolo J.R., Castagneto-Gissey L., et al. (2016). Obesity, Type 2 Diabetes, and the Metabolic Syndrome: Pathophysiologic Relationships and Guidelines for Surgical Intervention. Surgical Clinics of North America. Aug;96(4):681-701. (15) Lind L. (2019). Genome-Wide Association Study of the Metabolic Syndrome in UK Biobank. Metabolic Syndrome and Related Disorders. Vol. XX, Number XX:1-7.
Give a brief description of participants and procedure (methods, tests etc.)	Research Subjects: Dataset for participants with MetS-associated genotype will be extracted from the UK Biobank. A bespoke dataset for this research project will be created based on the instructions for users and coordination with UK Biobank administration. All information will be derived from the UK Biobank, and no other external data collection methods will be used in this study. Data extraction will include: (1) Participant demographics which will include age, sex, ethnicity, and anthropometric measurements (e.g. weight, height, BMI, waist-to-hip ratio) among others; smoking status, alcohol use, as well as diabetes status (2) Biomarkers for cardiovascular (e.g. AHA risk components), diabetes, renal, and liver; and BP measurement (3) Individuals with MetS-associated genes and their genotype (4) Physical activity, diet profile, digestive health, and food preferences (5) Ethnicity and socioeconomic status (where available) (6) Medication for treatment of single or combination of MetS components (7) Bariatric surgery details, if any (8) Health-related outcome, particularly first occurrence of diabetes and/or CVD

	Data preparation: Excel will be used to organise and code the data, as necessary, derived from the UK Biobank. Subgroup data sets may also be prepared. Data screening: SPSS will be used to screen the data. Details to check for assumptions are as follows: 1) Accuracy- Descriptive Statistics and Frequencies 2) Missing data- Missing Value Analysis 3) Outliers- Explore (z-scores, histogram, box plot) 4) Normality- Explore (normality plots with tests) 5) Homogeneity of variance- Graph (scatter-plot) 6) Linearity- Graph (scatter-plot) Statistical power: Since this is a retrospective study, the power of the study may be computed after an established number of research subjects have been confirmed from the UK Biobank, based on previous similar MetS or chronic disease studies in the UK population. This may be used as part of the results of the study. Data analyses proper: Master data sheet will be prepared and finalized. Exploratory sequential mixed method type may be used for Lifestyle: Quantitative equivalent of qualitative diet and physical activity data may be possible based on the type of data retrieved (e.g. physical activity levels can be translated as low, moderate, or high); how these were measured and reported will be analysed to conclude reliability. Quantitative analyses: SPSS will be used for quantitative data analysis. Descriptive statistics 1) Participant demographics 2) Biomarkers and BP measurement 3) MetS gene mutation types 4) Activity levels and diet profile 5) Types of medication Correlation: The use of Pearson or Spearman correlation will be checked to assess the association between the following parameters: - MetS mutation - FPG, RBS, triglycerides, HDL-c, BP - Diet and physical activity - Medication types and outcomes - T2D and CVD onset Non-parametric analyses: When assumptions are not met, non-parametric analyses (e.g. Mann-Whitney U, Wilcoxon signed rank, Kruskal-Wallis, Friedman) will be used. Qualitative analyses: *This may be used primarily for the qualitative data which will be derived from

	the output on interview-based diet questionnaire as well as the physical activity data depending on the type of results retrieved from the UK Biobank. Ethical consideration: As this project will be solely based on the UK Biobank and corresponding approvals have been sought and have now been duly approved, minimal potential ethical concerns are expected. Timescale: Data request / retrieval / organization / preliminary analysis: 6 months Data analysis and interpretation: 6 months Manuscript preparation: 9 months Thesis Defence Expected outcome/s: At least one publication to a high impact journal is expected as an outcome of this study.		
Expected Start Date:	January 2022	Expected End Date:	November 2023

Relevant professional body ethical guidelines should be consulted when completing this form.

Please seek guidance from the School Ethics Coordinator if you are uncertain about any ethical issues arising from this application.

There is an obligation on the researcher and supervisor (where applicable) to bring to the attention of the School Ethics Coordinator any issues with ethical implications not identified by this form.

Researcher Declaration

I consider that this project has no significant ethical implications requiring full ethical review	☒

I confirm that:		
1.	The research will **NOT** involve members of vulnerable groups. Vulnerable groups include but are not limited to: children and young people (under 18 years of age), those with a learning disability or cognitive impairment, patients, people in custody, people engaged in illegal activities (e.g. drug taking), or individuals in a dependent or unequal relationship.	☒
2.	The research will **NOT** involve sensitive topics. Sensitive topics include, but are not limited to: participants' sexual behaviour, their illegal or political behaviour, their experience of violence, their abuse or exploitation, their mental health, their gender or ethnic status. The research must not involve groups where permission of a gatekeeper is normally required for initial access to members, for example, ethnic or cultural groups, native peoples or indigenous communities.	☒
3.	The research will **NOT** deliberately mislead participants in any way.	☒
4.	The research will **NOT** involve access to records of personal or confidential information, including genetic or other biological information, concerning identifiable individuals.	☒
5.	The research will **NOT** induce psychological stress, anxiety or humiliation, cause more than minimal pain, or involve intrusive interventions. This includes, but is not limited to: the administration of drugs or other substances,	☒

	vigorous physical exercise, or techniques such as hypnotherapy which may cause participants to reveal information which could cause concern, in the course of their everyday life.	
6.	The research **WILL** be conducted with participants' full and informed consent at the time the study is carried out: • The main procedure will be explained to participants in advance, so that they are informed about what to expect. ☐ • Participants will be told their involvement in the research is voluntary. ☐ • Written consent will be obtained from participants. *(This is not required for self-completion questionnaires as submission of the completed questionnaire implies consent to participate).* ☐ • Participants will be informed about how they may withdraw from the research at any time and for any reason. ☐ • For questionnaires and interviews: Participants will be given the option of omitting questions they do not want to answer. ☐ • Participants will be told that their data will be treated with full confidentiality and that, if published, every effort will be made to ensure it will not be identifiable as theirs. ☐ • Participants will be given the opportunity to be debriefed i.e. to find out more about the study and its results. ☐	YES ☐ N/A ☒
7.	A risk assessment has been completed for this research project	YES ☐ N/A ☒

If you are unable to confirm any of the above statements, please complete a **Full Ethical Review Form**. If the research will include participants that are **patients,** please complete the Independent Peer Review process.

8. Information and Data Please provide answers to the following questions regarding the handling and storage of information and data:
a) How will research data be stored (manually or electronically)? **The research data will be stored electronically in a password-protected computer.**
b) How is protection given to the participants (e.g. by being made anonymous through coding and with a participant identifier code being kept separately and securely)? **This is a secondary anonymised data that will be obtained from the UK Biobank. This anonymised data is irreversible when given or used by researchers who are able to download the encrypted data when application has been approved. Data sets are password-protected individually, and the passwords are being sent separately by the UK Biobank.** **Only the UK Biobank retains an internal database which may enable reversing of the anonymised data.**

c) What assurance will be given to the participant about the confidentiality of this data and the security of its storage? Aside from the irreversible anonymisation of the data, researchers applying for data access have to meet requirements of the UK Biobank's Access Procedures overseen by the Access Sub-Committee and have to enter a legally-binding contract, the Material Transfer Agreement (MTA). Furthermore, the researchers are not entitled to republish or make available any data other than as part of researcher results. In addition, very limited number of UK Biobank staff have access to the systems which are also subject to internal and external audits. Under the MTA, security sytems of researchers may also be audited.
d) Is assurance given to the participant that they cannot be identified from any publication or dissemination of the results of the project? Assurance is given to participants that they cannot be identified from any publication or dissemination; this is specified in the MTA and explicitly stated in the UK Biobank Data Management and Sharing Plan section 1 ("A researcher is not entitled to publish any material which could lead to the identification (inadvertent or otherwise) of an individual.").
e) Who will have access to this data, and for what purposes? Myself as student and my supervisors- Prof. Roozbeh Naemi and Dr. Samina Naseeb will have access to the data purely for the purpose of my thesis. Prof. Naemi is my main thesis supervisor, and Dr. Naseeb will be involved primarily for supervision and advice on the genetics aspect of the study.
f) How will the data be stored, for how long, and how will it be discarded? The UK Biobank, as specified on the Data Management and Sharing Plan section 13 (Data preservation), does not proscribe a time limit to when researchers may retain or discard the data. This is supplementary to section 7 (data generated by UK Biobank and researchers-ownership) provisions stating that "The researcher owns the Researcher Analyses and Researcher Results, subject to providing such information back to UK Biobank...".

Supporting Documentation

All key documents e.g. consent form, information sheet, questionnaire/interview schedule are appended to this application.	☒

Signature of Researcher:		Date:	31 January 2022

NB: If the research departs from the protocol which provides the basis for this proportionate review, then further review will be required and the applicant and supervisor(s) should consider whether or not the proportionate review remains appropriate. If it is no longer appropriate a full ethical review form **MUST** be submitted for consideration by the School Ethics Coordinator .

Next Step:

STUDENTS: Please submit this form (and supporting documentation) for consideration by your Supervisor/ Module Tutor.

STAFF: Please submit this form to your Head of Department or a Senior Researcher in your School. Once they have reviewed the form, this should be forwarded to the Research Administrators in RIIS (ethics@staffs.ac.uk) who will arrange for it to be considered by an independent member of the School's College of Reviewers .

PART B: TO BE COMPLETED BY SUPERVISOR/MODULE TUTOR (If student) OR Head of Department/ Senior Researcher (if staff)

I consider that this project has no significant ethical implications requiring full ethical review by the Faculty Research Ethics Committee.	☒
I have checked and approved the key documents required for this proposal (e.g. consent form, information sheet, questionnaire, interview schedule).	☒

Signature of Supervisor/ Head of Department/ Senior Researcher:	R.Naemi	Date:	31.01.2022

Next Step: Please forward this form to the Research Administrators in RIIS (ethics@staffs.ac.uk) who will arrange for it to be considered by an independent member of the School's College of Ethical Reviewers , having no direct connection with the researcher or his/her programme of study.

PART C: TO BE COMPLETED BY A MEMBER OF THE SCHOOL'S COLLEGE OF ETHICAL REVIEWERS

This research proposal has been considered using agreed University Procedures and is now approved.	☐
Or	
This research proposal has not been approved due to the reasons given below.	☐
Recommendation (delete as appropriate): Approve/ Amendments required/ Reject	

Name of Reviewer:		Date:	
Signature:			

Signed (School Ethical Coordinator)		Date:	