ASSOCIATING DIET, CELL PROFILING, AND BIOMARKER DATASETS FROM UK BIOBANK TO UNDERSTAND THE TRAJECTORY OF NAFLD PATIENTS

By

GENEVIEVE MONAGHAN

A thesis submitted to the University of Birmingham for the degree of

MSC BY RESEARCH

Institute of Cancer and Genomic Sciences

College of Medical and Dental Sciences

University of Birmingham

September 2022

Abstract

Introduction: Non-alcoholic fatty liver disease (NAFLD) is a potentially serious liver disease that affects approximately one-quarter of the global adult population, becoming the leading global cause of chronic liver disease in the past few decades. Despite its growing prevalence, the underlying biological mechanisms are poorly understood with lifestyle factors such as nutrition and physical activity currently playing a key role as main therapeutic strategies.

Objective: To study the associations between nutritional and biological factors in NAFLD participants using UK Biobank.

Methods: In this study, we used participants diagnosed with NAFLD and their biomarker, whole blood count, clinical, and dietary data to find associations with disease course. NAFLD participants were divided into three categories by date of NAFLD diagnosis and matched controls included. Linear, logistic, and ordinal regression were used to assess the associations between datasets. Analysis of variance and Tukey's honestly significant difference tests were used to identify significant (corrected P value < 0.05) differences between cohorts.

Results: 990 associations were identified, of which 13 biomarker to whole blood count feature associations were mutual between the four cohorts. The control cohort and the longest NAFLD suffering participants shared the least associations. Of the 61 features examined, 53 had at least one significant difference between two cohorts.

Conclusions: Certain biomarkers can be used as predictors of whether someone will be diagnosed with NAFLD – triglycerides > 1.7 mmol/L, high density lipoprotein cholesterol < 1.39 mmol/L, gamma glutamyl transferase > 40 U/L, and aspartate aminotransferase > 33 U/L. Increased fibre intake is associated with a decrease in metabolic hallmarks of NAFLD.

Acknowledgements

We would like to thank Dr Animesh Acharjee and Prof Georgios Gkoutos for their supervision of this project, and Dr Animesh Acharjee for the management of the project and his contribution to the conceptualisation and strategy design of the project. We would also like to thank Laura Bravo for her systematic collection of the UK Biobank data.

Contents

List of Figures

List of Tables

Abbreviations

Abbreviation	Full name
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
ANOVA	Analysis of variance
Apo-A	Apolipoprotein A
Apo-B	Apolipoprotein B
AST	Aspartate aminotransferase
BMI	Body mass index
CI	Confidence intervals
CKD	Chronic kidney disease
CRP	C reactive protein
CVD	Cardiovascular disease
D2AF	Patients diagnosed more than 2 years after biomarker data collection
D2BF	Patients diagnosed more than 2 years before biomarker data collection
DW2	Patients diagnosed within +-2 years of biomarker data collection
FFA	Free fatty acids
GFR	Glomerular filtration rate
GGT	Gamma glutamyl transferase
HCC	Hepatocellular carcinoma
HDL-C	High density lipoprotein cholesterol
HLSRC	High light scatter reticulocyte count
HLSRP	High light scatter reticulocyte percentage
HSD	Honestly significant difference
LDL-C	Low density lipoprotein cholesterol
MCHC	Mean corpuscular haemoglobin concentration
MCV	Mean corpuscular volume of red blood cells
NAFLD	Non-alcoholic fatty liver disease
NASH	Non-alcoholic steatohepatitis
RBC	Red blood cell

RDW	Red blood cell distribution width
ROS	Reactive oxygen species
RPR	Red blood cell distribution width to platelet ratio
SCFA	Short chain fatty acids
SD	Standard deviation
SHBG	Sex hormone binding globulin
T2D	Type II diabetes mellitus
TC	Total cholesterol
TG	Triglycerides
Th17	IL-17 producing T helper cell
TLR	Toll-like receptor
Treg	Regulatory T cell
WBC	White blood cell

Chapter 1. Introduction

1.1 Definition of Non-alcoholic fatty liver disease

The liver performs a multitude of essential biological functions from detoxification to regulation of systemic metabolism, including the control of blood glucose and lipid metabolism (1, 2). Loss of liver function, as occurs with both acute and chronic liver injury, leads to the dysfunction and dysequilibrium of whole body metabolism and can lead to death (3, 4). Non-alcoholic fatty liver disease (NAFLD) is the most prevalent chronic liver disease in the world, currently affecting approximately one-quarter of the global adult population and predicted to increase by 21% in the next decade (2, 3). NAFLD is characterized by steatosis, an excess accumulation of hepatic fat (>5% fat content in the liver) in those who do not consume an excess of alcohol, and can progress into more severe diseases (2-4). More specifically, the spectrum of diseases progressing from a NAFLD diagnosis can be seen in **Figure 1**, with 12-40% of NAFLD patients developing non-alcoholic steatohepatitis (NASH), characterized by hepatic necroinflammation and fibrosis. From these, 15-25% go on to progress towards cirrhosis with increased predisposition for hepatocellular carcinoma (HCC) or liver failure (5-7).

NAFLD risk increases with age and body mass index (BMI), and, although the profiles of male and female livers have been established as metabolically distinct, the influence of biological sex on NAFLD prevalence and risk remains undecided (4, 8). There is elevated overall mortality risk in the NAFLD population, with cardiovascular disease (CVD) as the most common cause of death, accounting for ~13% of deaths (2-4).

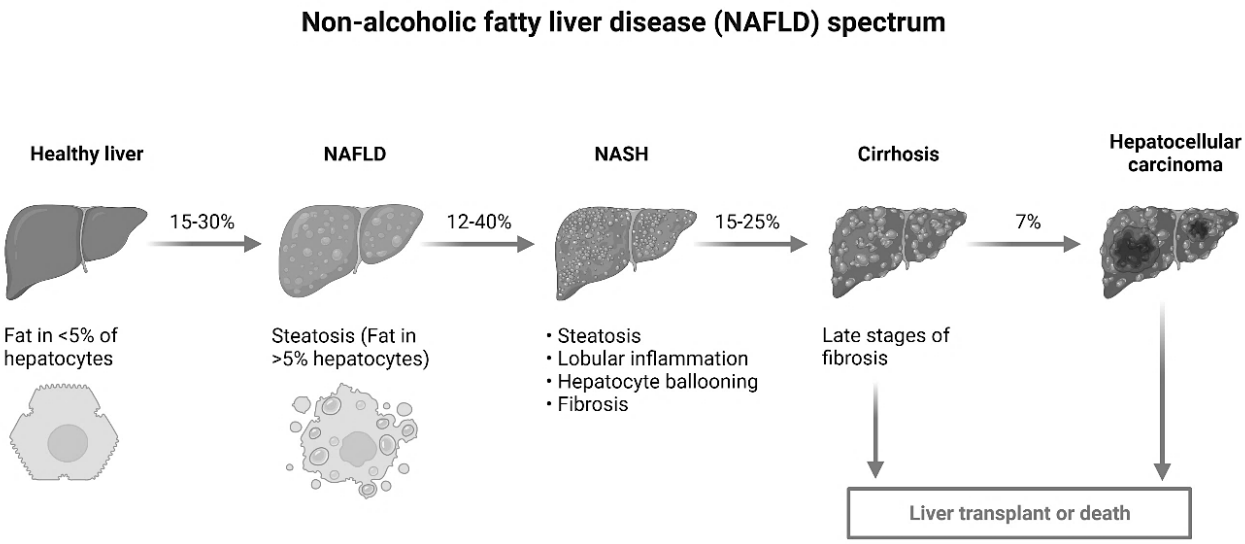


Figure 1: NAFLD Spectrum

NAFLD spectrum ranges from steatosis to, with the development of inflammation, steatohepatitis, cirrhosis, and, in a small percentage of cases, HCC. Changes to the cellular structure of hepatocytes accompanies the transition from healthy liver to steatosis with intracellular fat droplet accumulation, enlarged cells, and an altered cytoskeleton. The final stages of fatty liver disease see fibrosis of the liver and loss of function.

Abbreviations: Non-alcoholic fatty liver disease (NAFLD), non-alcoholic steatohepatitis (NASH)

Created with BioRender.com

1.2 NAFLD biomarkers

The key physiological hallmark of NAFLD is steatosis, which can be easily determined through ultrasound readings, although the gold standard for diagnosis is liver biopsy (9). Markers of metabolic dysfunction such as lipotoxicity, insulin insensitivity, and raised liver enzymes are gaining traction as key descriptors of NAFLD pathology, promising a better understanding of the underlying biological mechanisms, and offering alternatives for clinical outcome staging and progression (see **Figure 2**) (1, 6, 10, 11).

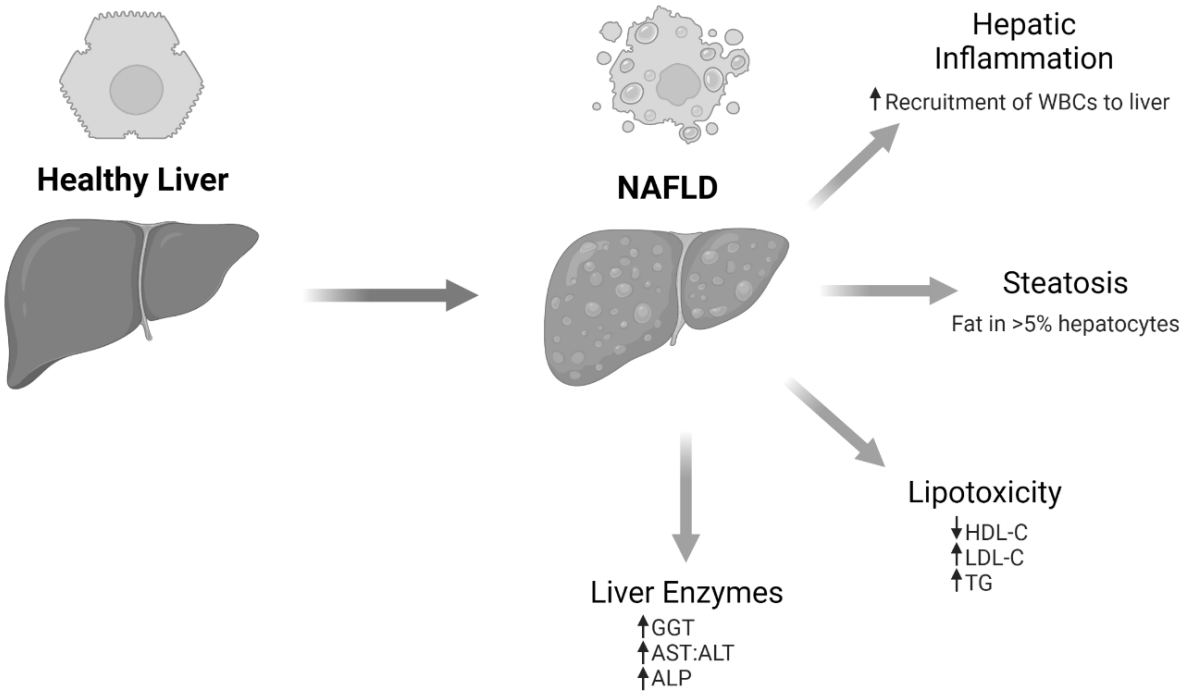


Figure 2: Cardinal features of NAFLD

The cardinal features of NAFLD compared to the function of the healthy liver. These include hepatocyte fat retention in beyond 5% of cells; immune cell mediated lobular inflammation characterised by increased hepatocyte death, influx of immune cells to the liver, and elevated inflammatory biomarkers including CRP; lipotoxicity, primarily reduced HDL-C and elevated LDL-C and TG, although other lipid species including sphingomyelins have been highlighted in relation to NAFLD-related lipotoxicity; abnormal liver enzymes, particularly GGT, AST, and ALT, although their functions within the liver remain unclear. Other common characteristics of NAFLD patients include insulin resistance and elevated bilirubin, but these features do not form part of the key diagnostic criteria.

Abbreviations: non-alcoholic fatty liver disease (NAFLD), white blood cell (WBC), high density lipoprotein cholesterol (HDL-C), low density lipoprotein cholesterol (LDL-C), triglycerides (TG), gamma glutamyl transferase (GGT), alanine transaminase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), C reactive protein (CRP)

Created with BioRender.com

1.2.1 Liver enzymes

Elevated levels of liver enzymes alanine aminotransferase (ALT), gamma glutamyl transferase (GGT), aspartate aminotransferase (AST), and alkaline phosphatase (ALP) have been associated with NAFLD pathogenesis and progression (12-15).

GGT elevation is indicative of liver or bile duct damage and as such GGT levels are one of the key clinical biomarkers used in diagnosis (13-15). Although the biochemical mechanics of GGT action within the liver are not fully understood at present, elevated GGT in NAFLD patients has been associated with increased mortality and further progression along the NAFLD spectrum due to its close correlation with fibrosis staging (4, 12). The use of GGT alone to confirm NAFLD is not advised as although GGT is predominantly expressed along the hepatobiliary tract, it is also expressed in multiple other organs (13, 14, 16).

In addition to GGT, AST and ALT measurements have been used both separately and in conjunction with each other for the study of NAFLD as the two enzymes play similar roles within hepatic catabolism (15, 16). Both enzymes have been observed to be elevated in NAFLD, although there is debate in the literature as to how the alterations to these two enzymes at the different stages of the NAFDL spectrum are reflected in the AST:ALT ratio in NAFLD patients (15, 17, 18). However, blood AST levels may also be influenced by injury to other organs, such as the heart, so ALT measurements are more indicative of liver specific damage (15, 16, 18). Regardless, fluctuations in AST and ALT values are associated with circadian rhythm, and so, whilst ALT measurements may be more liver specific than AST measurements, spot

measurements of both enzymes may not be a reliable representation of liver injury as a whole (15, 16, 18, 19). ALP, a further enzyme elevated in NAFLD, can be used in combination with AST and ALT to confirm liver dysfunction, due to its role in bile production (16, 17). However, standalone use of ALP is not recommended due to the distribution of tissue-nonspecific isoenzymes of ALP throughout the body (16). Although no one specific liver-expressed enzyme has been identified as a reliable solitary marker for NAFLD, the use of these enzymes in combination with each other provides the most dependable identification of chronic liver damage (15).

1.2.2 Lipotoxicity

Lipotoxicity is crucial in the aetiology and progression of NAFLD (2-4). Hepatic lipotoxicity happens when the capacity of the liver to store and metabolise lipids is overwhelmed (1). It plays a major role by instigating inflammation, hepatic steatosis, and insulin resistance (20-23). Lipotoxicity is associated with increased accumulation of total cholesterol (TC), low density lipoprotein cholesterol (LDL-C), triacylglycerols, sphingomyelins, and triglycerides (TG) with decreased accumulation of high density lipoprotein cholesterol (HDL-C) (see **Figure 2**) (20-26). This aberrant lipid profile is also a hallmark of metabolic syndrome and CVD which are common comorbidities associated with NAFLD (23). The sustaining of de novo lipogenesis through diet has been shown to hasten the progression of the clinically advanced stages of NAFLD (from NASH to HCC) through mediation of cell stress (24, 25). Lipid-mediated cell stress increases the susceptibility of hepatocytes to apoptosis, aggravating liver injury and fibrosis with major NAFLD risk factor genes such as *MBOAT7* and *PNPLA3* relating to control of the liver lipidome (24, 25, 27).

1.2.3 Hepatic Metabolic Dysfunction

Further metabolic hallmarks of NAFLD include increased glycated haemoglobin (HbA1c) and bilirubin (see **Figure 2**) (4, 16, 17). Decreased insulin sensitivity accompanies NAFLD, observed through elevated HbA1c, even when associated with normal glucose tolerance (1, 4, 17, 28). Reduced total bilirubin forms part of the cholestatic pattern of liver and bile duct injury (16, 29). As the end product of red blood cell (RBC) breakdown in the liver, aberrations to liver metabolism are reflected in alterations to the balance of unconjugated (indirect) bilirubin and conjugated (direct) bilirubin (16, 29). This decrease in serum bilirubin is a common part of the NAFLD diagnosis testing, and yet, the causal relationship between NAFLD and serum bilirubin is a subject of debate in NAFLD research (16, 29, 30).

1.3 Hepatic Inflammation in NAFLD

Hepatic inflammation is a pathogenic feature of chronic liver injury and a key driver of NAFLD progression into NASH, observed both through increased inflammatory markers, such as C reactive protein (CRP), and upregulated recruitment of pro-inflammatory immune cells to the liver (17, 31). The secretion of chemokines by hepatic cells following injury direct the migration and infiltration of various immune cell populations, including neutrophils, monocytes, monocyte-derived macrophages, and T lymphocytes (31).

The innate immune response has been the primary focus of cell driven NAFLD inflammation research (32-34). Liver injury stimulates the liver-resident macrophages

(Kupffer cells) which initiate the inflammatory process and, in turn, recruit circulating monocytes, mediated by the release of pro-inflammatory cytokines and chemokines (32-34). Pro-inflammatory signals linked to gut dysbiosis, owing to increased intestinal permeability in NAFLD, are the main factors mediating macrophage activation in NAFLD (32, 33). Both the liver-resident Kupffer cells and circulating monocytes, which undergo differentiation towards a pro-inflammatory phenotype, release pro-inflammatory signals, including cytokines, chemokines, and reactive oxygen species (ROS) to promote apoptosis and necrosis of hepatocytes (32, 34). This increase in cell death further stimulates the innate inflammatory response, exacerbating the NAFLD-induced damage to the liver, and in this way, Kupffer cells are critical mediators of NAFLD development and progression (33, 34).

Despite their role as the primary responders to acute inflammation, the role of neutrophils in chronic hepatic inflammation is less well understood – although neutrophil derived microRNAs have been shown to inhibit the NLRP3 inflammasome, an important contributor to the development of liver fibrosis, and increased neutrophil infiltration is closely correlated with development and severity of NASH (35, 36). Neutrophil infiltration in NAFLD, combined with the production of ROS, activates ASK-p38 apoptosis signalling pathway in hepatocytes, promoting hepatocyte death and fibrosis (34, 35). This corresponds to the increased hepatic expression of p38 in NAFLD patients (35).

The emerging role of the adaptive immune response in addition to the innate immune response in cell driven NAFLD hepatic inflammation indicates the key roles of both T and B lymphocytes (37, 38). There is variation in responses of different T lymphocytes

subtypes with a loss of CD4+ (T helper) and regulatory (Treg) T lymphocytes, but not CD8+ (cytotoxic) and IL-17 producing CD4+ (Th17) T lymphocytes in NAFLD (37, 39, 40). This disparity in T lymphocytes has been linked to lipotoxicity and ROS signalling pathways as a blockade of ROS signalling pathways reverses CD4+ T lymphocyte NAFLD-induced loss and delays progression to HCC (37). The hepatocyte killing activity of CD8+ polarised T lymphocytes is enhanced in NAFLD and contributes to reduced insulin sensitivity and increased fibrosis of the liver (37, 38). The accumulation of Th17 cells has been linked to the progression of NAFLD to NASH and both the increased proliferation of Th17 cells and the expression of proinflammatory cytokine IL-17 are associated with NAFLD-related HCC (39-41). Treg cells can mitigate the response of Th17 cells and protect against hepatic inflammation, but in NAFLD there is a loss of Treg cells and a progressive increase in the Th17:Treg ratio along the NAFLD spectrum (39, 40). Whilst there is a growing body of evidence for the role of different subtypes of T lymphocyte in NAFLD pathogenesis, the role of B lymphocytes is less well understood (37, 38). In NASH, there is an influx of B lymphocytes to the liver, but whether this association is causal or consequential to the triggering of other immune cells and the resulting inflammatory environment remains unclear (38).

1.4 Comorbidities and Risk Factors of NAFLD

NAFLD is the hepatic manifestation of metabolic syndrome, and as such has been associated with obesity, dyslipidaemia, CVD, and type II diabetes mellitus (T2D) as key comorbidities and an increased risk of hypertension, CVD, and other chronic diseases (3, 20, 42-44). However, the complex links underlying diseases are difficult to establish, with for instance, a bidirectional relationship existing between T2D and

NAFLD where NAFLD can be induced by the elevated hepatic de novo lipogenesis and hypertriglyceridaemia in T2D, but also predisposes patients to T2D due to increased hepatic fat storage and insulin resistance (43, 45, 46). Also, the development of insulin resistance in T2D is aided by free fatty acids (FFA) influx and lipotoxicity, with hepatocyte ballooning and inflammation inducing reduced insulin sensitivity (43, 45). Approximately, 50% of T2D patients are also diagnosed with NAFLD and synergies between diseases are key to better understand disease course development and inform patient clinical management (43). Glycaemic deterioration associated with T2D occurs in NAFLD patients with elevated fasting blood glucose and HbA1c in NAFLD vs non-NAFLD patients (46). Patients with both T2D and NAFLD show elevated fasting blood glucose and/or elevated HbA1c compared to patients with NAFLD alone and the presence of T2D has been shown to be a useful predictor for the long term clinical outcomes of NAFLD, including NASH development and increased overall mortality (43, 45).

Obesity is also a key factor in NAFLD patients, with the increased prevalence of NAFLD associated with the rising rates of obesity (4). Metabolic hallmarks of NAFLD, such as hypercholesterolaemia and insulin resistance, are similarly associated with obesity and genetic variants linked to increased risk for NAFLD, such as *PNPLA3* variant I148M, also confer greater genetic susceptibility to obesity (47-49).

Regarding downstream complications, both NAFLD and NASH have been established as independent risk factors for severe chronic kidney disease (CKD), the gradual loss of kidney function and reduced glomerular filtration rate (GFR) (20, 42, 50, 51). As with NAFLD, CKD is associated with high expenses and poor outcomes, with the incidence

and prevalence of both diseases projected to increase in the coming years (3, 20). The biochemical mechanisms underpinning the association between CKD and NAFLD are less clear, with a multifaceted crosstalk between the liver and the kidneys with increased levels of HbA1c, ALT, and haemoglobin and decreased HDL-C in comorbid patients (52, 53). This suggests that the dysregulated lipoprotein and lipid metabolism in NAFLD may affect renal injury and that NAFLD may be aggravated in turn by CKD (52, 53).

1.5 Nutrition and Lifestyle in NAFLD

It has been suggested that lifestyle factors contribute more to NAFLD progression than genetic markers, with physical activity and nutritional management through weight loss and calorie deficit shown to reduce disease impact significantly (5, 22, 44, 54). Currently, lifestyle management acts as the only treatment course available, with no other targeted pharmacological interventions approved to date (6). Attempts at developments for therapies include bile acid receptor agonists and targeted treatments for hypercholesterolaemia, from acetyl-CoA carboxylase suppressors to peroxisome proliferator-activated receptor agonists, as there is a need for the improvement of non-invasive methods of diagnosis and treatment (55).

Regarding nutritional management, several dietary factors, such as dietary fructose, have been highlighted as key in NAFLD development (25, 52). Fructose is the primary substrate for hepatic de novo lipogenesis and is associated with the increased intestinal permeability in NAFLD, which contributes to pro-inflammatory signalling to stimulate hepatic inflammation (25, 32, 33, 56). Current research indicates that the

balance of macronutrients in a NAFLD-targeting diet should weigh low in carbohydrates, low in sugars (e.g. fructose), and high in proteins; these characteristics align with the Mediterranean diet as it is largely plant-based and low in carbohydrates, but high in essential fatty acids (25, 44, 56). The Mediterranean diet has been shown to reduce intestinal permeability, as well as lower AST and ALT levels and BMI (56). Research has expanded on the effects of plant-based nutrition by observing the effect of a vegan diet on NAFLD patients as patients displayed improved liver enzymes, independent of weight loss (10). Therefore understanding the specific epigenetic, inflammatory, and lipidomic changes induced by a certain diet and its associations to patient prognosis can improve our understanding of the disease and its management guiding the development of personalised nutrition (57).

Physical activity is inversely associated with NAFLD risk, mediated by smaller waist circumference, improved insulin sensitivity, and reduced intrahepatic lipid content (58-60). Research into physical activity in NAFLD patients in the UK Biobank supports this inverse association, even in those with a high genetic predisposition for NAFLD, and physical activity has also been found to be inversely associated with hepatic inflammation in the UK Biobank (22, 61, 62). Although an increase in physical activity is recommended as part of lifestyle treatment for NAFLD patients, the specific recommendations for the type of exercise or duration is not consistent between health bodies and studies (60, 63, 64). The combination of physical activity with nutritional management has produced the most effective reduction in NAFLD risk and NAFLD markers (63).

1.6 Genetics of NAFLD

The heterogeneity in NAFLD phenotypes may be explained by patient genetics, which play a role in NAFLD susceptibility with genetic variants driving NAFLD progression and risk (6, 10, 11, 21). The heritability of NAFLD is estimated between 20%-70% with specific variants of genes acting as major genetic determinants of NAFLD risk (6, 48). The discovery of novel loci in recent years has focused on the genes for ALT and AST – a recently identified that loss of function variant HSD17B13 reduced ALT and AST levels as well as reduced NASH risk but did not reduce NAFLD risk, suggesting the role of ALT and AST in more clinically advanced stages of the NAFLD spectrum (48, 65, 66).

A literature search was conducted on PubMed for papers including the terms 'UK,' 'biobank' and 'NAFLD' on 13/10/2021 and received 32 papers. The UK Biobank was chosen as the data source as this is the database which we have utilised in our research. Of the 32 papers, 26 were identified as using UK Biobank data and a subset of nine papers utilised UK Biobank data to investigate NAFLD genetic risk, focusing on identifying genes which confer genetic predisposition to NAFLD (see **Table 1**) (67-75). Many of the genes identified through UK Biobank affect the liver lipidome and the hepatic storage of fat (67-75).

Gene identified	Protein effected	Variants	Effect of variant	References
APOE	Apolipoprotein E	rs429358	Increased serum lipids	(67, 68)
CELSR2-PSRC1-SORT1 gene cluster		rs599839 A>G		(69)
GCKR	Glucokinase	rs1260326 P446L	Loss of function	(67, 70, 72)
GPAM	GPAT1	rs2792751		(68)
HSD17B13	17β-Hydroxysteroid dehydrogenase type 13	rs72613567	Loss of function	(71, 72)
MARC1	mARC1	rs2642442		(67)
MAU2	MAU2	rs73001065		(67)
MBOAT7	MBOAT7	rs641738 C>T	Reduced gene expression	(70, 72, 73)
PCSK7	PCSK7	rs236918	Loss of membrane transferrin receptor	(74)
PNPLA3	Adiponutrin	rs738409 C>G I148M	Increased lipid synthesis and cellular lipid accumulation	(67, 70-73)
		rs3747207		
TM6SF2	TM6SF2	rs58542926 C>T E167K	Increased hepatic retention of lipids	(67, 70-73)
TMBIM1	TMBIM1	rs2288464		(75)
		rs9389268		
TRIB1	TRIB1	rs17321515		(67)

Table 1

Genes identified from UK Biobank data as increasing risk of developing or progressing NAFLD. Each risk-conferring gene is listed with its corresponding protein and NAFLD-related variants along with the biochemical and physiological effects of these mutations, where known.

PNPLA3 I148M is the most researched, most common genetic mutation linked to NAFLD (67, 70-73). *PNPLA3* codes for Adiponutrin, a protein which promotes lipid

synthesis, in which the gain of function mutation I148M causes excessive cellular lipid accumulation (67, 70-73). *PNPLA3* variant I148M confers greater genetic susceptibility to NAFLD, independent of metabolic syndrome, and is associated with more aggressive progression of NAFLD and greater fold risk of developing fibrosis (47, 48). *PNPLA3* I148M polymorphism is also associated with increased ALT levels, liver fat accumulation, and elevated hepatic TG levels (47, 48). Genetic variants of *PNPLA3*, including rs738409, are also associated with insulin resistance, predisposition to obesity, and greater susceptibility to CKD (20, 47, 48).

1.7 Research Aims

1. To assess the significant associations between any of the biomarker, dietary, and whole blood count features and whether these associations are positive or negative
2. To evaluate the significant differences between cohorts in the levels of biomarkers and whole blood count features and whether these markers could be useful in NAFLD diagnosis

Chapter 2. Methods

In this study, we used NAFLD patient biomarker, whole blood count, clinical, and dietary data extracted from the UK biobank and used multiple statistical methods to associate the different data types.

2.1 Datasets and participants description

UK Biobank was established in 2006 and is a large-scale biomedical database with in-depth genetic and health information from around 502,000 volunteers in the UK (76). With an average age at recruitment of 57 years and ages ranging from 37 to 73, approximately 40% are multimorbid, with ~600 having NAFLD as main diagnosis, offering a rich resource of information for our studies. The phenotypic and health-related information available includes biological measurements, lifestyle indicators such as diet and nutrition, biomarkers in blood and urine, genome-wide genotype data or follow-up longitudinal information provided by linking health and medical records (76).

For our study, we selected all patients with NAFLD (ICD10 code K760), and NASH diagnosis (ICD10 code K758) as established by main ICD10 code diagnosis fields in UK Biobank (fields: 41280 and 41270) as well as healthy controls (participants with no disease diagnosis) (76). A total of 21,879 patients were included for consideration with information on clinical, biomarker (category 717), metabolomic, and nutrition data (category 1004). The extracted dataset was comprised of 644 features and after deleting features missing 30% or more of values and identifying relevant features to

our study, we were left with 91 features from UK Biobank to study (see **Appendix Table 1** for a list of all features included in analysis).

Given most data (biological, nutritional and lifestyle) has been collected at baseline and NAFLD diagnosis may precede or follow this date by as much as 10 years, participants were categorized based on the time difference between data collection and diagnosis. From previous studies, biomarker data has been shown to stay constant in a 2 year time frame and participants were divided into: patients diagnosed more than 2 years before from biomarker data collection (D2BF) (n=326), patients diagnosed within +/- 2 years of biomarker data collection (DW2) (n=450), and patients diagnosed more than 2 years after biomarker data collection (D2AF) (n=4210) (77). In total 4986 NAFLD participants were studied after excluding those diagnosed with NASH, those having missing values for more than 30% of the features studied, and those who had made a significant change to their diet in the last five years, due to illness or otherwise (see **Figure 3**). The final cohort included 12,293 participants out of the 21,879 extracted from UK Biobank (see **Appendix Figure 1**).

NAFLD participants were included in the analysis according to the exclusion criteria (n=4986) .We matched these NAFLD patients with 7307 matched controls in UK Biobank with no reported ICD10 diagnosis and repeated the analysis (see **Table 2**). Matching was done where patients were matched using the “nearest” method which utilises a greedy search to match each sample with their nearest neighbour (78). The distance was calculated using the Mahalanobis distance, which estimates the distribution closest for each point (79). This procedure was performed in R (R Core Team 2021) using the MatchIt package (80).

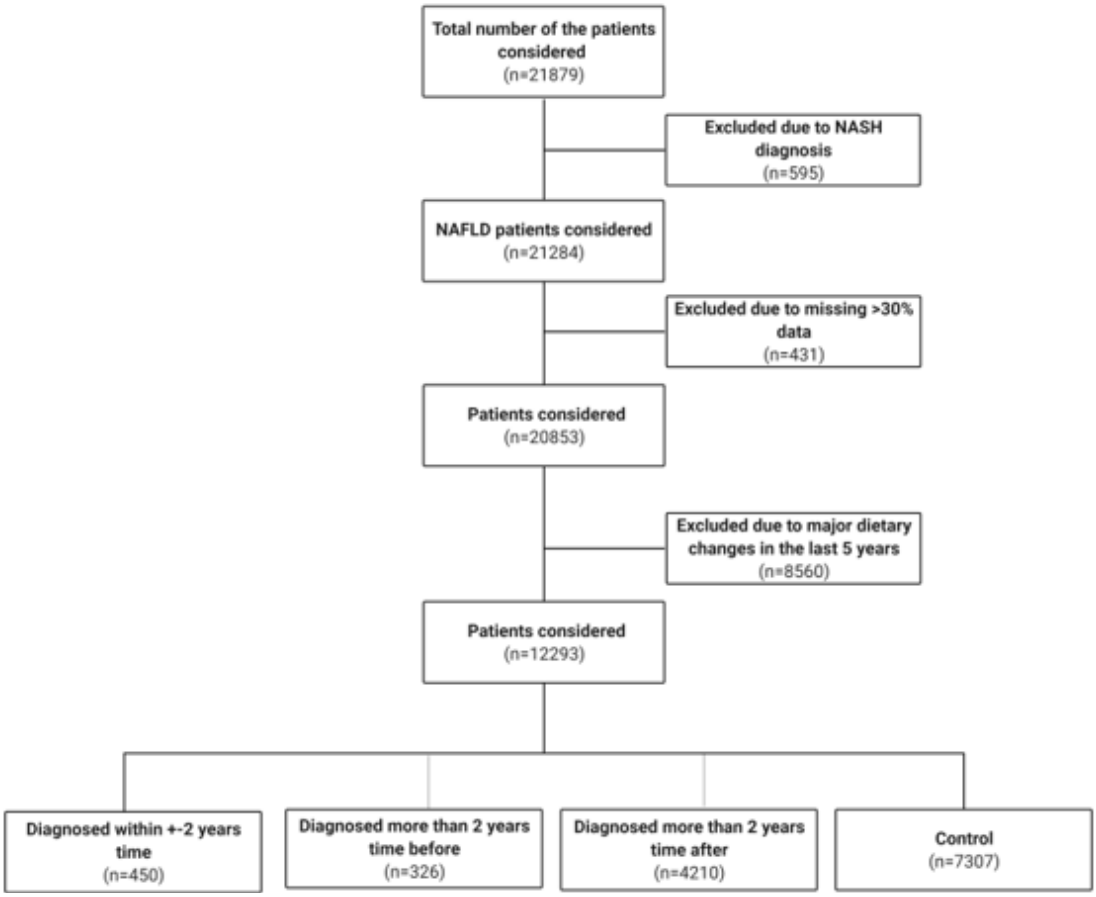


Figure 3

Workflow of the patient inclusion and exclusion in UK Biobank data analysis

	Levels	Control	NAFLD	p-value
Age at recruitment (years)	Median (IQR)	57.0 (51.0 to 62.0)	60.0 (53.0 to 65.0)	<0.001
Gender	Female	3561 (48.7)	2339 (46.9)	0.049
	Male	3746 (51.3)	2647 (53.1)	
BMI	Median (IQR)	26.5 (23.9 to 29.4)	28.4 (25.3 to 31.7)	<0.001
Time to Death (months)	Median (IQR)	114.2 (60.0 to 128.3)	96.9 (69.4 to 118.1)	0.01
Binary Death	No	7174 (98.2)	3993 (80.1)	<0.001
	Yes	133 (1.8)	993 (19.9)	
Townsend deprivation index	Median (IQR)	-2.3 (-3.8 to 0.1)	-1.6 (-3.3 to 1.5)	<0.001
Smoking status	Current	766 (10.5)	805 (16.1)	<0.001
	Never	4364 (59.7)	2305 (46.2)	
	Previous	2177 (29.8)	1876 (37.6)	
Follow up time (years)	Median (IQR)	11.9 (11.2 to 12.7)	11.7 (10.8 to 12.4)	<0.001

Table 2

Summary of the demographic information and general description of patients after exclusion criteria was applied

2.2 Ethics

The research used the UK Biobank resource (approved application number: 31224).

2.3 Data pre-processing

The final dataset contains 12,293 participants (4986 NAFLD participants and 7307 controls) and is comprised of 91 features (see **Appendix Figure 1** and **Appendix Table 1**).

Outliers were evaluated as those readings with values 4 times outside their standard deviation (SD) range and set to NA. Imputation was performed in the following way: categorical variables, such as type of cereal consumed, using the most frequent category and either the mean or median used to impute numerical values based on the data distribution. In specific cases, such as microalbumin, readings of <6.7 were imputed with 3.35 (half of the minimum sensitivity) and potassium readings of >200 were imputed with 300.

Important associations were studied between dietary features, biomarkers, and whole blood count features in the four different cohorts of participants previously mentioned based on their time of data collection in relation to their NAFLD diagnosis (Control, D2BF, DW2, and D2AF).

2.4 Univariate Regression

We have used univariate regression analysis to identify important associations between nutritional data, biomarkers, and whole blood count datasets. We estimated regression coefficients with Bonferroni corrected P values ($P<0.05$) and accounted for confounders sex, age, and BMI in our analysis. To estimate the relationship correlation between continuous numerical datasets, linear regression was used. For frequency of intake categorical dietary data, ordinal regression model was used utilising the polr function in the R package MASS. For categorical data, categorical variables were divided into their individual components for analysis and logistic regression was used. P values were adjusted using Bonferroni correction. Values are reported as (mean [95% confidence intervals (CI)]).

Network graphs were created using the ggplot2, ggraph, and igraph packages.
All codes are available at: https://github.com/genrmonaghan/NAFLD_MSc_codes .

2.5 ANOVA

One way analysis of variance (ANOVA) model (degrees of freedom = 3) was conducted with corrected P value < 0.05 to identify significantly altered features from the biomarker and whole blood count datasets in between the four patient cohorts. All assumptions of one-way ANOVA were met for all cohorts. Tukey's honestly significant difference (HSD) tests were used to find significant (corrected P value < 0.05) differences in biomarker and whole blood count levels between cohorts.

Chapter 3. Results

In this study, a total of 32 biomarkers, 29 whole blood count features, and 39 dietary features were investigated, and multiple methods were used to associate them. Linear, logistic, and ordinal univariate regression models were used to identify the significant ($P<0.05$) associations between diet, biomarker, and whole blood count features. In addition to this, ANOVA and Tukey HSD tests were conducted to find significant ($P<0.05$) differences in biomarker and cell features between cohorts.

3.1 Univariate Regression Analysis

3.1.1 D2AF Cohort

In the D2AF cohort (n=4210), 385 associations were identified (see **Appendix Figure 2a**), involving 82 unique features across the three data types, of which the most significant was a positive association between mean sphered cell volume and HDL-C ($P<0.0001$). Across the three analysis groups, AST had the highest number of total associations, accounting for 6.75% of all associations across the three analyses groups for the D2AF cohort (see **Appendix Figure 3a**).

3.1.1.1 Diet ~ Biomarkers

Univariate regression of the D2AF cohort (n = 4210) detected 70 significant diet to biomarker associations. Of the 32 biomarkers studied, 14 biomarkers were associated with cereal consumption and seven biomarkers were associated with fresh fruit intake. Cereal consumption ($P<0.05$) and fresh fruit intake ($P<0.05$) were each significantly associated with four out of the seven cholesterol-related biomarkers (see **Figure 4a**).

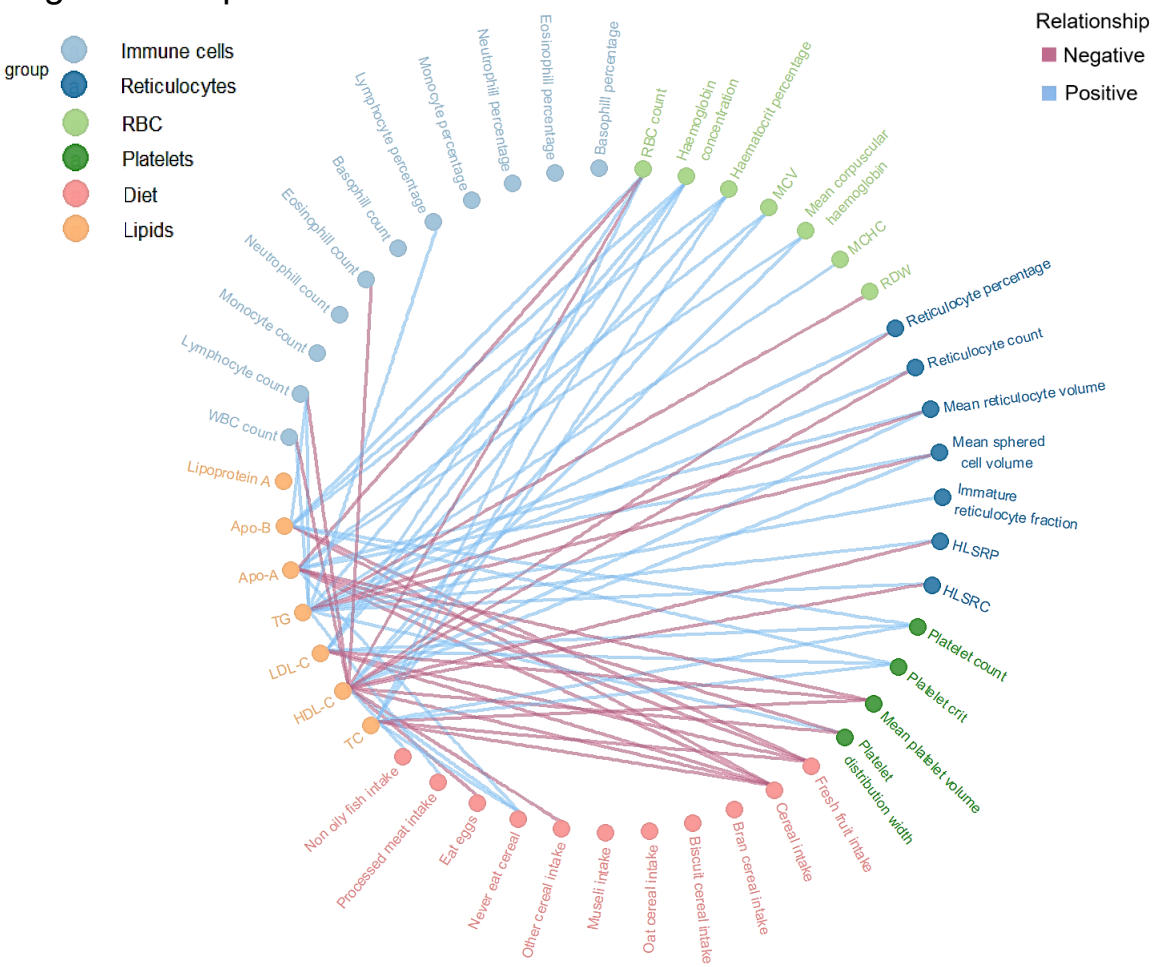
3.1.1.2 Diet ~ Whole blood count

Univariate regression of the D2AF cohort (n = 4210) identified 53 cell count to diet associations that were significant. Of the 29 whole blood count features studied, 12 were associated with cereal consumption, including all six measures of reticulocyte characteristics and function (P<0.0001). White blood cell (WBC) count had the highest number of significant associations with dietary features, followed by neutrophil count (see **Figure 4c**).

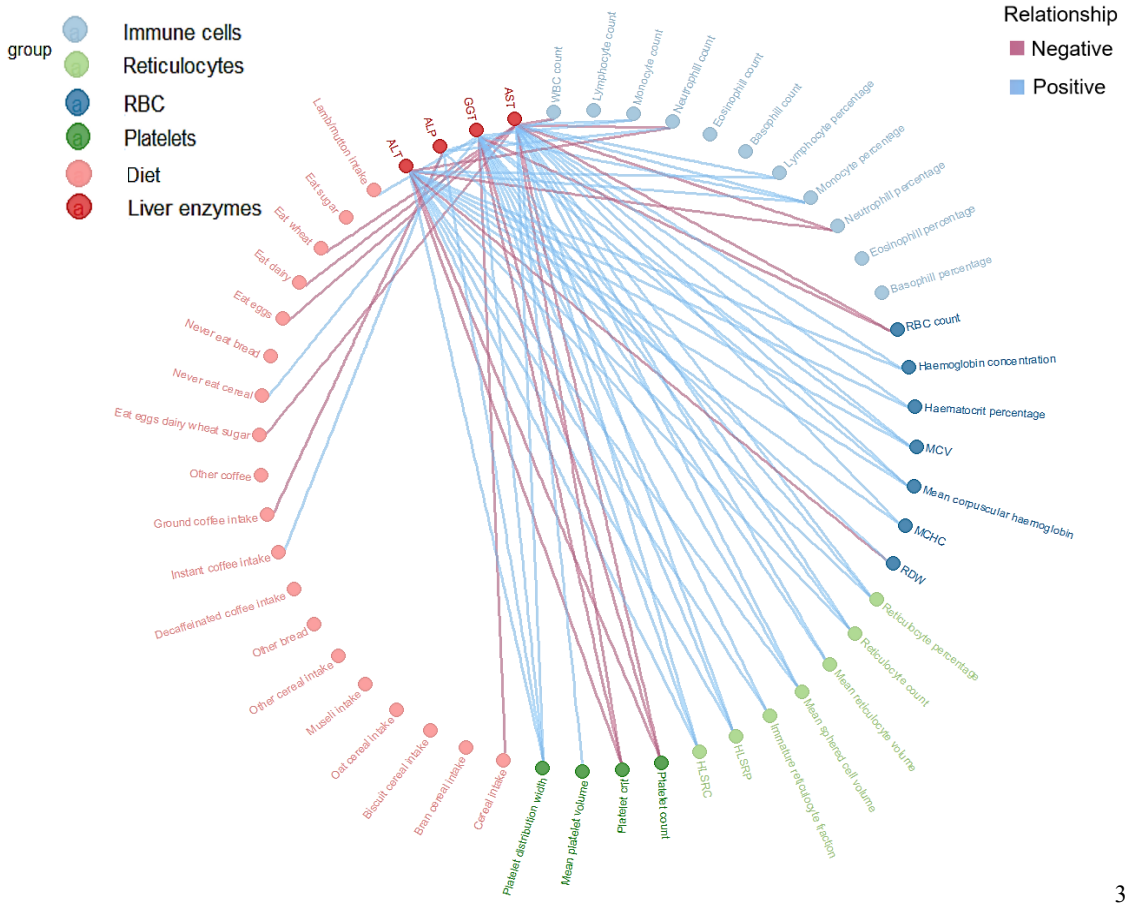
3.1.1.3 Biomarkers ~ Whole blood count

Analysis of the D2AF (n = 4210) cohort identified 262 significant biomarker to cell associations. Of these associations, 61 involved WBC features (see **Figure 4c**), relating to 19 unique biomarkers, and 60 involved liver enzymes (see **Figure 4b**), linking to 24 unique whole blood count features.

A) Significant Lipid Associations



B) Significant Liver Enzyme Associations



C) Significant Immune Cell Associations

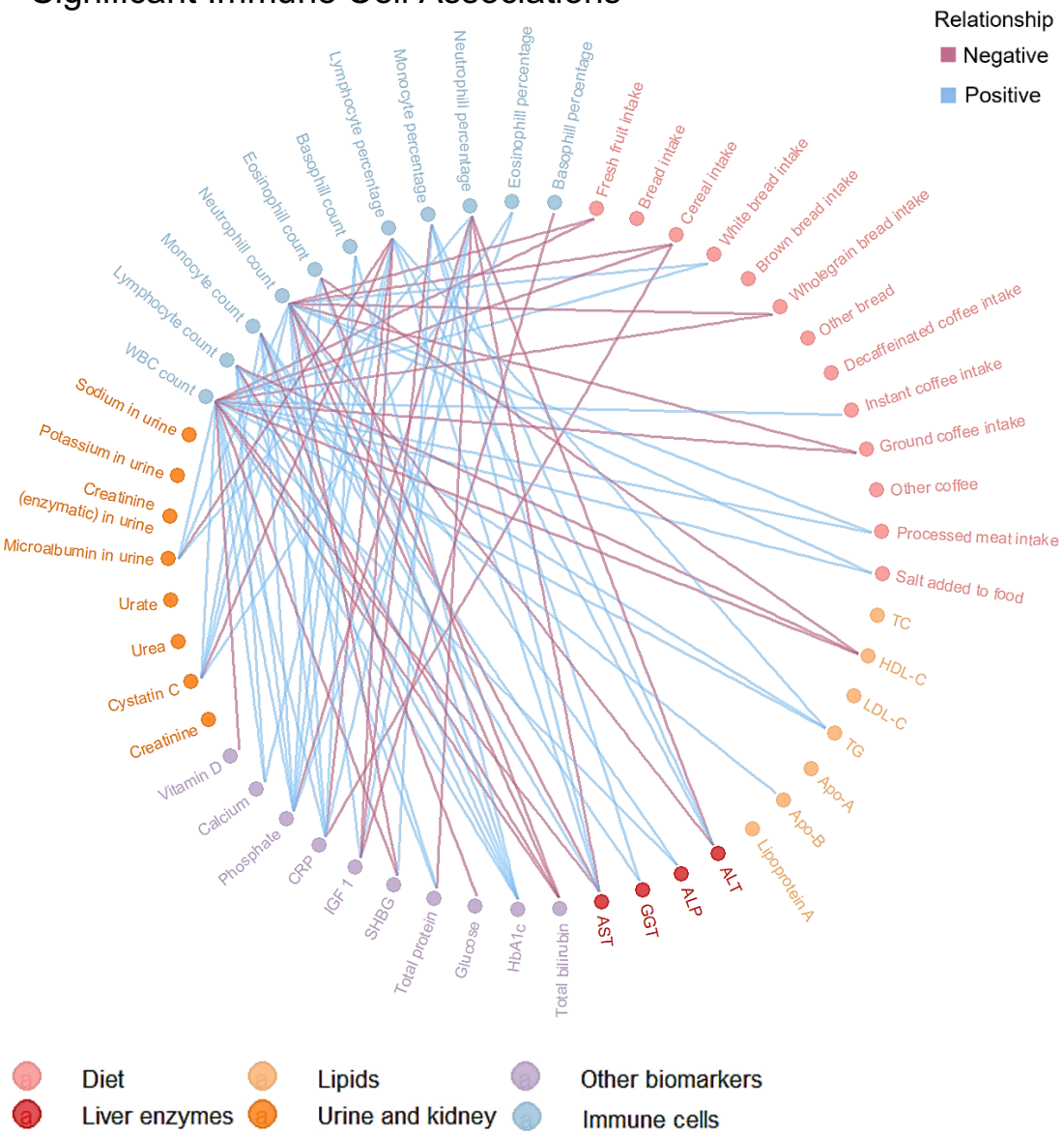


Figure 4: Graphical representations of associations in the D2AF cohort

Network graphs showing significant associations in the D2AF cohort for **A)** lipids, **B)** liver enzymes, and **C)** immune cells

Abbreviations: white blood cell (WBC), high light scatter reticulocyte percentage (HLSRP), high light scatter reticulocyte count (HLSRC), red blood cell (RBC), mean corpuscular volume of RBC (MCV), mean corpuscular haemoglobin concentration (MCHC), RBC distribution width (RDW), aspartate aminotransferase (AST), alanine aminotransferase (ALT) , alkaline phosphatase (ALP), gamma glutamyl transferase (GGT), total cholesterol (TC), high density lipoprotein cholesterol (HDL-C), low density lipoprotein cholesterol (LDL-C), triglycerides (TG), apolipoprotein A (Apo-A), apolipoprotein B (Apo-B), sex hormone binding globulin (SHBG), C reactive protein (CRP), glycated haemoglobin (HbA1c), insulin-like growth factor 1 (IGF-1)

3.1.2 DW2 Cohort

In the DW2 cohort (n=450), 81 associations were identified (see **Appendix Figure 2b**), of which a positive association between biomarker albumin and RBC count (P<0.0001) was the most significant. Whole blood count feature mean sphered cell volume had the highest number of total associations across the three analyses pools, accounting for 13.6% of all association detected in the DW2 cohort (see **Appendix Figure 3b**).

3.1.2.1 Diet ~ Biomarkers

Univariate regression modelling of the DW2 cohort (n = 450) identified no significant (P<0.05) diet to biomarker associations.

3.1.2.2 Diet ~ Whole blood count

Univariate regression analysis of the DW2 cohort (n = 450) identified one significant diet to cell feature association, a negative association between mean sphered cell volume and cereal intake (P<0.05).

3.1.2.3 Biomarkers ~ Whole blood count

Analysis of the DW2 cohort (n = 450) identified 80 significant biomarker to whole blood count feature associations. Of these associations, 20 involved lipids or lipid-related features (see **Figure 5a**) and 16 involved liver enzymes (see **Figure 5b**). Mean reticulocyte volume and mean sphered cell volume were significantly linked to the most biomarkers, including three out of four liver enzymes (P<0.0001).

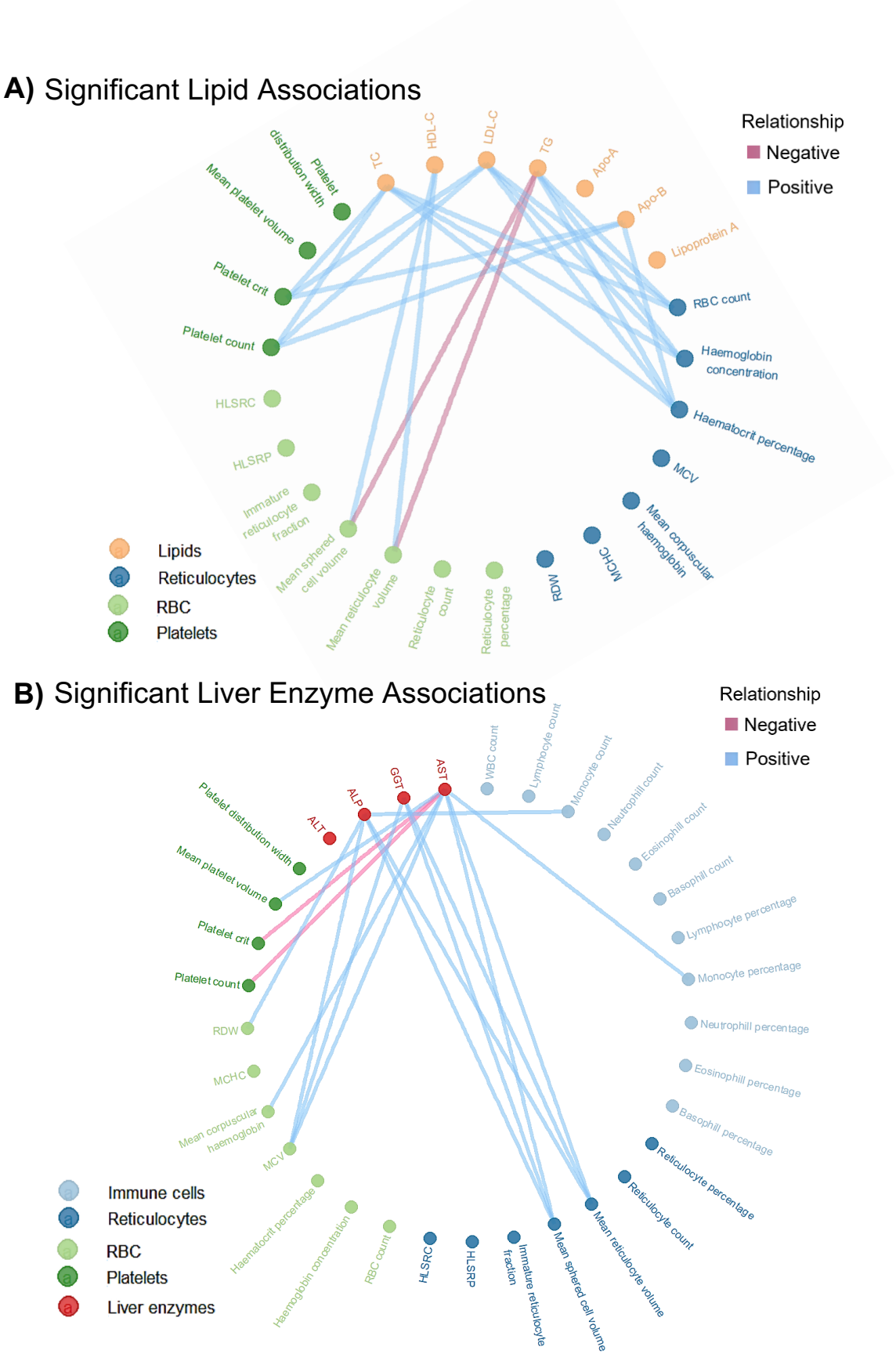


Figure 5: Graphical representations of associations in the DW2 cohort

Network graphs showing significant associations in the DW2 cohort for **A)** lipids and **B)** liver enzymes

Abbreviations: white blood cell (WBC), high light scatter reticulocyte percentage (HLSRP), high light scatter reticulocyte count (HLSRC), red blood cell (RBC), mean corpuscular volume of RBC (MCV), mean corpuscular haemoglobin concentration (MCHC), RBC distribution width (RDW), aspartate aminotransferase (AST), alanine aminotransferase (ALT) , alkaline phosphatase (ALP), gamma glutamyl transferase (GGT), total cholesterol (TC), high density lipoprotein cholesterol (HDL-C), low density lipoprotein cholesterol (LDL-C), triglycerides (TG), apolipoprotein A (Apo-A), apolipoprotein B (Apo-B)

3.1.3 D2BF Cohort

27 associations were detected in the D2BF cohort (n=326) (see **Appendix Figure 2c**), of which a positive association between biomarker albumin and RBC count (P<0.0001) was the most significant. Whole blood count feature mean sphered cell volume had the highest number of total associations across the three analyses pools, accounting for 18.5% of all association detected in the D2BF cohort (see **Appendix Figure 3c**).

3.1.3.1 Diet ~ Biomarkers

Univariate regression modelling of the D2BF cohort (n = 326) identified no significant (P<0.05) diet to biomarker associations.

3.1.3.2 Diet ~ Whole blood count

Univariate regression analysis of the D2BF cohort (n = 326) identified one significant diet to cell feature association, a negative association between RBC count and eating sugar (P<0.0001).

3.1.3.3 Biomarkers ~ Whole blood count

Analysis of the D2BF cohort (n = 326) identified 21 significant biomarker to cell association, of which the most significant association was a positive association between albumin and RBC count (P<0.0001). Blood biomarker albumin was significantly associated with the most lymphocyte profiling features out of the 11 biomarkers identified in the biomarker ~ whole blood count analysis, although the lipid (see **Figure 6a**) and liver enzyme (see **Figure 6b**) groups had the most associations out of the biomarker specific groups.

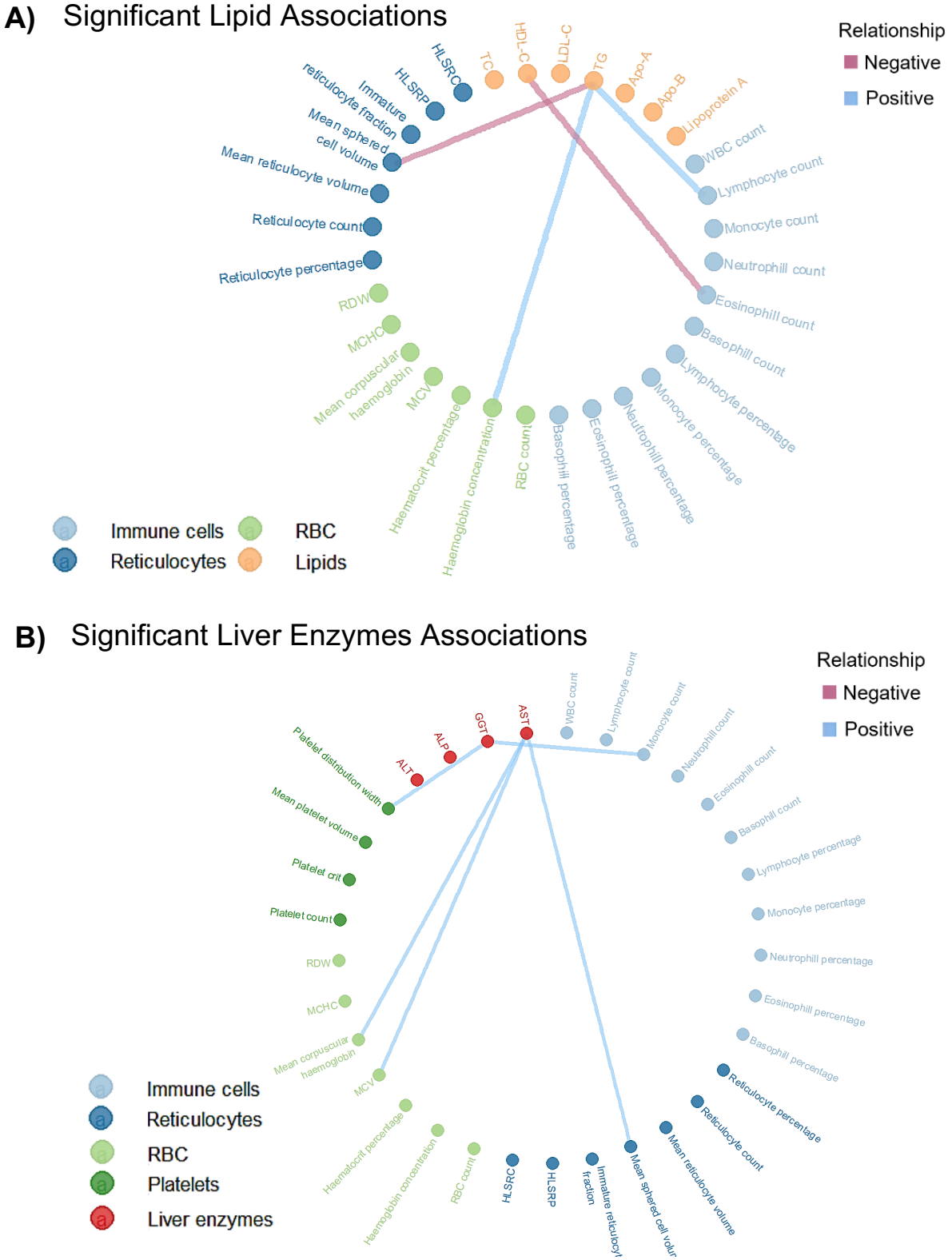


Figure 6: Graphical representations of associations in the D2BF cohort

Network graphs showing significant associations in the D2BF cohort for **A)** lipids and **B)** liver enzymes

Abbreviations: white blood cell (WBC), high light scatter reticulocyte percentage (HLSRP), high light scatter reticulocyte count (HLSRC), red blood cell (RBC), mean corpuscular volume of RBC (MCV), mean corpuscular haemoglobin concentration (MCHC), RBC distribution width (RDW), aspartate aminotransferase (AST), alanine aminotransferase (ALT) , alkaline phosphatase (ALP), gamma glutamyl transferase (GGT), total cholesterol (TC), high density lipoprotein cholesterol (HDL-C), low density lipoprotein cholesterol (LDL-C), triglycerides (TG), apolipoprotein A (Apo-A), apolipoprotein B (Apo-B)

3.1.4 Control cohort

497 associations were detected in the control cohort (n=7307), (see **Appendix Figure 2d**), involving 90 unique features across three data types, of which a positive association between reticulocyte count and TG (P<0.0001) was the most significant. Biomarker cystatins C had the highest number of total associations across the three analyses pools, accounting for 6.04% of all association detected in the control cohort (see **Appendix Figure 3d**).

3.1.4.1 Diet ~ Biomarker

Univariate regression of the control cohort (n = 7307) identified 121 significant diet to biomarker associations, of which an inverse association between sodium in urine and cereal intake was the most significant (P<0.0001. Of the 32 biomarkers studied, 14 biomarkers were associated with cereal consumption, 11 biomarkers were associated with bread consumption, and 12 biomarkers were associated with coffee consumption. Lipid markers TC, HDL-C, and apolipoprotein A (Apo-A) were significantly associated with cereal consumption (P<0.05) and LDL-C, HDL-C, TC, Apo-A were significantly associated with coffee consumption (P<0.05) (see **Figure 7b**). Liver enzymes GGT and AST were both significantly associated with fresh fruit intake (P<0.05) (see **Figure 7a**). Of the 121 associations, 56 related to urinary and kidney markers, of which cystatin C accounted for 28.6% of these.

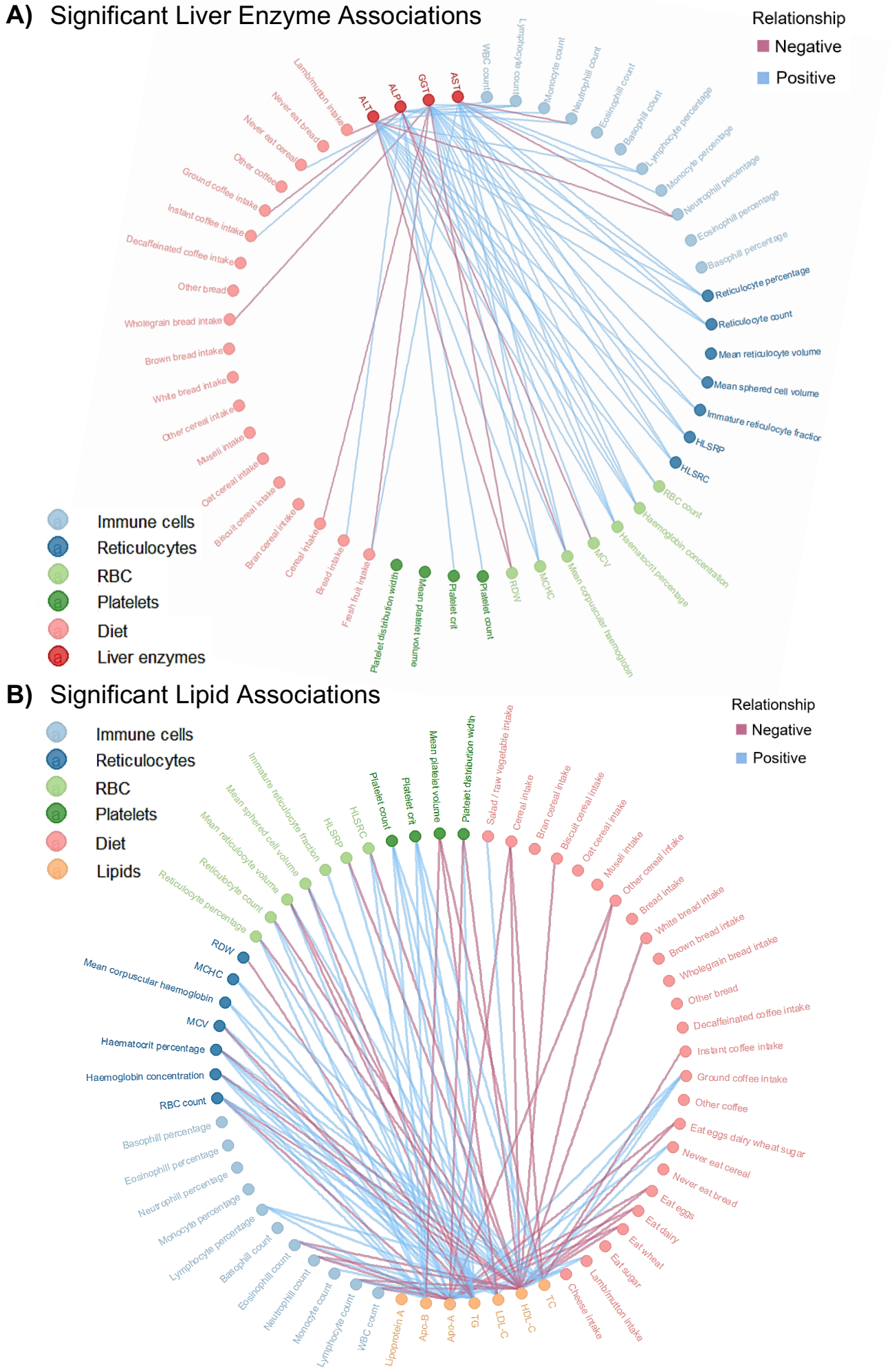
3.1.4.2 Diet ~ Whole blood count

Univariate regression of the control cohort (n = 7307) identified 65 diet to cell associations that were significant. Of the 29 whole blood count features studied, 15 were associated with cereal intake. Of the 65 detected associations, 22 related to

reticulocytes and a further 22 related to RBCs and haemoglobin, including nine involving mean corpuscular volume (MCV).

3.1.4.3 Biomarker ~ Whole blood count

Analysis of the control (n = 7307) cohort identified 311 significant biomarker to cell associations. Of the 311 associations, 74 involved measurements of lipids (see **Figure 7b**), 21 of which were TG, and 101 involved measurements of WBCs (see **Figure 7c**), relating to 26 unique biomarkers.


A) Significant Liver Enzyme Associations
Relationship
Negative
Positive
Immune cells
Reticulocytes
RBC
Platelets
Diet
Liver enzymes
ALT
ALP
GGT
AST
WBC count
Lymphocyte count
Monocyte count
Neutrophil count
Eosinophil count
Basophil count
Lymphocyte percentage
Monocyte percentage
Neutrophil percentage
Eosinophil percentage
Basophil percentage
Reticulocyte percentage
Reticulocyte count
Mean reticulocyte volume
Mean sphered cell volume
Immature reticulocyte fraction
HLSRP
HLSRC
RBC count
Haemoglobin concentration
Haematocrit percentage
MCV
Mean corpuscular haemoglobin
MCHC
RDW
Platelet count
Platelet crit
Mean platelet volume
Platelet distribution width
Fresh fruit intake
Bread intake
Cereal intake
Bran cereal intake
Biscuit cereal intake
Oat cereal intake
Muesli intake
Other cereal intake
White bread intake
Brown bread intake
Wholegrain bread intake
Other bread
Decaffeinated coffee intake
Instant coffee intake
Ground coffee intake
Other coffee
Never eat cereal
Never eat bread
Lamb/mutton intake
B) Significant Lipid Associations
Relationship
Negative
Positive
Immune cells
Reticulocytes
RBC
Platelets
Diet
Lipids
Immature reticulocyte fraction
Mean sphered cell volume
Mean reticulocyte volume
Reticulocyte count
Reticulocyte percentage
HLSRP
HLSRC
Platelet count
Platelet crit
Mean platelet volume
Platelet distribution width
Salad / raw vegetable intake
Cereal intake
Bran cereal intake
Biscuit cereal intake
Oat cereal intake
Muesli intake
Other cereal intake
Bread intake
White bread intake
Brown bread intake
Wholegrain bread intake
Other bread
Decaffeinated coffee intake
Instant coffee intake
Ground coffee intake
Other coffee
Eat eggs dairy wheat sugar
Never eat cereal
Never eat bread
Eat eggs
Eat dairy
Eat wheat
Eat sugar
Lamb/mutton intake
Cheese intake
RDW
MCHC
Mean corpuscular haemoglobin
MCV
Haematocrit percentage
Haemoglobin concentration
RBC count
Basophil percentage
Eosinophil percentage
Neutrophil percentage
Monocyte percentage
Lymphocyte percentage
Basophil count
Eosinophil count
Neutrophil count
Monocyte count
Lymphocyte count
WBC count
Lipoprotein A
Apo-B
Apo-A
TG
LDL-C
HDL-C
TC

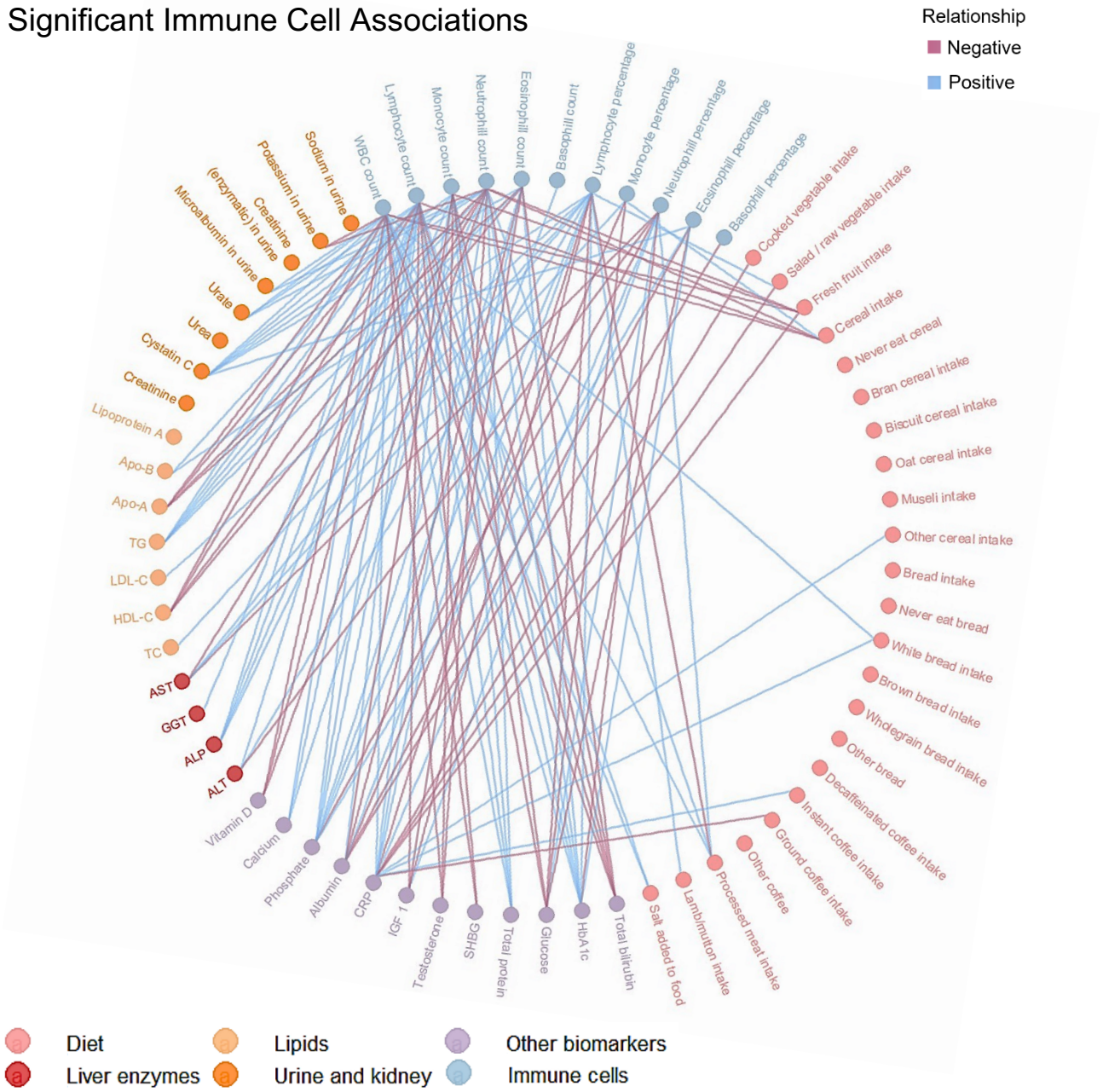


Figure 7: Graphical representations of associations in the control cohort

Network graphs showing significant associations in the control cohort for **A)** liver enzymes, **B)** lipids, and **C)** immune cells

Abbreviations: white blood cell (WBC), high light scatter reticulocyte percentage (HLSRP), high light scatter reticulocyte count (HLSRC), red blood cell (RBC), mean corpuscular volume of RBC (MCV), mean corpuscular haemoglobin concentration (MCHC), RBC distribution width (RDW), aspartate aminotransferase (AST), alanine aminotransferase (ALT) , alkaline phosphatase (ALP), gamma glutamyl transferase (GGT), total cholesterol (TC), high density lipoprotein cholesterol (HDL-C), low density lipoprotein cholesterol (LDL-C), triglycerides (TG), apolipoprotein A (Apo-A), apolipoprotein B (Apo-B), sex hormone binding globulin (SHBG), C reactive protein (CRP), glycated haemoglobin (HbA1c), insulin-like growth factor 1 (IGF-1)

3.1.5 Overlap between cohorts

In total, 990 associations were identified across the four cohorts, and 13 were shared between all cohorts (see **Figure 8a**). These were all associations between biomarkers and whole blood count features, involving measurements of RBCs, reticulocytes, and platelets with albumin, AST, TG, total bilirubin, and sex hormone binding globulin (SHBG). No mutual diet associations were found between all four cohorts (see **Figures 8c-d**). Control and D2AF cohorts shared the highest number of mutual associations (69.61% of total D2AF cohort associations and 53.92% of total control cohort associations), with 196 mutual associations in the biomarker ~ whole blood count association analysis alone (see **Figure 8b**). The D2BF and control cohorts shared the fewest number of overall associations (17 associations, accounting for 3.42% of total associations detected in the control cohort). Both the DW2 cohort and D2BF shared the most associations (83.95% and 92.59% of their total associations, respectively) with the D2AF cohort. For a full list of all significant associations identified through univariate regression in all four datasets, see **Appendix Table 2**.

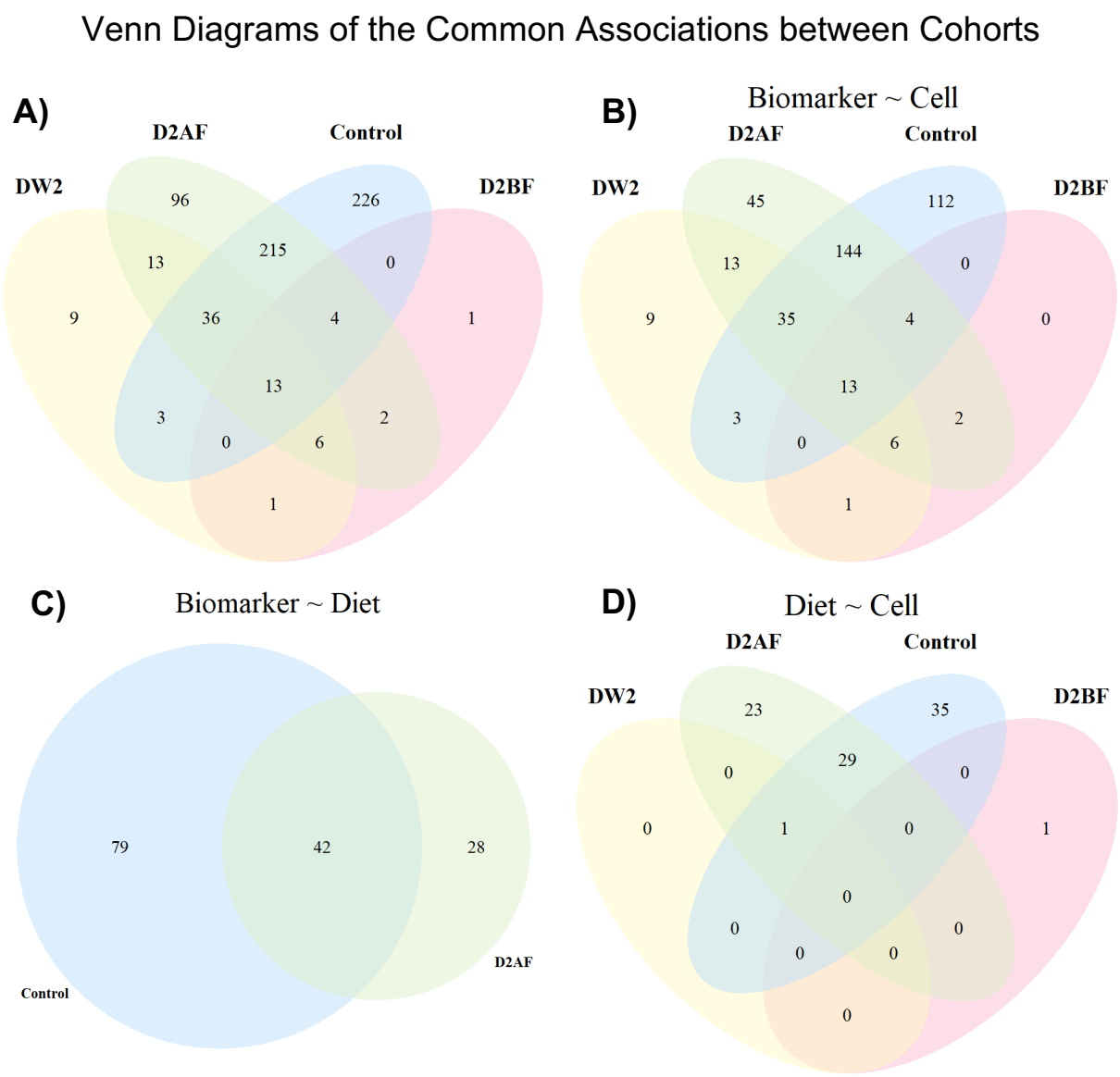


Figure 8: Venn diagrams of the common associations between cohorts

Venn diagrams of the common associations between the four cohorts for **a)** all analysis, **b)** biomarker to whole blood count analysis, **c)** biomarker to diet analysis, and **d)** diet to whole blood count analysis.

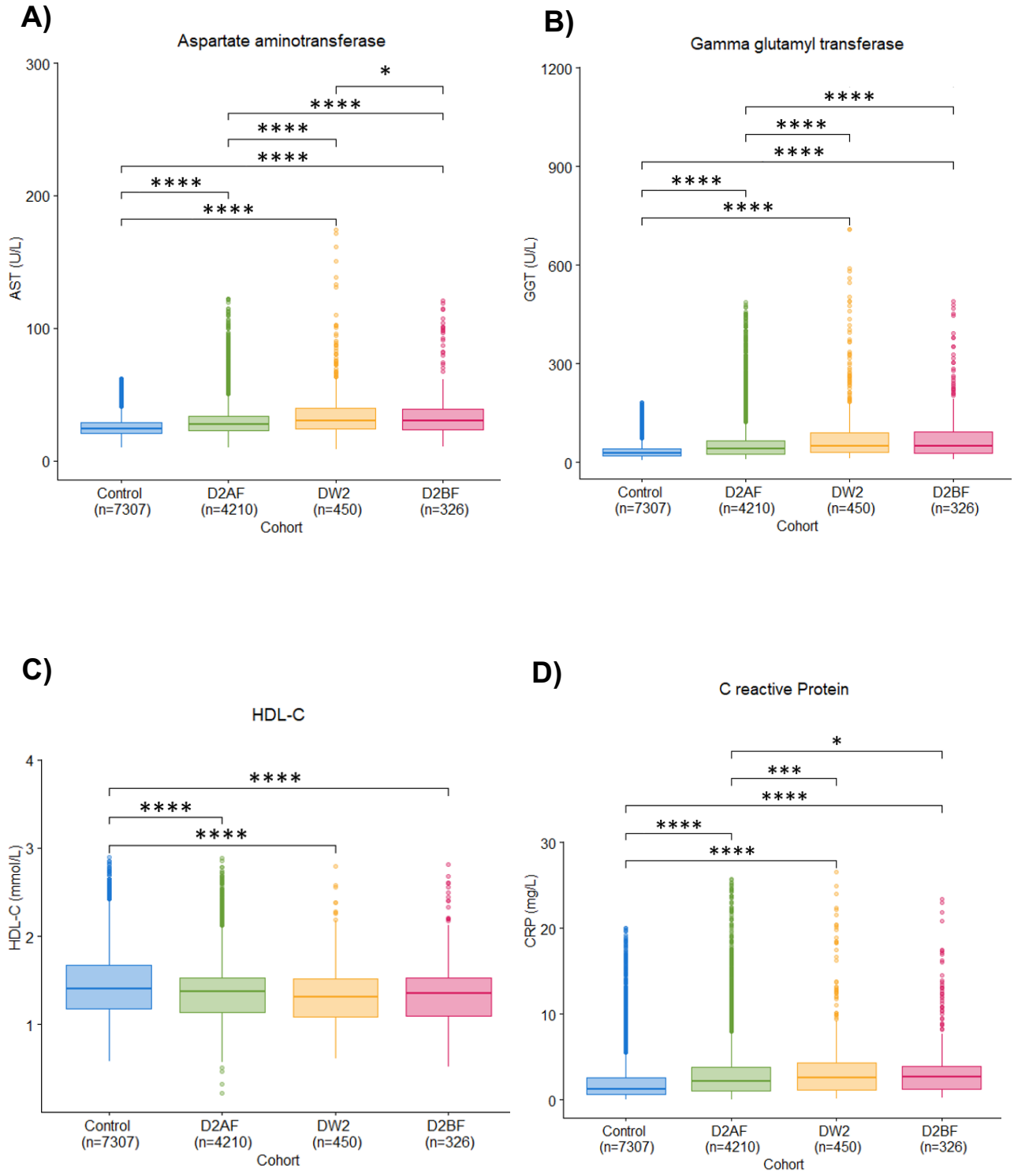
3.2 ANOVA

A total of 32 biomarkers and 29 whole blood count features were investigated in this study. Of the 61 features, 53 had at least one significant different between two cohorts, of which 29 were biomarkers and 24 were whole blood count features. All of these features, apart from haemoglobin concentration, had at least one significant difference

between a NAFLD cohort and the control cohort. Biomarkers AST and direct bilirubin had the highest number of significant differences between cohorts. Select examples are shown in **Figures 9a-h** to highlight the differences between cohorts in select clinically relevant features. See **Appendix Table 3** for all significant differences between cohorts.

GGT and AST levels showed a progressive increase from control to D2AF, to DW2 and D2BF cohorts. Values are given as (mean [95%CI]). GGT levels were within the normal range (5 to 40 U/L) for the control cohort (34.3 U/L [33.7 to 34.8]) but were well above the normal range in the D2AF (62.0 U/L [60.0 to 64.0]), DW2 (88.9 U/L [79.0 to 98.8]), and D2BF (81.6 U/L [72.1 to 91.1]) cohorts. There was no significant difference between DW2 and D2BF cohort GGT levels (see **Figure 9b**). Whilst AST levels were within normal range (8 to 33 U/L) for both control (25.7 U/L [25.5 to 25.9]) and D2AF cohorts (31.4 U/L [30.9 to 31.8]), they were above normal range for DW2 (37.2 U/L [35.2 to 39.3]) and D2BF (35.1 U/L [33.1 to 37.2]) cohorts (see **Figure 9a**).

ANOVA box plots of significant differences between cohorts



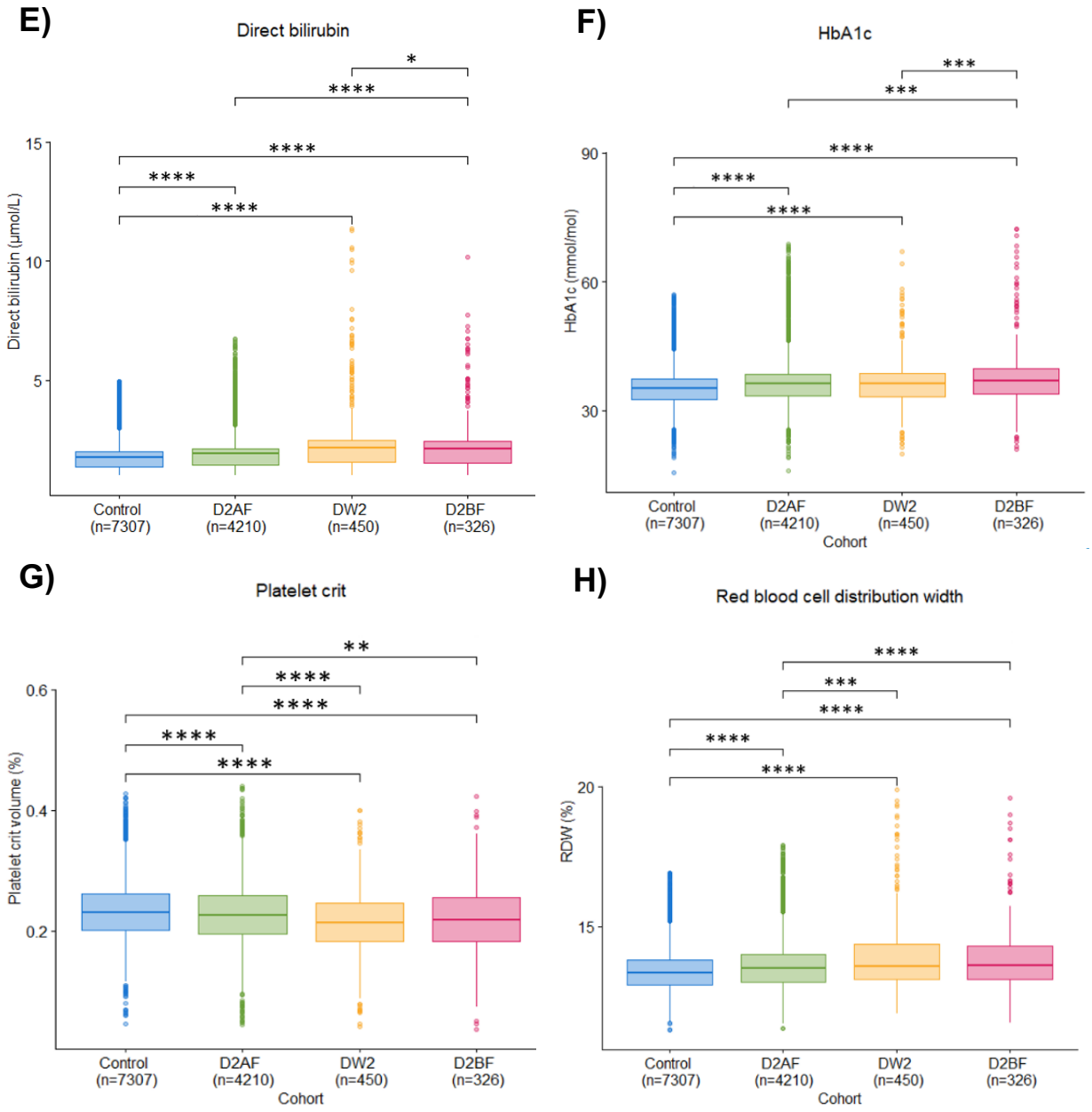


Figure 9: ANOVA box plots of significant differences between cohorts

Box plots display median and interquartile range (IQR) of **a)** AST, **b)** GGT, **c)** HDL-C, **d)** CRP, **e)** direct bilirubin, **f)** HbA1c, **g)** platelet crit, and **h)** RDW between the D2BF (n=326), DW2 (n=450), D2AF (n=4210), and control (n=7307) cohorts. A one-way ANOVA and Tukey HSD tests were performed with corrected P values obtained at $p<0.05$. Comparison lines between box plots are marked to indicate between cohort significance with * = $p<0.05$, ** = $p<0.01$, *** = $p<0.001$, and **** = $p<0.0001$

Abbreviations: aspartate aminotransferase (AST), gamma glutamyl transferase (GGT), high density lipoprotein cholesterol (HDL-C), C reactive protein (CRP), glycated haemoglobin (HbA1c), red blood cell (RBC), red blood cell distribution width (RDW)

Chapter 4. Discussion

In this study, we used a UK Biobank data set to associate between multiple data types including biomarkers, dietary features, and whole blood count features, with the aim to find interesting and novel associations that could add value in the field of NAFLD.

As previously stated in **Chapter 1**, a literature search was conducted on PubMed for papers including the terms ‘UK,’ ‘biobank’ and ‘NAFLD’ on 13/10/2021 and identified 26 published papers utilising the UK Biobank Data on NAFLD. Of these papers, nine were genomic studies and many studies investigated comorbidities for NAFLD, identifying obesity, dyslipidaemia, CVD, poor muscle function, and T2D (81-86). As far as this author is aware, nutritional UK Biobank data for NAFLD patients has only been used for one study, to investigate the relationship between caffeine and NAFLD (22, 87). UK Biobank was chosen for this study because of its large sample size, array of variables available, and the range of diagnosis dates. More information on this cohort is available in **section 2.1**.

4.1 Liver Enzymes

Diagnosis of NAFLD patients is currently based upon ultrasound readings and liver function blood tests, which include liver enzymes, although the gold standard of diagnosis remains liver biopsy (9).

GGT levels were dramatically raised in the NAFLD cohorts compared to the control cohorts, indicative of liver or bile duct damage (12, 13). Those in the D2AF cohort had

GGT levels well above normal, suggesting that GGT levels may be an early indicator for NAFLD. Our results align with the use of GGT as part of the biomarker panel for NAFLD diagnosis, but our results suggest a stronger association and greater importance of GGT in NAFLD development than originally believed (4, 12, 15, 16).

In our study, AST levels showed a progressive increase from control to D2AF, to DW2 and D2BF cohorts. This corresponds to current understanding and AST usage as diagnostic marker (16, 17). However, although ALT levels were elevated in the NAFLD cohorts compared to the control cohort, contradictory to its use as a biomarker for NAFLD, levels were still within normal range (4 to 36 U/L) for all four cohorts (16, 17). The AST:ALT ratios were above one, indictive of advanced liver disease, for all four cohorts, and even in the DW2 cohort, AST:ALT ratios were not significantly different from that of the control, so it would not prove useful as diagnostic marker for NAFLD patients in this study (15, 18). This is suggestive that AST alone may be a better predictor of NAFLD diagnosis than ALT or the AST:ALT ratio.

ALP levels followed a similar pattern to ALT – although levels were significantly higher in all three NAFLD cohorts than in the control cohort, all four cohorts still came within normal range (30 to 120 U/L). However, ALP has been reported to be overexpressed in cancers of the liver (65, 88). As HCC and NASH diagnoses were not accounted for in our NAFLD cohorts, it may be that ALP and ALT, although not clinically relevant for the diagnosis of NAFLD in this study, may be useful in later diagnoses of related liver diseases and detection of clinically advanced stages of NAFLD e.g. HCC (65).

The D2AF and control cohorts were the only two cohorts which detected significant dietary associations for any of the four enzymes. In these cohorts, the detrimental effect of lamb/mutton intake and instant coffee intake on liver enzymes GGT and ALP levels and positive effect of cereal intake, fresh fruit intake, and ground coffee intake were observed. The impact of increased fibre intake and fewer processed foods on GGT levels in the D2AF cohort demonstrates how diet can be used to reduce key biomarkers of NAFLD and, given the correlation between GGT levels and increased mortality in NAFLD patients, could reduce risk of long term adverse clinical outcomes (4, 12). However, as these nutritional exposures are not observed in those who have been diagnosed with NAFLD in our study, it may be that the susceptibility of liver enzymes to diet alters with the developing pathology of NAFLD.

4.2 Lipids

The global alterations by NAFLD to the liver lipidome were observed through HDL-C and TG as although HDL-C levels were still within normal range of >1.3 mmol/L (women) or >1.0 mmol/L (men) for all four cohorts, according to a recent study, HDL-C <1.39 mmol/L is associated with NAFLD diagnosis after 4.6 years (26). Given D2AF cohort HDL-C levels (1.37 mmol/L [1.36 to 1.38]) were below this threshold, this suggests that prolonged HDL-C <1.39 mmol/L predisposes patients to NAFLD, and levels do not recover after diagnosis, although still within a clinically determined healthy range (26). The use of statins, which reduce TC and LDL-C concentrations via HMG-CoA reductase inhibition, in NAFLD have recently been shown to reduce cancer-related and overall mortality, but not CVD mortality, highlighting the importance of

lipotoxicity in NAFLD pathophysiology, but also suggesting the role of lipids in NAFLD beyond atherosclerotic risk (89).

LDL-C, TC, apolipoprotein B (Apo-B), and TG followed a similar pattern of associations whilst HDL-C and Apo-A tended to have opposite associations to the other lipid species. Across different cohorts and groups HDL-C and TG showed repeated opposite associations to the same dietary features and whole blood count features, such as reticulocyte count, RBC count, and lymphocyte count. The sustained positive association between TG and lymphocytes and negative association between HDL-C and lymphocytes further corroborates the role of lipotoxicity in hepatic inflammation (37). Positive associations were also detected between elevated atherosclerotic lipid parameters LDL-C, Apo-B, TC, and TG, and RBC parameters RBC count, haemoglobin concentration, and haematocrit percentage in the DW2, D2AF, and control cohorts. The effect of cholesterol on RBCs has been studied and elevated TC forces a left shift in haemoglobin dissociation curve, reducing oxygen transport (90). In compensation for this, RBC production may be increased which could explain the increase in RBC count seen with elevated TC and LDL-C (90).

The D2AF and control cohorts were the only two cohorts which detected significant dietary associations for any of the lipid parameters measured. In these cohorts, the detrimental effects of never eating cereal and eating cereal such as Frosties (other cereal intake), opposed the positive effects of higher cereal intake. Cereal consumption features were the only retained associations between the two cohorts, whilst the control cohort also detected the positive effects of ground coffee versus instant coffee on the liver lipidome, and the D2AF cohort found fresh fruit reduced

atherogenic lipid parameters. These dietary associations are similar to those observed for the liver enzyme biomarkers – with greater fibre intake there is a reduction in biomarkers predisposing to NAFLD.

LDL-C and TC levels were above normal range (3 mmol/L and 5 mmol/L respectively) in all cohorts, including the control cohort; whilst elevated cholesterol has been reported in NAFLD patients, as described in **section 1.2.2**, the control cohort would be expected to have cholesterol within the normal range (20-26). This may be explained by the patient demographic (see **Table 2**), as average patient age at recruitment was 57 in the control cohort and patients were averaging an overweight BMI of 26.5. Increased age and BMI are known risk factors for hyperlipidaemia, but this lipidomic profile of the control cohort is cause for concern and signifies a systemic hyperlipidaemia problem in the general population (4). However, although LDL-C and TC levels for the three NAFLD cohorts were also above their respective normal ranges, they were lower in the NAFLD cohorts than in the control cohort, in direct contradiction to current knowledge and understanding of NAFLD. Why these two lipid species did not follow the expected pattern for lipids in NAFLD, but HDL-C and TG did, is unclear and more comprehensive lipidomic analysis is needed.

Beyond hepatic cholesterol, TG, and FFA accumulation, more specific alterations to the liver lipidome with NAFLD have been found, including disturbances to the balance of sphingomyelins and cholesteryl esters, suggesting the role of ceremide synthesis pathways to cholesterol and TG accumulation seen in NAFLD (27, 91, 92). These lipid species have been used to distinguish between NAFLD and NASH lipidomes and although these parameters are available through UK Biobank, there was a lack of

sufficient data in this study to measure these species in our research (27). Recent studies have gone beyond this and highlighted the role of ceremide and sphingolipids in HCC and fibrosis (91-93). Exploration of these lipid species in UK Biobank NAFLD participants may explain the lowered TC and LDL-C in NAFLD cohorts compared to the control cohort and should be the focus of future work.

Although no pharmacological treatments are currently in use for treating steatosis in NAFLD, drugs targeting lipotoxicity are under development such as peroxisome proliferator-activated receptor agonists to increase fatty acid oxidation and oppose hypercholesterolaemia (55).

4.3 Inflammatory markers and Immune Cells

The evolution of the inflammatory hepatic immune landscape plays an integral role in NAFLD pathogenesis as a driver of hepatocyte death (17, 31). Inflammation in this study was observed both through WBCs and through inflammation serum biomarker CRP (17, 31). In the UK Biobank, levels of cytokines and interleukins are not available, and CRP was the only inflammatory biomarker we had available in this study, limiting our ability to understand inflammation in this NAFLD cohort.

In our research, CRP levels were significantly lower in the control cohort than all three NAFLD cohorts and significantly increased the longer the patient had been suffering from NAFLD, with the D2BF cohort having the highest CRP levels (see **Figure 9d**). An increase in CRP levels is associated with higher risk of NAFLD, even when CRP

levels are within normal range and a recent study has suggested elevated CRP can be used as predictor for NAFLD – certain studies have even gone further to suggest that CRP can be used as a surrogate marker for NAFLD disease severity, although our results do not support this (17, 94). Given that CRP levels for all three NAFLD cohorts, as well as the control cohort, fell well within the normal range for CRP (<10.0 mg/L), our results do not support the clinical application of CRP as a diagnostic biomarker for NAFLD, but do denote an increase in inflammation with prolonged NAFLD suffering. However, hepatic inflammation is a more defining characteristic of NASH pathology and as we excluded NASH participants, this may explain why CRP levels remained within normal range for NAFLD patients.

Although measurements for other proinflammatory markers are not available from the UK Biobank, the observed increase in CRP in NAFLD patients corresponds to previously reported increases in pro-inflammatory markers IL-4, IL-6, and TNF-α which occur in tandem with increased CRP (40, 95). Expression of these pro-inflammatory markers has been shown to increase along the NAFLD spectrum, and these markers have been suggested as biomarkers for increased NAFLD risk (40, 95). IL-4, IL-6, and TNF-α promote fibrosis and mediate oxidative stress and hepatic inflammation through the recruitment and activation of Kupffer cells (95-98). Anti-inflammatory cytokine IL-10 has been associated with reduced NAFLD risk, due to its ability to inhibit secretion of inflammatory cytokines, such as IL-6 and TNF-α (95, 98). Future research into inflammation and NAFLD would benefit from in depth study of interleukin and cytokine data alongside biomarker and dietary data to further inform the process of inflammation in NAFLD.

The increased presence of biomarkers of inflammation in NAFLD patients is reflected by increased inflammatory cell recruitment to the liver. In this retrospective research we observed elevated WBCs, particularly monocytes, the levels of which were positively associated with the levels of GGT in the DW2, D2AF, and control cohorts and AST in the D2BF, D2AF, and control cohorts. Additional WBCs, including neutrophils, eosinophils, and basophils, were all significantly higher in abundance in the D2AF and DW2 cohorts than the control cohort, aligning with the current understanding of the influx of WBCs to the liver in NAFLD and the following immune cell mediated liver injury. However, lymphocytes were only significantly different between the D2AF cohort and the control cohort when measuring lymphocyte count – there was no significant difference between the cohorts when measuring lymphocyte percentage. The opposing reactions of different lymphocyte subtypes to NAFLD may explain the homogeneity between cohorts in lymphocyte abundance (37, 38). The lack of lymphocyte subtype data in our study limits our ability to detect specific effects and pathways involved in hepatic inflammation, but the relationship between lymphocytes and TG does correspond to pre-existing understanding of the role of lipotoxicity in triggering CD8+ T lymphocytes (37). For future research into the associations between biomarkers of NAFLD and the immune system, lymphocyte subtype data, including CD4+, CD8+, Th17, and Treg cells is needed.

The role of diet in reducing inflammatory markers was observed clearly through the control (n=7307) and D2AF (n=4210) cohorts, but not in cohorts of longer NAFLD suffering. This may be due to smaller cohort sizes in the DW2 (n=450) and D2BF (n=326) cohorts or the disturbed metabolism and liver dysfunction characteristic of NAFLD. Of the 12 features of inflammation measured in this study, only CRP, WBC

count, and neutrophil count were associated with dietary features in both the D2AF and control cohorts. Across the two cohorts, wholegrain bread intake, fresh fruit intake, and cereal intake reduced all measures of inflammation, whilst processed meat intake, instant coffee intake, and white bread intake caused an increase in immune cell abundance.

The contribution of immune cell mediated inflammation to the development of the steatotic liver is becoming increasingly relevant in NAFLD research as deeper insight into the role of WBCs in driving hepatocyte death, one of the cardinal features of fatty liver disease pathogenesis, and HCC is revealed (24, 32-34). Moving away from the previous focus on innate immune cells in NAFLD, the recent focus on adaptive immunity in NAFLD research, particularly the role of T lymphocytes, has highlighted the complexity and severity of inflammation in NAFLD, for example, the disrupted balance of T lymphocyte subtypes promotes the development of NAFLD at all ages (37, 38, 99). The relationship between the liver lipidome and lobular inflammation has recently been shown to play a key role in HCC development through CD8+ T lymphocytes, as aberrant cholesterol signalling causes CD8+ T lymphocyte dysfunction and impairs anti-tumour immunosurveillance (100). Beyond the liver, a recent study showed that the increased prevalence of cognitive impairment in NAFLD patients compared to non-NAFLD patients was exacerbated by elevated WBCs (101). WBCs and inflammation in NAFLD have long term and far reaching consequences around the body which we are only starting to understand the mechanics behind (99-101).

4.4 Blood components and composition

The liver is responsible for the breakdown of old and damaged RBCs as well as converting haemoglobin breakdown product, indirect bilirubin, into direct bilirubin (102, 103). Alterations to RBC characteristics in NAFLD have been observed in both previous research and our study with increased red blood cell distribution width (RDW) as well as elevated RBC count in NAFLD patients (102-104). In addition to this, direct bilirubin levels were also raised in those with NAFLD, suggestive of increased RBC breakdown in line with increased RBC count, although the causal relationship between bilirubin and NAFLD is under debate (29, 30).

One of the key mechanisms behind these alterations to RBC-related parameters in NAFLD is oxidative stress, induced by ROS (105-107). Oxidative stress is associated with the progression of NAFLD, induced by excessive ROS signalling, which is a crucial part of the induction of an uninhibited inflammatory environment in NAFLD, as described in **section 1.3** (105-107). One of the suggested mechanisms for the promotion of excessive RBC breakdown and death by oxidative stress is inhibition of Ca-ATPase, resulting in elevated intracellular calcium, which impairs the deformability of RBC and induces cell shrinkage, observed through raised RDW, as well as prematurely ageing RBCs (102, 106-108).

The interplay between lipid species, RBCs, and oxidative stress is complex with oxidative stress and excessive ROS production as key drivers for both RBC alterations and dysregulated lipid metabolism in NAFLD (109). In this way, oxidative stress and ROS mediates the positive association between atherosclerotic lipid parameters and

RBC parameters described in **section 4.2** (90, 108, 109). Haemoglobin, in particular, is more suspectable to ROS-induced damage and has been connected to TG, both in our study and previous research, and positively associated with increased fibrosis (88, 104, 108, 110). However, there is very little research into the mechanisms underlying the relationships between ROS, dyslipidaemia, and RBCs in NAFLD and further exploration of this research through biochemical analysis is needed to understand the role of oxidative stress in NAFLD.

The connection between NAFLD and RBCs was observed more keenly observed in this study in immature RBCs (reticulocytes) with their connection to liver enzymes and the liver lipidome. AST was positively associated with mean reticulocyte volume in all cohorts, and GGT, ALT, and ALP were positively associated with additional reticulocyte measurements in DW2, D2AF, and control cohorts. HDL-C and TG had opposing effects on reticulocyte parameters, with TG negatively associated with mean reticulocyte volume across all cohorts and HDL-C positively associated with mean reticulocyte volume and mean sphered cell volume of reticulocytes in the DW2, D2AF, and control cohorts. This surprising relationship with lipids opposes our results for reticulocytes with liver enzymes, as we would expect to observe TG elevated with mean reticulocyte volume in the same pattern as liver enzymes. This relationship may be explained through reticulocyte count, which was inversely associated with HDL-C in the D2AF and control cohorts, so although reticulocyte volume parameters may not behave as expected, reticulocyte count does in respect to lipids and liver enzymes. But as no associations with reticulocyte count were found in those diagnosed with NAFLD, detailed reticulocyte analysis in NAFLD is needed. Over findings with reticulocytes are completely novel as, as with RBCs, reticulocytes have been poorly

studied in relation to NAFLD and this problem is only exacerbated when attempting to look at the biochemical mechanics behind the role of reticulocytes in NAFLD.

In NAFLD patients, the increase in RDW is also observed through the increase in RDW to platelet ratio (RPR), with a significant association between advance fibrosis and elevated RPR (111). Thrombocytopenia in NAFLD patients has been explored and low platelet count is more prevalent in NAFLD patients than non-NAFLD patients, occurring in up to 25% of NAFLD patients and can be indicative of end stage live disease, but nearly all studies of platelet count and NAFLD use non-invasive methods and do not occur at point of diagnosis (112-114). In the first study of thrombocytopenia related to histological diagnosis of NAFLD, it was found that low platelet levels in NAFLD patients without cirrhosis correlated to high haemoglobin levels (114). Previous studies also have found that platelet count is not associated with insulin resistance, cholesterol, AST, and ALT levels in NAFLD patients (112, 114). This corresponds to our results for platelet count for the D2BF and DW2 cohorts, however, in the D2AF and control cohorts, platelet count was inversely linked with GGT, AST, and ALT levels as well as positively associated with TC, LDL-C, and Apo-B levels. The mechanisms behind why these connections are broken in NAFLD are unclear, but platelet count and platelet crit were both associated with other biomarkers made in the liver – both platelet parameters are positively associated with total bilirubin and negatively associated with SHBG in all four cohorts. Both total bilirubin and SHBG levels were elevated in NAFLD cohorts and positively associated with reticulocyte and haemoglobin parameters in the DW2 cohort. The causes for thrombocytopenia in NAFLD are unclear, but our study highlights to the contrasting relationships between platelets and different liver-related biomarkers and how liver dysfunction, but further

research is needed to define the mechanisms behind thrombocytopenia in NAFLD (113, 114).

In the ANOVA and Tukey HSD tests, both platelet count and platelet crit levels were highly significantly different between patient groups with levels lower in the D2BF and DW2 NAFLD cohorts than the control and D2AF cohorts. But of the two features, only platelet crit, a measure of the volume that platelets occupy in the blood, was clinically relevant as platelet crit levels were within normal range (0.220% to 0.240%) in the control (0.233% [0.232 to 0.234]) and D2AF (0.227% [0.226 to 0.229]) cohorts but significantly below normal range in DW2 (0.213% [0.208 to 0.219]) and D2BF (0.218% [0.211 to 0.224]) cohorts. Platelet crit is a less common measurement of platelet function, but its similarities with the associations observed for platelet count in this study in addition to its clinical relevance for NAFLD patients suggests the potential use of platelet crit in the diagnosis of thrombocytopenia in NAFLD. The clinical significance of platelet crit in NAFLD is a novel finding as previous research has not highlighted the connection between platelet crit and NAFLD before. However, platelet crit not part of common whole blood count panels, which may limit its usefulness as a diagnostic tool. Further research is needed into thrombocytopenia in NAFLD, particularly the mechanisms behind this association, and potentially utilising platelet crit in conjunction with platelet crit in future studies of thrombocytopenia in NAFLD (113, 114).

4.5 Urinary and Kidney Function Markers

The close association between markers of CKD and NAFLD was observed in this study by measuring an array of urine and kidney-related biomarkers. Whist there was not direct GFR data available, there are indirect ways of estimating kidney function through biomarkers associated with kidney function, primarily urinary cystatin C, serum creatinine, and albuminuria (51, 115, 116). Additional measures of kidney function include urinary creatinine and potassium, with newer biomarkers such as monocyte chemoattractant protein one being currently investigated (51, 115, 117, 118).

Cystatin C is produced by cells in the body and is an indicator of GFR, and therefore kidney function, as it is directly and freely filtered by the glomerulus in the kidneys (51, 119). The normal range for cystatin C is 0.62 – 1.15 mg/L and cystatin C levels above of this range are indicative of kidney dysfunction (51, 116, 119). Compared to the control cohort, urinary cystatin C levels were significantly increased the longer a patient had been diagnosed with NAFLD, suggesting a close association between NAFLD progression and renal dysfunction. Cystatin C levels were positively correlated with multiple WBC features in all four patient cohorts and the linear regression models for cystatin C levels in all three NAFLD cohorts demonstrated that when cystatin C levels were above normal range, WBC count levels were also above normal range (4.50 - 11.0 x 10^9 cells/L). This pattern was repeated for neutrophil, eosinophil, and monocyte counts. In line with other studies, our results suggest that elevated signs of renal dysfunction are linked to increased inflammation in NAFLD which could contribute to the faster progression of NAFLD in patients with CKD (120, 121).

Contradictory to previous studies on NAFLD and the known characteristics of CKD patients, in this research we found that the levels multiple secondary measurement of kidney function, including urinary creatine, improved the longer a patient had been suffering with NAFLD (120-122). Although blood creatinine was not significantly different between all four patient group, creatinine in urine was significantly higher in the control patients than in the three NAFLD patient groups. There were no significant differences in urinary creatinine levels between the three NAFLD patient groups. This is contradictory to current research on the deterioration of urinary markers of kidney function and understanding of hepatorenal syndrome wherein further progression of NAFLD to more clinically advanced stages is associated with a greater degree of renal dysfunction and injury (120, 121). Both potassium in urine and sodium in urine levels followed this pattern of decreasing in the NAFLD cohorts compared to the control cohort, contrasting the previously observed increase with NAFLD and CKD (118, 120-122). However, whilst these urinary markers can be used as additional indictors of kidney function alongside GFR, they are also highly variable with diet, and spot urinary measurements are less accurate representations of true ion excretion values (118, 120-123). As there were no significant associations between urinary and kidney function biomarkers with dietary features in the DW2 an D2BF cohorts, diet may not be a main contributor to this pattern between cohorts, but nutritional break down of dietary information to ascertain specific nutrient intakes would be needed to determine this.

4.6 Gut Microbiome and Diet

Disruption of the gut microbiome homeostasis and metabolic reprogramming accompany metabolic disorders and chronic liver diseases, including NAFLD (26, 124). The gut microbiome influences the host metabolome due to connections with host lymphocyte gene expression and close ties to the liver through both venous blood supply and secretion of bile (124, 125). The gut microbiome is heavily influenced by dietary patterns (124, 125). The predominant bacterial species in the gut belong to Bacteroidetes, Firmicutes, and Proteobacteria at the phylum level (124, 125). The dysequilibrium of the symbiotic relationship between the gut microbiota and host in obesity and NAFLD results in a significantly lower availability of bacteria belonging to the Bacteroidetes taxonomic phylum with overgrowth of member of the Firmicutes phylum (4, 124). Studies have demonstrated the association between the pathogenesis of NAFLD and dysbiosis of the gut microbiome by showcasing the transmission of NAFLD through faecal matter transplants, with linkage to disease severity and impaired mucosal immunity (124-126). Recent studies have even found that the gut microbiome can be used as an early clinical warning sign for NAFLD development and dysbiosis of the gut microbiome fuels HCC development (26, 127). Bacteroidetes phyla relative abundance is decreased in NAFLD patients and negatively correlated to hepatic steatosis staging, ALT, AST, and GGT levels, aligning with the liver enzyme results from our study (124, 126).

Anaerobic bacteria in the gastrointestinal tract, such as Bacteroidetes and Firmicutes, can metabolise polysaccharides, which are found in abundance in fibre rich foods, to form short chain fatty acids (SCFA) (124, 128). Although the underlying mechanisms

behind the metabolic effects of the gut microbiome are not fully established, SCFA produced by gut microbiome are known to play a key role in mediating metabolic interactions between the host metabolome and immune system with the gut microbiome (124, 128). Novel microbiome centric therapies for NAFLD and cirrhosis include probiotics focusing on SCFA butyrate producing bacteria and several clinical trials have shown the value of both probiotic and prebiotic use in liver disease (125, 129). SCFA interact with immune cells to promote anti-inflammatory immune processes and allow tolerance for gut microbiota to enable symbiosis (124). Toll-like receptor (TLR) regulation is a key role of SCFA, which interact with TLR expressing cells to cause the release of immunoregulatory cytokines, including interleukin-2 which promotes the trans-differentiation of naïve T lymphocytes into CD4+ T lymphocytes (124, 130). Increased TLR ligands have been highlighted in relation to obesity as this leads to prolonged TLR signalling stimulation, promoting host T lymphocyte damage and inflammation (124, 130). In our results, we observed a corresponding increase in CRP and a variety of WBCs in NAFLD patients, described in further detail in **section 4.3**, denoting the expected increase in inflammation with dysbiosis in NAFLD (124). Additionally, in the D2AF cohort, WBC count was positively associated with white bread intake and processed meat intake and negatively associated with wholegrain bread intake, cereal intake, and fresh fruit intake, indicative of reduced inflammation with increased intake of fibre rich foods, which may be mediated via polysaccharide breakdown into SCFA by the gut microbiome (124, 128). Although other studies have also observed increased inflammatory markers and proinflammatory immune cells in NAFLD, research around the gut microbiome, dietary patterns, and the immune system in combination is more limited (17, 31, 124, 130).

An influential factor in altering the gut microbiome is dietary fibre, which promotes the growth of SCFA-producing taxa due the high content of non-digestible polysaccharides (124, 128). All cardinal features of NAFLD measured in our analysis, inflammation, lipids, and liver enzymes, were reduced with increased intake of dietary fibre sources, especially cereal. Wholegrain bread, ground coffee, and fresh fruit were also frequently highlighted and shown to reduce biomarker and whole blood count feature risk factors for NAFLD, whilst white bread and instant coffee intake increased them. Our results are supported by previous research showing significantly lower availability of bacteria belonging to the SCFA-producing Bacteroidetes phylum in NAFLD (4, 124). These effect observed in our research may have been mediated by the gut microbiome and SCFA, but due to the lack of microbiome data in this study we cannot drawn conclusive mechanistic pathways between these variables. Further research is needed linking the gut microbiome, diet, blood biomarkers, and whole blood count features to illuminate the pathways involved between the four groups and how each of the groups affect each other simultaneously.

The only previous research using the dietary UK Biobank data and NAFLD found no significant causal relationship between coffee consumption and NAFLD, which does not align with both previous epidemiological studies on coffee intake in NAFLD and our results (87, 131, 132). Our results were novel in exploring the effects of different types of coffee on NAFLD and showing contrasting effects between instant and ground coffee, with ground coffee reducing NAFLD markers. However, we did not look at the degree of coffee intake for each coffee type. Previous epidemiological studies on the effect of coffee intake on NAFLD risk have shown that increased coffee intake significantly reduces NAFLD risk, which aligns with our results for ground coffee, but

the effects of different types of coffee have not been accounted for (131, 132). Future research into the association between coffee and NAFLD should study in the intake of different types of coffee in relation to NAFLD risk and biomarkers of NAFLD in other cohorts to confirm our findings in this study.

Dietary intervention research for NAFLD has looked in detail at the beneficial effects of the Mediterranean diet, which is high in fish and fresh fruit and vegetables and low in saturated fats and processed foods (25, 44, 56). Whilst our results do support the use of Mediterranean diet in NAFLD treatment in terms of a higher fresh fruit and vegetable intake being beneficial, we found little evidence to support the intake of fish reducing NAFLD biomarkers. However, in this study we only had sufficient data to study non-oily fish dietary intake, but not oily fish intake and blood omegas, so our research cannot be used to draw complete conclusions about the Mediterranean diet. Further dietary research into NAFLD could look at the intake of different types of fish as well as both micronutrient and macronutrient intake to inform our understanding of the Mediterranean diet and its NAFLD-risk reducing effects.

Our results suggest that future dietary interventions for NAFLD should focus upon increasing fibre content through high intake of fruit and vegetables and wholegrain carbohydrates. Additionally, more research is needed into the effects of different intake of different types of coffee and different types of fish. For a more comprehensive understanding of the role diet plays in NAFLD, more studies are needed which measure dietary patterns, the gut microbiome, and NAFLD biomarkers simultaneously.

4.7 Limitations

One of the main limitations of our study lies with the lack of longitudinal biomarker, dietary, and whole blood count data for each participant. With just a single measurement of nutrition, biomarker, and whole blood count data available necessitated split of NAFLD participants into three groups by time to diagnosis in order to make reliable associations between data types in relation to NAFLD diagnosis time.

The grouping of cohorts based upon dates of enrolment and diagnosis of NAFLD relies upon accurate date of diagnosis and we cannot be sure how long a patient had underlying NAFLD before being diagnosed (26). Additionally, the methods used to determine NAFLD diagnosis, whether diagnosis is based upon steatosis identified through liver biopsy, the gold standard for NAFLD diagnosis, versus ultrasound, may affect how quickly or whether a patient gets diagnosed with NAFLD (9, 26).

The use of 24-hour dietary recall and food frequency questionnaires to capture highly detailed information about quantities of each specific dietary component in a singular day and look at usual intakes over time can be problematic with measurement errors and within-person variability (133). The reliability of the UK Biobank 24-hour dietary survey and food frequency questionnaire data has been previously explored and showed moderate to substantial reproducibility over the course of 4 years, depending on the dietary variable (133, 134).

4.8 Conclusions

NAFLD is a multifactorial disease of which many aspects are poorly understood. A multitude of contributing factors work in tandem to induce the hallmarks of NAFLD, primarily steatosis and metabolic dysfunction, which in turn contribute to the onset of fibrosis, cirrhosis, and potentially HCC. These stressors, such as lobular inflammation and lipotoxicity, combine to induce the onset of NAFLD and NASH via triggering of cellular stress pathways, leading to hepatocyte death and the hallmark rise in liver enzymes and excessive lobular fat deposition. Oxidative stress and excessive ROS production, induced through hepatic inflammation, mediates many of the adverse effects observed both in this study and in previous research, including RBC dysfunction and hepatocyte death.

There is a clear disparity in biomarker concentrations between the D2BF and DW2 cohorts with the D2AF and control cohorts denoting the far reaching physiological and biochemical effects of NAFLD. However, our study found that certain biomarkers can be used as predictors of whether someone will be diagnosed with NAFLD. The best predictive biomarkers of NAFLD were GGT > 40 U/L and HDL-C < 1.39 mmol/L, although other markers, platelet crit < 0.220%, TG > 1.7mmol/L, and AST > 33 U/L, were also good diagnostic markers of NAFLD. For the most reliable diagnosis of NAFLD, we would recommend using these five markers in combination, although platelet crit is not part of all common blood panels, which may limit its usefulness, and the association between NAFLD and platelet crit is a completely novel finding for our paper so should be validated by repeating this analysis with other datasets before use as part of NAFLD diagnostic criteria. The other four markers have been found to be

consistently altered in NAFLD in previous studies, aligning with our results. In addition to this, our results showed that NAFLD severely upsets the balance of blood composition with low platelet count, low platelet crit, elevated WBC count, raised RBC count, and increased RDW, although these parameters would not be reliable for NAFLD diagnosis. Although the literature has shown that AST:ALT ratio and ALP may be good indicators of other forms of liver damage, including hepatitis C and alcoholic fatty liver disease, in this study we found that they are not accurate predictors for NAFLD, although this does not rule out their use in clinically advanced stages of NAFLD.

Many of the most significant associations between biomarkers, diet, and whole blood count features in the NAFLD participants involve biomarkers of CKD, reiterating the link between the two diseases. This suggests that CKD patients should be screened for NAFLD and vice versa as, although there is an association found in this research, more research into biochemical pathway underpinning these association are needed to determine a causal relationship. This data suggests that both the presence and severity of NAFLD and NASH are associated with an increase in risk for CKD. Other studies on NAFLD and CKD support our results and the close associations between the two diseases (20, 42, 50, 51).

Many of our dietary association results are supported by the literature, although several of our findings were completely novel, including the differing effects of ground and instant coffee types on markers of NAFLD (25, 44, 56). The role of diet in NAFLD is still being understood and the lack of targeted pharmaceutical therapies leaves lifestyle intervention as the current most viable form of treatment for the rapidly

accelerating fatty liver epidemic. Current dietary advice for NAFLD focuses on the reduction of saturated fats, carbohydrates, and sugars (25, 44, 56). In our study, increased fibre intake through consumption of cereal, wholegrain bread, and fruit and vegetables decreased markers which have been associated in the past with NAFLD pathogenesis and markers that we have identified in this paper as significantly different between NAFLD cohorts. There is potential for this relationship to be mediated by the gut microbiome, but this cannot be confirmed from this study alone and further research is needed.

Future directions of research should focus on diet and the gut microbiome in NAFLD and understanding the biochemical interactions between different aspects of NAFLD, such as inflammation and lipotoxicity. Some whole blood count features should be looked at in more in depth, such as platelets and reticulocytes, as our findings for these groups were completely novel and these results need validation in other cohorts. Additionally, future research looking at NAFLD and diet could also look at physical activity, which is a known modulator of NAFLD risk and the gut microbiome (58-60). Data which could be highly useful in informing future research, such as cytokines, lymphocyte subtyping, and gut microbiome data, is missing from the UK Biobank. The UK Biobank is a highly useful resource for NAFLD research, so further research into NAFLD using the UK Biobank could combine with other cohorts which contain these missing variables, but there are limitations with this such as potential data loss, different formatting of variables, and inconstancy in data collection across cohorts. As we found several novel findings in this study and have proposed several key biomarkers which could be used as predictors of NAFLD or as part of NAFLD

diagnostic criteria, experimental validation of these findings in another cohort would be prudent.

The field of hepatology is rapidly working to combat the emergence of the NAFLD epidemic, and further research is needed into the biochemical mechanisms underpinning the hallmark features of NAFLD and the connection between the different cardinal features instead of researching each feature in isolation as the multifactorial nature of NAFLD demands.

List of References

1. Jones JG. Hepatic glucose and lipid metabolism. Diabetologia. 2016;59(6):1098-103.
2. Pouwels S, Sakran N, Graham Y, Leal A, Pintar T, Yang W, et al. Non-alcoholic fatty liver disease (NAFLD): a review of pathophysiology, clinical management and effects of weight loss. BMC Endocr Disord. 2022;22(1):63.
3. Estes C, Razavi H, Loomba R, Younossi Z, Sanyal AJ. Modeling the epidemic of nonalcoholic fatty liver disease demonstrates an exponential increase in burden of disease. Hepatology. 2018;67(1):123-33.
4. Martinou E, Pericleous M, Stefanova I, Kaur V, Angelidi AM. Diagnostic Modalities of Non-Alcoholic Fatty Liver Disease: From Biochemical Biomarkers to Multi-Omics Non-Invasive Approaches. Diagnostics (Basel). 2022;12(2).
5. Wree A, Broderick L, Canbay A, Hoffman HM, Feldstein AE. From NAFLD to NASH to cirrhosis-new insights into disease mechanisms. Nat Rev Gastroenterol Hepatol. 2013;10(11):627-36.
6. Del Campo JA, Gallego-Duran R, Gallego P, Grande L. Genetic and Epigenetic Regulation in Nonalcoholic Fatty Liver Disease (NAFLD). Int J Mol Sci. 2018;19(3).
7. Teramoto T, Shinohara T, Takiyama A. Computer-aided classification of hepatocellular ballooning in liver biopsies from patients with NASH using persistent homology. Comput Methods Programs Biomed. 2020;195:105614.
8. Lonardo A, Nascimbeni F, Ballestri S, Fairweather D, Win S, Than TA, et al. Sex Differences in Nonalcoholic Fatty Liver Disease: State of the Art and Identification of Research Gaps. Hepatology. 2019;70(4):1457-69.
9. Papagianni M, Sofogianni A, Tziomalos K. Non-invasive methods for the diagnosis of nonalcoholic fatty liver disease. World J Hepatol. 2015;7(4):638-48.
10. Chiarioni G, Popa SL, Dalbeni A, Senore C, Leucuta DC, Baroni L, et al. Vegan Diet Advice Might Benefit Liver Enzymes in Nonalcoholic Fatty Liver Disease: an Open Observational Pilot Study. J Gastrointestin Liver Dis. 2021;30(1):81-7.
11. Vos MB. Nutrition, nonalcoholic fatty liver disease and the microbiome: recent progress in the field. Curr Opin Lipidol. 2014;25(1):61-6.
12. Tahan V, Canbakan B, Balci H, Dane F, Akin H, Can G, et al. Serum gamma-glutamyltranspeptidase distinguishes non-alcoholic fatty liver disease at high risk. Hepatogastroenterology. 2008;55(85):1433-8.
13. Banderas DZ, Escobedo J, Gonzalez E, Liceaga MG, Ramirez JC, Castro MG. gamma-Glutamyl transferase: a marker of nonalcoholic fatty liver disease in patients with the metabolic syndrome. Eur J Gastroenterol Hepatol. 2012;24(7):805-10.
14. Fujii H, Doi H, Ko T, Fukuma T, Kadono T, Asaeda K, et al. Frequently abnormal serum gamma-glutamyl transferase activity is associated with future development of fatty liver: a retrospective cohort study. BMC Gastroenterol. 2020;20(1):217.
15. Lala V, Goyal A, Minter DA. Liver Function Tests. StatPearls. Treasure Island (FL)2022.
16. Kalas MA, Chavez L, Leon M, Taweesedt PT, Surani S. Abnormal liver enzymes: A review for clinicians. World J Hepatol. 2021;13(11):1688-98.
17. Kumar R, Porwal YC, Dev N, Kumar P, Chakravarthy S, Kumawat A. Association of high-sensitivity C-reactive protein (hs-CRP) with non-alcoholic fatty liver disease (NAFLD) in Asian Indians: A cross-sectional study. J Family Med Prim Care. 2020;9(1):390-4.
18. Sattar N, Forrest E, Preiss D. Non-alcoholic fatty liver disease. BMJ. 2014;349:g4596.
19. Kozlova MA, Kirillov YA, Makartseva LA, Chernov I, Areshidze DA. Morphofunctional State and Circadian Rhythms of the Liver under the Influence of Chronic Alcohol Intoxication and Constant Lighting. Int J Mol Sci. 2021;22(23).
20. Byrne CD, Targher G. NAFLD as a driver of chronic kidney disease. J Hepatol. 2020;72(4):785-801.
21. Meroni M, Longo M, Rustichelli A, Dongiovanni P. Nutrition and Genetics in NAFLD: The Perfect Binomium. Int J Mol Sci. 2020;21(8).
22. Schneider CV, Zandvakili I, Thaiss CA, Schneider KM. Physical activity is associated with reduced risk of liver disease in the prospective UK Biobank cohort. JHEP Rep. 2021;3(3):100263.
23. Khadge S, Sharp JG, Thiele GM, McGuire TR, Klassen LW, Duryee MJ, et al. Dietary omega-3 and omega-6 polyunsaturated fatty acids modulate hepatic pathology. J Nutr Biochem. 2018;52:92-102.
24. Huby T, Gautier EL. Immune cell-mediated features of non-alcoholic steatohepatitis. Nat Rev Immunol. 2021.
25. Todoric J, Di Caro G, Reibe S, Henstridge DC, Green CR, Vrbanac A, et al. Fructose stimulated de novo lipogenesis is promoted by inflammation. Nat Metab. 2020;2(10):1034-45.
26. Leung H, Long X, Ni Y, Qian L, Nychas E, Siliceo SL, et al. Risk assessment with gut microbiome and metabolite markers in NAFLD development. Sci Transl Med. 2022;14(648):eabk0855.
27. Vvedenskaya O, Rose TD, Knittelfelder O, Palladini A, Wodke JAH, Schuhmann K, et al. Nonalcoholic fatty liver disease stratification by liver lipidomics. J Lipid Res. 2021;62:100104.
28. Yoo JH, Kang M, Kim G, Hur KY, Kim JH, Sinn DH, et al. Mean and visit-to-visit variability of glycated hemoglobin, and the risk of non-alcoholic fatty liver disease. J Diabetes Investig. 2021;12(7):1252-62.
29. Jang BK. Elevated serum bilirubin levels are inversely associated with nonalcoholic fatty liver disease. Clin Mol Hepatol. 2012;18(4):357-9.
30. Luo L, An P, Jia X, Yue X, Zheng S, Liu S, et al. Genetically Regulated Bilirubin and Risk of Non-alcoholic Fatty Liver Disease: A Mendelian Randomization Study. Front Genet. 2018;9:662.
31. Loomba R, Friedman SL, Shulman GI. Mechanisms and disease consequences of nonalcoholic fatty liver disease. Cell. 2021;184(10):2537-64.

32. Kazankov K, Jorgensen SMD, Thomsen KL, Moller HJ, Vilstrup H, George J, et al. The role of macrophages in nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. Nat Rev Gastroenterol Hepatol. 2019;16(3):145-59.
33. Cha JY, Kim DH, Chun KH. The role of hepatic macrophages in nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. Lab Anim Res. 2018;34(4):133-9.
34. Brenner C, Galluzzi L, Kepp O, Kroemer G. Decoding cell death signals in liver inflammation. J Hepatol. 2013;59(3):583-94.
35. Herrero-Cervera A, Soehnlein O, Kenne E. Neutrophils in chronic inflammatory diseases. Cell Mol Immunol. 2022;19(2):177-91.
36. Mridha AR, Wree A, Robertson AAB, Yeh MM, Johnson CD, Van Rooyen DM, et al. NLRP3 inflammasome blockade reduces liver inflammation and fibrosis in experimental NASH in mice. J Hepatol. 2017;66(5):1037-46.
37. Ma C, Kesarwala AH, Eggert T, Medina-Echeverz J, Kleiner DE, Jin P, et al. NAFLD causes selective CD4(+) T lymphocyte loss and promotes hepatocarcinogenesis. Nature. 2016;531(7593):253-7.
38. Barrow F, Khan S, Wang H, Revelo XS. The Emerging Role of B Cells in the Pathogenesis of NAFLD. Hepatology. 2021;74(4):2277-86.
39. Chackelevicius CM, Gambaro SE, Tiribelli C, Rosso N. Th17 involvement in nonalcoholic fatty liver disease progression to non-alcoholic steatohepatitis. World J Gastroenterol. 2016;22(41):9096-103.
40. Rau M, Schilling AK, Meertens J, Hering I, Weiss J, Jurowich C, et al. Progression from Nonalcoholic Fatty Liver to Nonalcoholic Steatohepatitis Is Marked by a Higher Frequency of Th17 Cells in the Liver and an Increased Th17/Resting Regulatory T Cell Ratio in Peripheral Blood and in the Liver. J Immunol. 2016;196(1):97-105.
41. Gomes AL, Teijeiro A, Buren S, Tummala KS, Yilmaz M, Waisman A, et al. Metabolic Inflammation-Associated IL-17A Causes Non-alcoholic Steatohepatitis and Hepatocellular Carcinoma. Cancer Cell. 2016;30(1):161-75.
42. Musso G, Gambino R, Tabibian JH, Ekstedt M, Kechagias S, Hamaguchi M, et al. Association of non-alcoholic fatty liver disease with chronic kidney disease: a systematic review and meta-analysis. PLoS Med. 2014;11(7):e1001680.
43. Younossi ZM, Golabi P, de Avila L, Paik JM, Srishord M, Fukui N, et al. The global epidemiology of NAFLD and NASH in patients with type 2 diabetes: A systematic review and meta-analysis. J Hepatol. 2019;71(4):793-801.
44. Miller EF. Nutrition Management Strategies for Nonalcoholic Fatty Liver Disease: Treatment and Prevention. Clin Liver Dis (Hoboken). 2020;15(4):144-8.
45. Tanase DM, Gosav EM, Costea CF, Ciocoiu M, Lacatusu CM, Maranduca MA, et al. The Intricate Relationship between Type 2 Diabetes Mellitus (T2DM), Insulin Resistance (IR), and Nonalcoholic Fatty Liver Disease (NAFLD). J Diabetes Res. 2020;2020:3920196.
46. Wang B, Li M, Zhao Z, Wang S, Lu J, Chen Y, et al. Glycemic Measures and Development and Resolution of Nonalcoholic Fatty Liver Disease in Nondiabetic Individuals. J Clin Endocrinol Metab. 2020;105(5).
47. Sookoian S, Pirola CJ. Meta-analysis of the influence of I148M variant of patatin-like phospholipase domain containing 3 gene (PNPLA3) on the susceptibility and histological severity of nonalcoholic fatty liver disease. Hepatology. 2011;53(6):1883-94.
48. Du X, DeForest N, Majithia AR. Human Genetics to Identify Therapeutic Targets for NAFLD: Challenges and Opportunities. Front Endocrinol (Lausanne). 2021;12:777075.
49. Mann JP, Anstee QM. NAFLD: PNPLA3 and obesity: a synergistic relationship in NAFLD. Nat Rev Gastroenterol Hepatol. 2017;14(9):506-7.
50. Mantovani A, Petracca G, Beatrice G, Csermely A, Lonardo A, Schattenberg JM, et al. Non-alcoholic fatty liver disease and risk of incident chronic kidney disease: an updated meta-analysis. Gut. 2022;71(1):156-62.
51. Waheed S, Matsushita K, Sang Y, Hoogeveen R, Ballantyne C, Coresh J, et al. Combined association of albuminuria and cystatin C-based estimated GFR with mortality, coronary heart disease, and heart failure outcomes: the Atherosclerosis Risk in Communities (ARIC) Study. Am J Kidney Dis. 2012;60(2):207-16.
52. Musso G, Cassader M, Cohney S, Pinach S, Saba F, Gambino R. Emerging Liver-Kidney Interactions in Nonalcoholic Fatty Liver Disease. Trends Mol Med. 2015;21(10):645-62.
53. Pan B, Wan X, Ma M, Cao C. Complement C3 and Nonalcoholic Fatty Liver Disease in Chronic Kidney Disease Patients: A Pilot Study. Kidney Blood Press Res. 2020;45(1):61-9.
54. Liu Z, Suo C, Zhao R, Yuan H, Jin L, Zhang T, et al. Genetic predisposition, lifestyle risk, and obesity associate with the progression of nonalcoholic fatty liver disease. Dig Liver Dis. 2021.
55. Wong VW, Singal AK. Emerging medical therapies for non-alcoholic fatty liver disease and for alcoholic hepatitis. Transl Gastroenterol Hepatol. 2019;4:53.
56. Biolato M, Manca F, Marrone G, Cefalo C, Racco S, Miggiano GA, et al. Intestinal permeability after Mediterranean diet and low-fat diet in non-alcoholic fatty liver disease. World J Gastroenterol. 2019;25(4):509-20.
57. Waniek S, di Giuseppe R, Plachta-Danielzik S, Ratjen I, Jacobs G, Koch M, et al. Association of Vitamin E Levels with Metabolic Syndrome, and MRI-Derived Body Fat Volumes and Liver Fat Content. Nutrients. 2017;9(10).
58. van Kleef LA, Hofman A, Voortman T, de Knegt RJ. Objectively Measured Physical Activity Is Inversely Associated With Nonalcoholic Fatty Liver Disease: The Rotterdam Study. Am J Gastroenterol. 2022;117(2):311-8.
59. Orci LA, Gariani K, Oldani G, Delaune V, Morel P, Toso C. Exercise-based Interventions for Nonalcoholic Fatty Liver Disease: A Meta-analysis and Meta-regression. Clin Gastroenterol Hepatol. 2016;14(10):1398-411.
60. Semmler G, Datz C, Reiberger T, Trauner M. Diet and exercise in NAFLD/NASH: Beyond the obvious. Liver Int. 2021;41(10):2249-68.

61. Schnurr TM, Katz SF, Justesen JM, O'Sullivan JW, Saliba-Gustafsson P, Assimes TL, et al. Interactions of physical activity, muscular fitness, adiposity, and genetic risk for NAFLD. Hepatol Commun. 2022;6(7):1516-26.
62. Sherry AP, Willis SA, Yates T, Johnson W, Razieh C, Sargeant JA, et al. Physical activity is inversely associated with hepatic fibro-inflammation: A population-based cohort study using UK Biobank data. JHEP Rep. 2023;5(1):100622.
63. Gelli C, Tarocchi M, Abenavoli L, Di Renzo L, Galli A, De Lorenzo A. Effect of a counseling-supported treatment with the Mediterranean diet and physical activity on the severity of the non-alcoholic fatty liver disease. World J Gastroenterol. 2017;23(17):3150-62.
64. Franco I, Bianco A, Mirizzi A, Campanella A, Bonfiglio C, Sorino P, et al. Physical Activity and Low Glycemic Index Mediterranean Diet: Main and Modification Effects on NAFLD Score. Results from a Randomized Clinical Trial. Nutrients. 2020;13(1).
65. Abul-Husn NS, Cheng X, Li AH, Xin Y, Schurmann C, Stevis P, et al. A Protein-Truncating HSD17B13 Variant and Protection from Chronic Liver Disease. N Engl J Med. 2018;378(12):1096-106.
66. Vujkovic M, Ramdas S, Lorenz KM, Guo X, Darlay R, Cordell HJ, et al. A multiancestry genome-wide association study of unexplained chronic ALT elevation as a proxy for nonalcoholic fatty liver disease with histological and radiological validation. Nat Genet. 2022;54(6):761-71.
67. Fairfield CJ, Drake TM, Pius R, Bretherick AD, Campbell A, Clark DW, et al. Genome-Wide Association Study of NAFLD Using Electronic Health Records. Hepatol Commun. 2021.
68. Jamialahmadi O, Mancina RM, Ciociola E, Tavaglione F, Luukkonen PK, Baselli G, et al. Exome-Wide Association Study on Alanine Aminotransferase Identifies Sequence Variants in the GPAM and APOE Associated With Fatty Liver Disease. Gastroenterology. 2021;160(5):1634-46.e7.
69. Meroni M, Longo M, Paolini E, Alisi A, Miele L, De Caro ER, et al. The rs599839 A>G Variant Disentangles Cardiovascular Risk and Hepatocellular Carcinoma in NAFLD Patients. Cancers (Basel). 2021;13(8).
70. De Vincentis A, Tavaglione F, Jamialahmadi O, Picardi A, Antonelli Incalzi R, Valenti L, et al. A Polygenic Risk Score to Refine Risk Stratification and Prediction for Severe Liver Disease by Clinical Fibrosis Scores. Clin Gastroenterol Hepatol. 2021.
71. Gellert-Kristensen H, Richardson TG, Davey Smith G, Nordestgaard BG, Tybjaerg-Hansen A, Stender S. Combined Effect of PNPLA3, TM6SF2, and HSD17B13 Variants on Risk of Cirrhosis and Hepatocellular Carcinoma in the General Population. Hepatology. 2020;72(3):845-56.
72. Bianco C, Jamialahmadi O, Pelusi S, Baselli G, Dongiovanni P, Zanoni I, et al. Non-invasive stratification of hepatocellular carcinoma risk in non-alcoholic fatty liver using polygenic risk scores. J Hepatol. 2021;74(4):775-82.
73. Liu Z, Suo C, Shi O, Lin C, Zhao R, Yuan H, et al. The Health Impact of MAFLD, a Novel Disease Cluster of NAFLD, Is Amplified by the Integrated Effect of Fatty Liver Disease-Related Genetic Variants. Clin Gastroenterol Hepatol. 2020.
74. Dongiovanni P, Meroni M, Baselli G, Mancina RM, Ruscica M, Longo M, et al. gene variation bridges atherogenic dyslipidemia with hepatic inflammation in NAFLD patients. J Lipid Res. 2019;60(6):1144-53.
75. Ganel L, Chen L, Christ R, Vangipurapu J, Young E, Das I, et al. Mitochondrial genome copy number measured by DNA sequencing in human blood is strongly associated with metabolic traits via cell-type composition differences. Hum Genomics. 2021;15(1):34.
76. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018;562(7726):203-9.
77. Tebani A, Gummesson A, Zhong W, Koistinen IS, Lakshmikanth T, Olsson LM, et al. Integration of molecular profiles in a longitudinal wellness profiling cohort. Nat Commun. 2020;11(1):4487.
78. Singhal R, Cardoso VR, Wiggins T, Super J, Ludwig C, Gkoutos GV, et al. 30-day morbidity and mortality of sleeve gastrectomy, Roux-en-Y gastric bypass and one anastomosis gastric bypass: a propensity score-matched analysis of the GENEVA data. Int J Obes (Lond). 2022;46(4):750-7.
79. Mahalanobis PC. On the generalized distance in statistics In: Proceedings of the National Institute of Science, India. 1936;2(1):49-55.
80. Ho D, Imai K, King G, Stuart EA. MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. Journal of Statistical Software. 2011;42(8):1 - 28.
81. Linge J, Ekstedt M, Dahlqvist Leinhard O. Adverse muscle composition is linked to poor functional performance and metabolic comorbidities in NAFLD. JHEP Rep. 2021;3(1):100197.
82. Censin JC, Peters SAE, Bovijn J, Ferreira T, Pulit SL, Mägi R, et al. Causal relationships between obesity and the leading causes of death in women and men. PLoS Genet. 2019;15(10):e1008405.
83. Liu Z, Zhang Y, Graham S, Wang X, Cai D, Huang M, et al. Causal relationships between NAFLD, T2D and obesity have implications for disease subphenotyping. J Hepatol. 2020;73(2):263-76.
84. Zou B, Yeo YH, Cheung R, Ingelsson E, Nguyen MH. Fatty Liver Index and Development of Cardiovascular Disease: Findings from the UK Biobank. Dig Dis Sci. 2021;66(6):2092-100.
85. Inan-Eroglu E, Huang BH, Ahmadi MN, Johnson N, El-Omar EM, Stamatakis E. Joint associations of adiposity and alcohol consumption with liver disease-related morbidity and mortality risk: findings from the UK Biobank. Eur J Clin Nutr. 2021.
86. Linge J, Whitcher B, Borga M, Dahlqvist Leinhard O. Sub-phenotyping Metabolic Disorders Using Body Composition: An Individualized, Nonparametric Approach Utilizing Large Data Sets. Obesity (Silver Spring). 2019;27(7):1190-9.

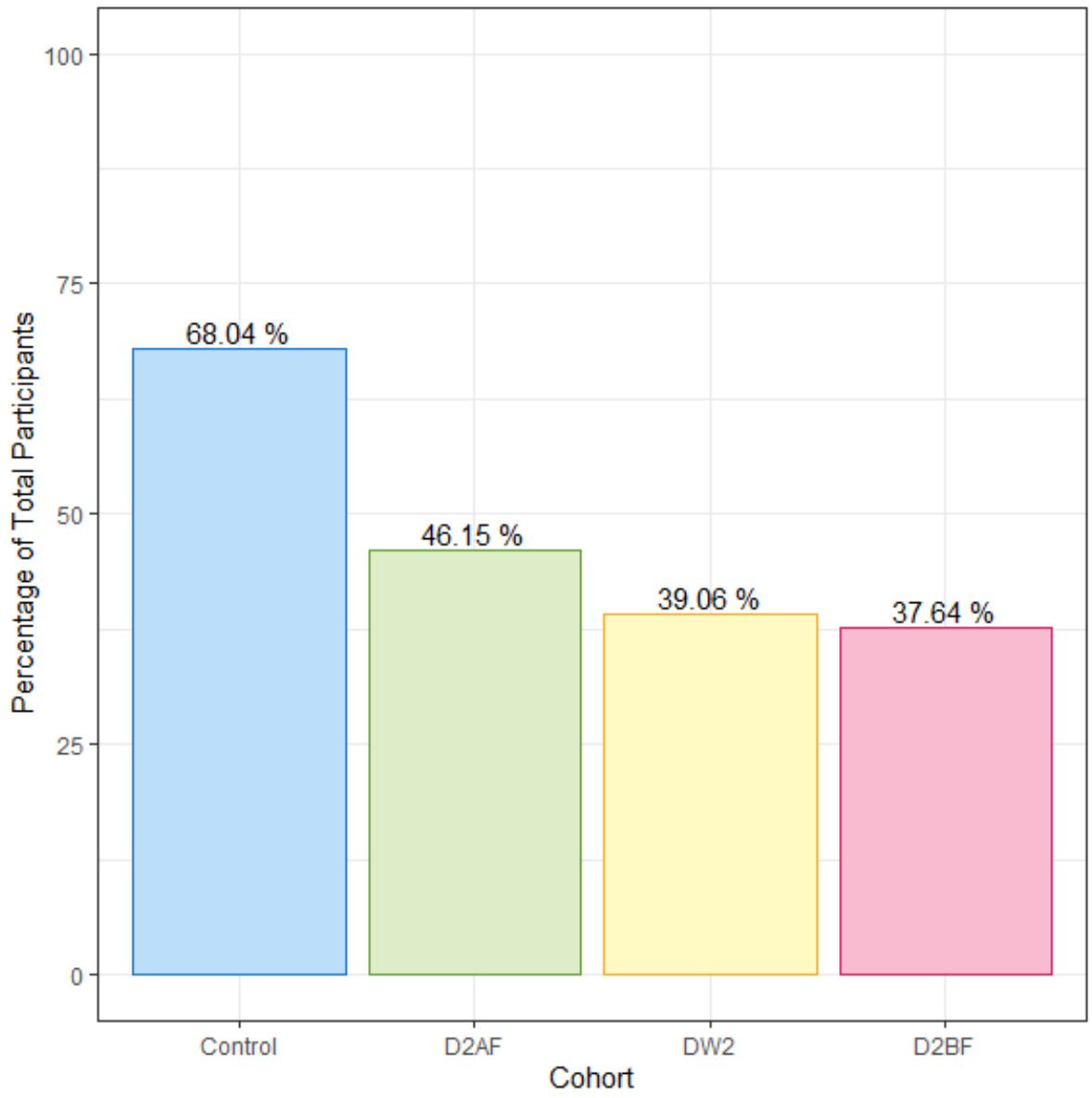
87. Zhang Y, Liu Z, Choudhury T, Cornelis MC, Liu W. Habitual coffee intake and risk for nonalcoholic fatty liver disease: a two-sample Mendelian randomization study. Eur J Nutr. 2021;60(4):1761-7.
88. Chicco D, Oneto L. Computational intelligence identifies alkaline phosphatase (ALP), alpha-fetoprotein (AFP), and hemoglobin levels as most predictive survival factors for hepatocellular carcinoma. Health Informatics J. 2021;27(1):1460458220984205.
89. Ng CH, Teng ML, Chew NW, Chan KE, Yong JN, Quek J, et al. Statins decrease overall mortality and cancer related mortality but are underutilized in NAFLD: a longitudinal analysis of 12,538 individuals. Expert Rev Gastroenterol Hepatol. 2022.
90. Buchwald H, O'Dea TJ, Menchaca HJ, Michalek VN, Rohde TD. Effect of plasma cholesterol on red blood cell oxygen transport. Clin Exp Pharmacol Physiol. 2000;27(12):951-5.
91. Simon J, Ouro A, Ala-Ibanibo L, Presa N, Delgado TC, Martinez-Chantar ML. Sphingolipids in Non-Alcoholic Fatty Liver Disease and Hepatocellular Carcinoma: Ceramide Turnover. Int J Mol Sci. 2019;21(1).
92. Tanase DM, Gosav EM, Petrov D, Jucan AE, Lacatusu CM, Floria M, et al. Involvement of Ceramides in Non-Alcoholic Fatty Liver Disease (NAFLD) Atherosclerosis (ATS) Development: Mechanisms and Therapeutic Targets. Diagnostics (Basel). 2021;11(11).
93. Jiang M, Li C, Liu Q, Wang A, Lei M. Inhibiting Ceramide Synthesis Attenuates Hepatic Steatosis and Fibrosis in Rats With Non-alcoholic Fatty Liver Disease. Front Endocrinol (Lausanne). 2019;10:665.
94. Lee J, Yoon K, Ryu S, Chang Y, Kim HR. High-normal levels of hs-CRP predict the development of non-alcoholic fatty liver in healthy men. PLoS One. 2017;12(2):e0172666.
95. Duan Y, Pan X, Luo J, Xiao X, Li J, Bestman PL, et al. Association of Inflammatory Cytokines With Non-Alcoholic Fatty Liver Disease. Front Immunol. 2022;13:880298.
96. Cobbina E, Akhlaghi F. Non-alcoholic fatty liver disease (NAFLD) - pathogenesis, classification, and effect on drug metabolizing enzymes and transporters. Drug Metab Rev. 2017;49(2):197-211.
97. Weng SY, Wang X, Vijayan S, Tang Y, Kim YO, Padberg K, et al. IL-4 Receptor Alpha Signaling through Macrophages Differentially Regulates Liver Fibrosis Progression and Reversal. EBioMedicine. 2018;29:92-103.
98. Nguyen-Lefebvre AT, Horuzsko A. Kupffer Cell Metabolism and Function. J Enzymol Metab. 2015;1(1).
99. Cairoli V, De Matteo E, Rios D, Lezama C, Galoppo M, Casciato P, et al. Hepatic lymphocytes involved in the pathogenesis of pediatric and adult non-alcoholic fatty liver disease. Sci Rep. 2021;11(1):5129.
100. Tang W, Zhou J, Yang W, Feng Y, Wu H, Mok MTS, et al. Aberrant cholesterol metabolic signaling impairs antitumor immunosurveillance through natural killer T cell dysfunction in obese liver. Cell Mol Immunol. 2022;19(7):834-47.
101. Kang S, Kim E, Cho H, Kim DJ, Kim HC, Jung SJ. Associations between non-alcoholic fatty liver disease and cognitive impairment and the effect modification of inflammation. Sci Rep. 2022;12(1):12614.
102. Yang W, Huang H, Wang Y, Yu X, Yang Z. High red blood cell distribution width is closely associated with nonalcoholic fatty liver disease. Eur J Gastroenterol Hepatol. 2014;26(2):174-8.
103. Papadopoulos C, Tentes I, Anagnostopoulos K. Red Blood Cell Dysfunction in Non-Alcoholic Fatty Liver Disease: Marker and Mediator of Molecular Mechanisms. Maedica (Bucur). 2020;15(4):513-6.
104. Jiang Y, Zeng J, Chen B. Hemoglobin combined with triglyceride and ferritin in predicting non-alcoholic fatty liver. J Gastroenterol Hepatol. 2014;29(7):1508-14.
105. Farzanegi P, Dana A, Ebrahimpoor Z, Asadi M, Azarbayjani MA. Mechanisms of beneficial effects of exercise training on non-alcoholic fatty liver disease (NAFLD): Roles of oxidative stress and inflammation. Eur J Sport Sci. 2019;19(7):994-1003.
106. Pasdar Y, Oubari F, Zarif MN, Abbasi M, Pourmahmoudi A, Hosseinikia M. Effects of Quercetin Supplementation on Hematological Parameters in Non-Alcoholic Fatty Liver Disease: a Randomized, Double-Blind, Placebo-Controlled Pilot Study. Clin Nutr Res. 2020;9(1):11-9.
107. Shimomura Y, Takaki A, Wada N, Yasunaka T, Ikeda F, Maruyama T, et al. The Serum Oxidative/Anti-oxidative Stress Balance Becomes Dysregulated in Patients with Non-alcoholic Steatohepatitis Associated with Hepatocellular Carcinoma. Intern Med. 2017;56(3):243-51.
108. Mohanty JG, Nagababu E, Rifkind JM. Red blood cell oxidative stress impairs oxygen delivery and induces red blood cell aging. Front Physiol. 2014;5:84.
109. Chen Z, Tian R, She Z, Cai J, Li H. Role of oxidative stress in the pathogenesis of nonalcoholic fatty liver disease. Free Radic Biol Med. 2020;152:116-41.
110. Giorgio V, Mosca A, Alterio A, Alisi A, Grieco A, Nobili V, et al. Elevated Hemoglobin Level Is Associated With Advanced Fibrosis in Pediatric Nonalcoholic Fatty Liver Disease. J Pediatr Gastroenterol Nutr. 2017;65(2):150-5.
111. Zhou WJ, Yang J, Zhang G, Hu ZQ, Jiang YM, Yu F. Association between red cell distribution width-to-platelet ratio and hepatic fibrosis in nonalcoholic fatty liver disease: A cross-sectional study. Medicine (Baltimore). 2019;98(30):e16565.
112. Lopez-Trujillo MA, Olivares-Gazca JM, Cantero-Fortiz Y, Garcia-Navarrete YI, Cruz-Mora A, Olivares-Gazca JC, et al. Nonalcoholic Fatty Liver Disease and Thrombocytopenia III: Its Association With Insulin Resistance. Clin Appl Thromb Hemost. 2019;25:1076029619888694.
113. Olivares-Gazca JC, Nunez-Cortes AK, Mendez-Huerta MA, Cantero-Fortiz Y, Orea-Martinez JG, Ruiz-Arguelles GJ. More on the thrombocytopenia of the non-alcoholic fatty liver disease. Hematology. 2017;22(5):316-9.
114. Panke CL, Tovo CV, Villela-Nogueira CA, Cravo CM, Ferreira FC, Rezende GFM, et al. Evaluation of thrombocytopenia in patients with non-alcoholic fatty liver disease without cirrhosis. Ann Hepatol. 2020;19(1):88-91.

115. Lopez-Giacoman S, Madero M. Biomarkers in chronic kidney disease, from kidney function to kidney damage. World J Nephrol. 2015;4(1):57-73.
116. Tavares AR, Jr. Cystatin C versus creatinine for kidney function-based risk. N Engl J Med. 2013;369(25):2457-8.
117. Ix JH, Shlipak MG. The Promise of Tubule Biomarkers in Kidney Disease: A Review. Am J Kidney Dis. 2021;78(5):719-27.
118. Luo X, Li Y, Zhou Y, Zhang C, Li L, Luo Y, et al. Association of Non-alcoholic Fatty Liver Disease With Salt Intake and Dietary Diversity in Chinese Medical Examination Adults Aged 18-59 Years: A Cross-Sectional Study. Front Nutr. 2022;9:930316.
119. Benoit SW, Ciccia EA, Devarajan P. Cystatin C as a biomarker of chronic kidney disease: latest developments. Expert Rev Mol Diagn. 2020;20(10):1019-26.
120. Targher G, Bertolini L, Rodella S, Lippi G, Zoppini G, Chonchol M. Relationship between kidney function and liver histology in subjects with nonalcoholic steatohepatitis. Clin J Am Soc Nephrol. 2010;5(12):2166-71.
121. Park CW, Tsai NT, Wong LL. Implications of worse renal dysfunction and medical comorbidities in patients with NASH undergoing liver transplant evaluation: impact on MELD and more. Clin Transplant. 2011;25(6):E606-11.
122. He J, Mills KT, Appel LJ, Yang W, Chen J, Lee BT, et al. Urinary Sodium and Potassium Excretion and CKD Progression. J Am Soc Nephrol. 2016;27(4):1202-12.
123. van den Berg EH, Gruppen EG, Blokzijl H, Bakker SJL, Dullaart RPF. Higher Sodium Intake Assessed by 24 Hour Urinary Sodium Excretion Is Associated with Non-Alcoholic Fatty Liver Disease: The PREVEND Cohort Study. J Clin Med. 2019;8(12).
124. Madatali Abuwani A, Priyadarshini Dash S, Ganesan R, Renu K, Vellingiri B, Kandasamy S, et al. Gut microbiome and metabolic response in non-alcoholic fatty liver disease. Clin Chim Acta. 2021;523:304-14.
125. Bajaj JS, Ng SC, Schnabl B. Promises of microbiome-based therapies. J Hepatol. 2022;76(6):1379-91.
126. Zeybel M, Arif M, Li X, Altay O, Yang H, Shi M, et al. Multiomics Analysis Reveals the Impact of Microbiota on Host Metabolism in Hepatic Steatosis. Adv Sci (Weinh). 2022;9(11):e2104373.
127. Schneider KM, Mohs A, Gui W, Galvez EJC, Candels LS, Hoenicke L, et al. Imbalanced gut microbiota fuels hepatocellular carcinoma development by shaping the hepatic inflammatory microenvironment. Nat Commun. 2022;13(1):3964.
128. Zhang S, Zhao J, Xie F, He H, Johnston LJ, Dai X, et al. Dietary fiber-derived short-chain fatty acids: A potential therapeutic target to alleviate obesity-related nonalcoholic fatty liver disease. Obes Rev. 2021;22(11):e13316.
129. Loman BR, Hernandez-Saavedra D, An R, Rector RS. Prebiotic and probiotic treatment of nonalcoholic fatty liver disease: a systematic review and meta-analysis. Nutr Rev. 2018;76(11):822-39.
130. Valentini M, Piermattei A, Di Sante G, Migliara G, Delogu G, Ria F. Immunomodulation by gut microbiota: role of Toll-like receptor expressed by T cells. J Immunol Res. 2014;2014:586939.
131. Wijarnpreecha K, Thongprayoon C, Ungprasert P. Coffee consumption and risk of nonalcoholic fatty liver disease: a systematic review and meta-analysis. Eur J Gastroenterol Hepatol. 2017;29(2):e8-e12.
132. Chen YP, Lu FB, Hu YB, Xu LM, Zheng MH, Hu ED. A systematic review and a dose-response meta-analysis of coffee dose and nonalcoholic fatty liver disease. Clin Nutr. 2019;38(6):2552-7.
133. Carter JL, Lewington S, Piernas C, Bradbury K, Key TJ, Jebb SA, et al. Reproducibility of dietary intakes of macronutrients, specific food groups, and dietary patterns in 211 050 adults in the UK Biobank study. J Nutr Sci. 2019;8:e34.
134. Bradbury KE, Young HJ, Guo W, Key TJ. Dietary assessment in UK Biobank: an evaluation of the performance of the touchscreen dietary questionnaire. J Nutr Sci. 2018;7:e6.

Appendix

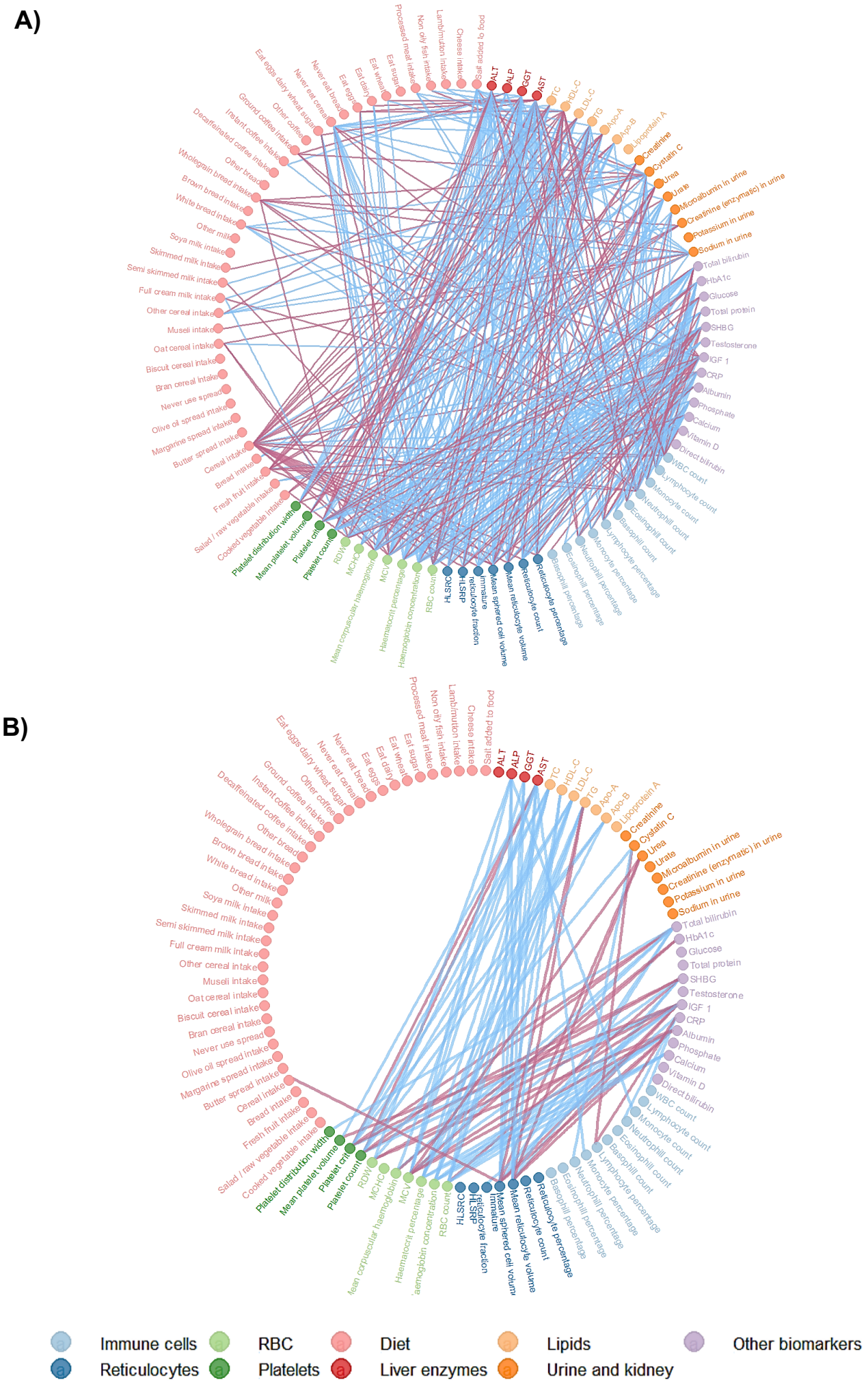
List of Figures

List of Tables

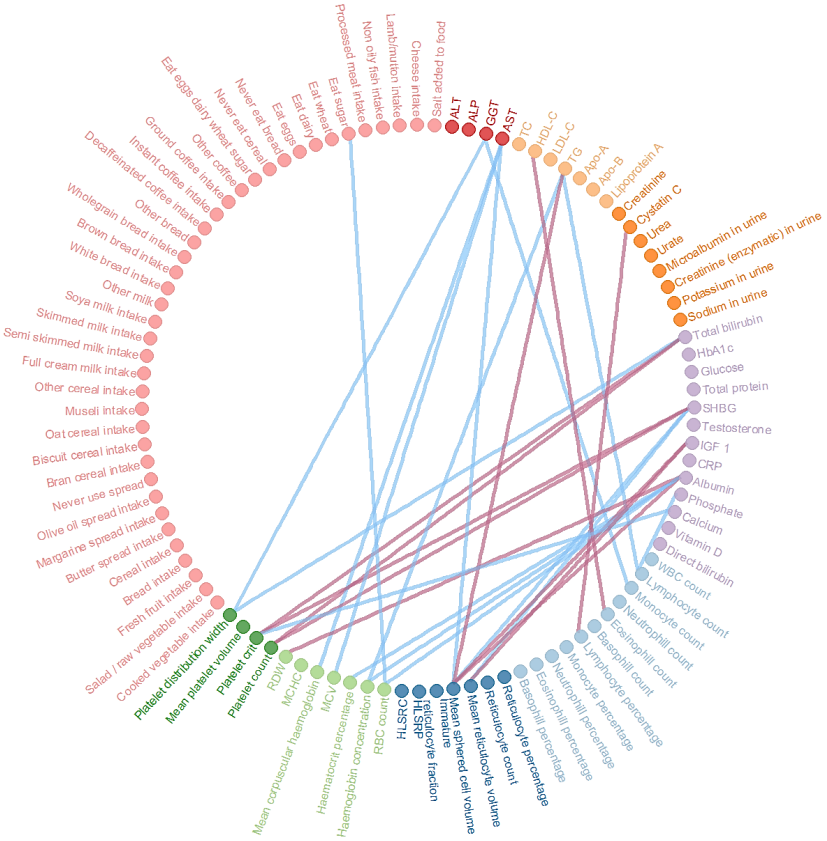


Appendix Figure 1

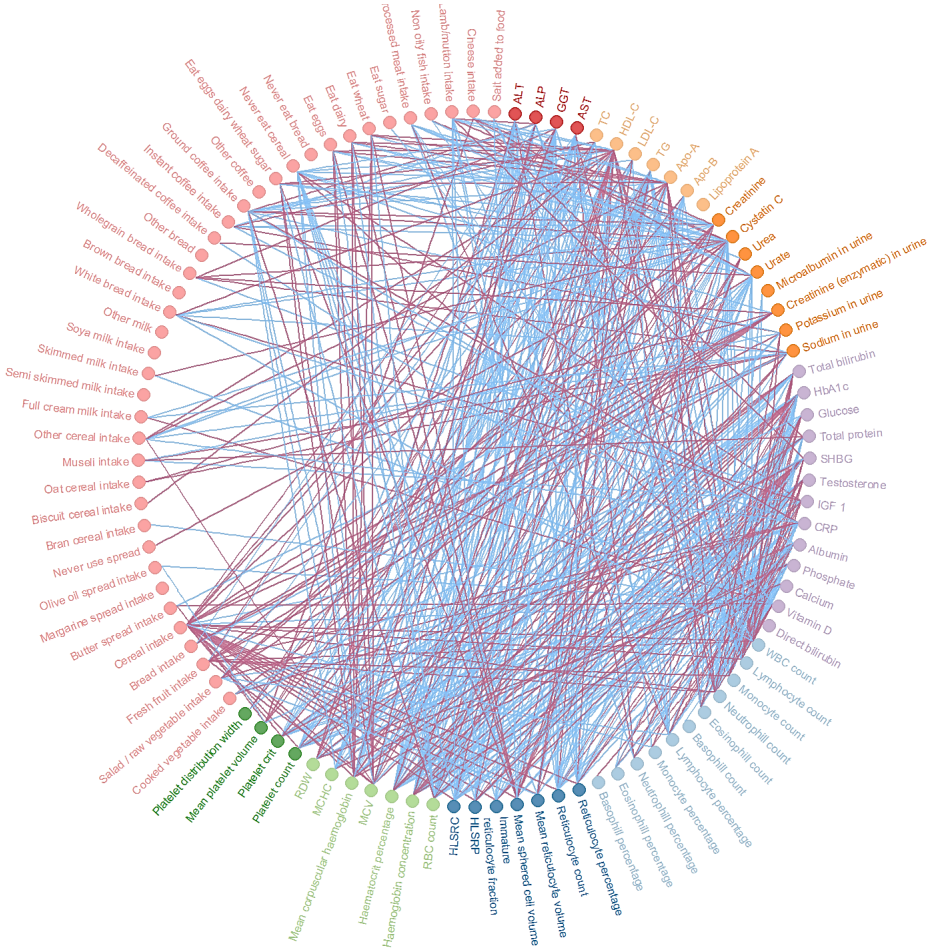
The percentage of participants from each cohort used in the final analysis after the application of the exclusion criteria. Values are expressed as percentages out of the total number of participants extracted for each respective cohort.


A)
B)
Processed meat intake
Non oily fish intake
Lamb/mutton intake
Cheese intake
Salt added to food
ALT
ALP
GGT
AST
TC
HDL-C
LDL-C
TG
Apo-A
Apo-B
Lipoprotein A
Creatinine
Cystatin C
Urea
Urate
Microalbumin in urine
Creatinine (enzymatic) in urine
Potassium in urine
Sodium in urine
Total bilirubin
HbA1c
Glucose
Total protein
SHBG
Testosterone
IGF 1
CRP
Albumin
Phosphate
Calcium
Vitamin D
Direct bilirubin
WBC count
Lymphocyte count
Monocyte count
Neutrophil count
Eosinophil count
Basophil count
Lymphocyte percentage
Monocyte percentage
Neutrophil percentage
Eosinophil percentage
Basophil percentage
Reticulocyte percentage
Reticulocyte count
Mean reticulocyte volume
Mean sphered cell volume
Immature reticulocyte fraction
HLSRC
CRSTL
RBC count
Haemoglobin concentration
Haematocrit percentage
MCV
Mean corpuscular haemoglobin
MCHC
RDW
Platelet count
Platelet crit
Mean platelet volume
Platelet distribution width
Salad / raw vegetable intake
Cooked vegetable intake
Fresh fruit intake
Bread intake
Cereal intake
Butter spread intake
Margarine spread intake
Olive oil spread intake
Never use spread
Bran cereal intake
Biscuit cereal intake
Oat cereal intake
Muesli intake
Other cereal intake
Full cream milk intake
Semi skimmed milk intake
Skimmed milk intake
Soya milk intake
Other milk
White bread intake
Brown bread intake
Wholegrain bread intake
Other bread
Decaffeinated coffee intake
Instant coffee intake
Ground coffee intake
Other coffee
Never eat cereal
Never eat bread
Eat eggs dairy wheat sugar
Eat eggs
Eat dairy
Eat wheat
Eat sugar
Immune cells
Reticulocytes
RBC
Platelets
Diet
Liver enzymes
Lipids
Urine and kidney
Other biomarkers

C)



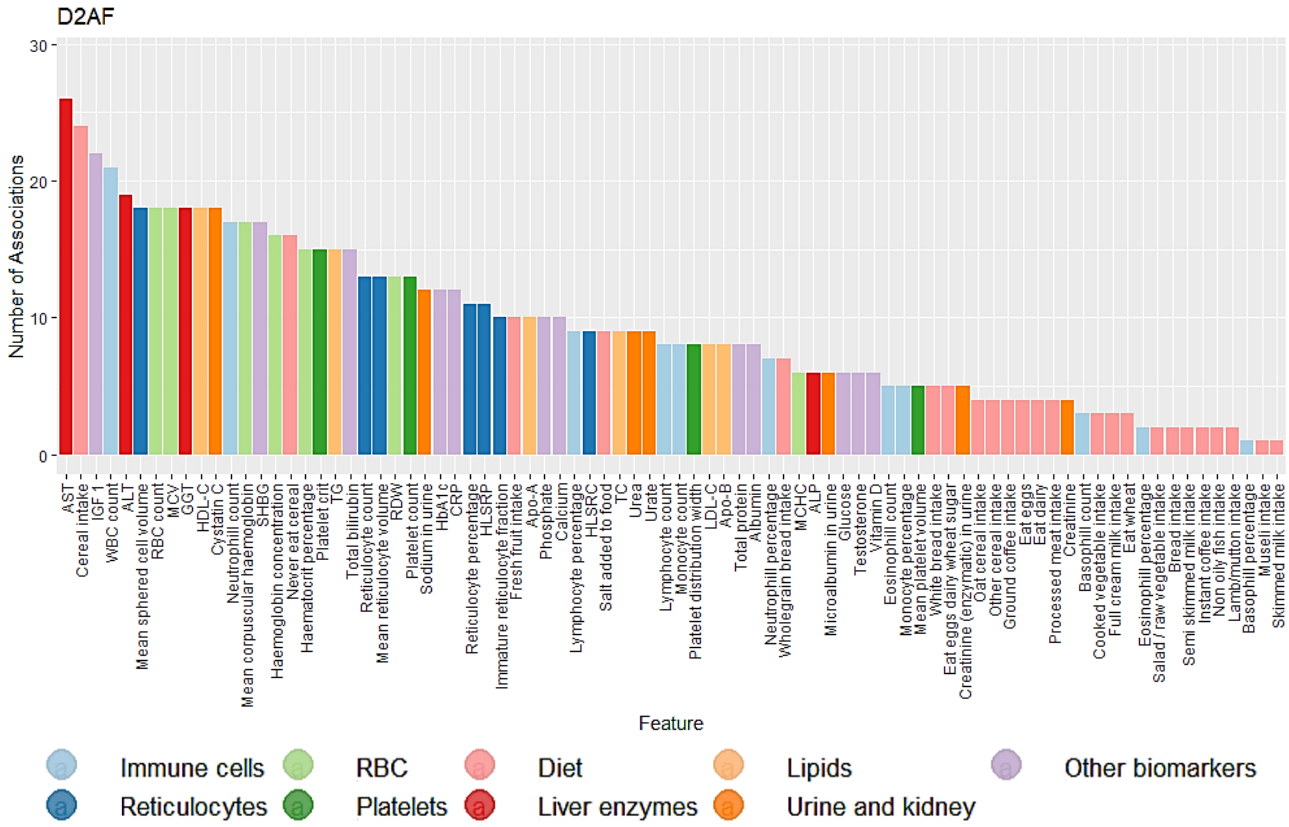
D)



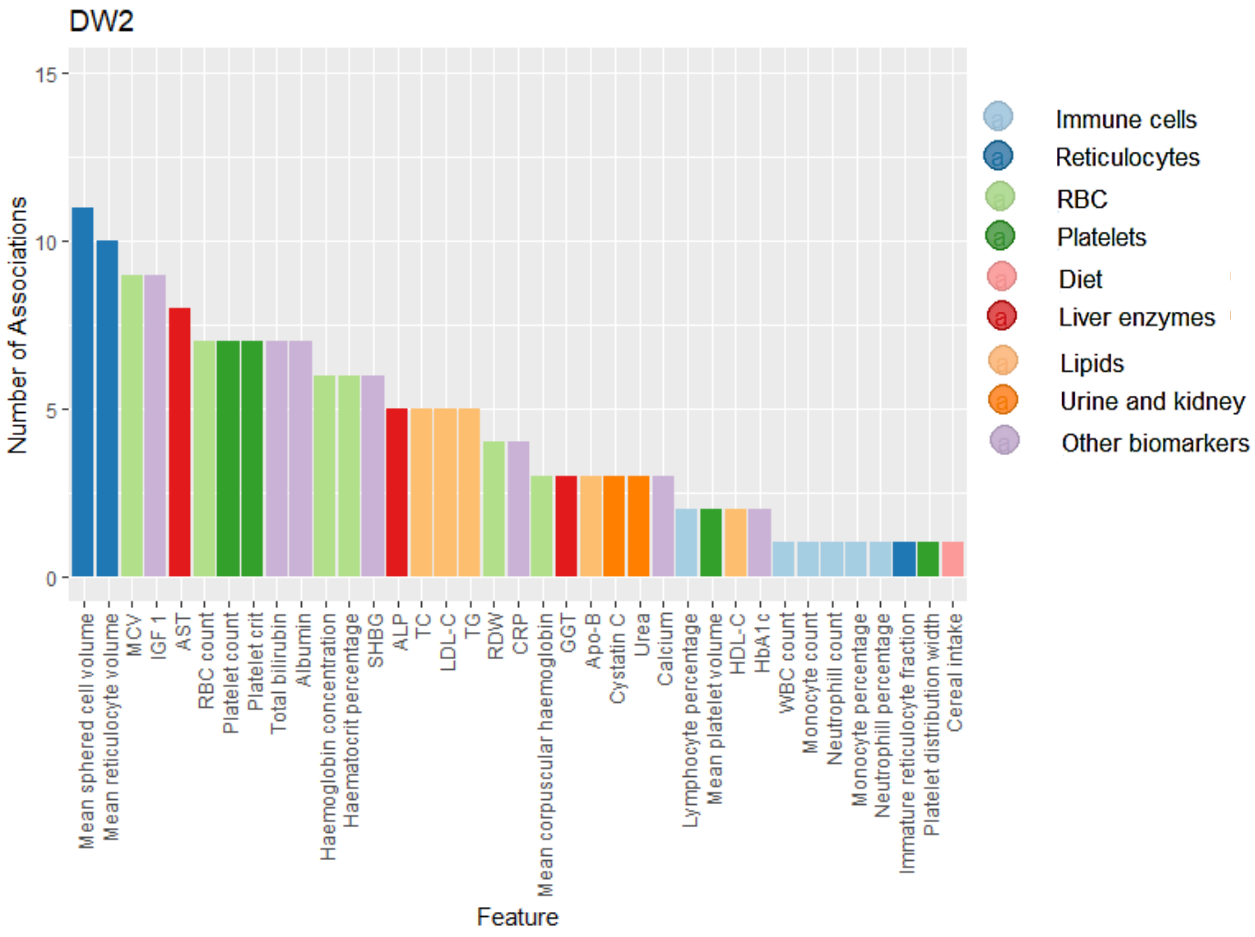
Appendix Figure 2

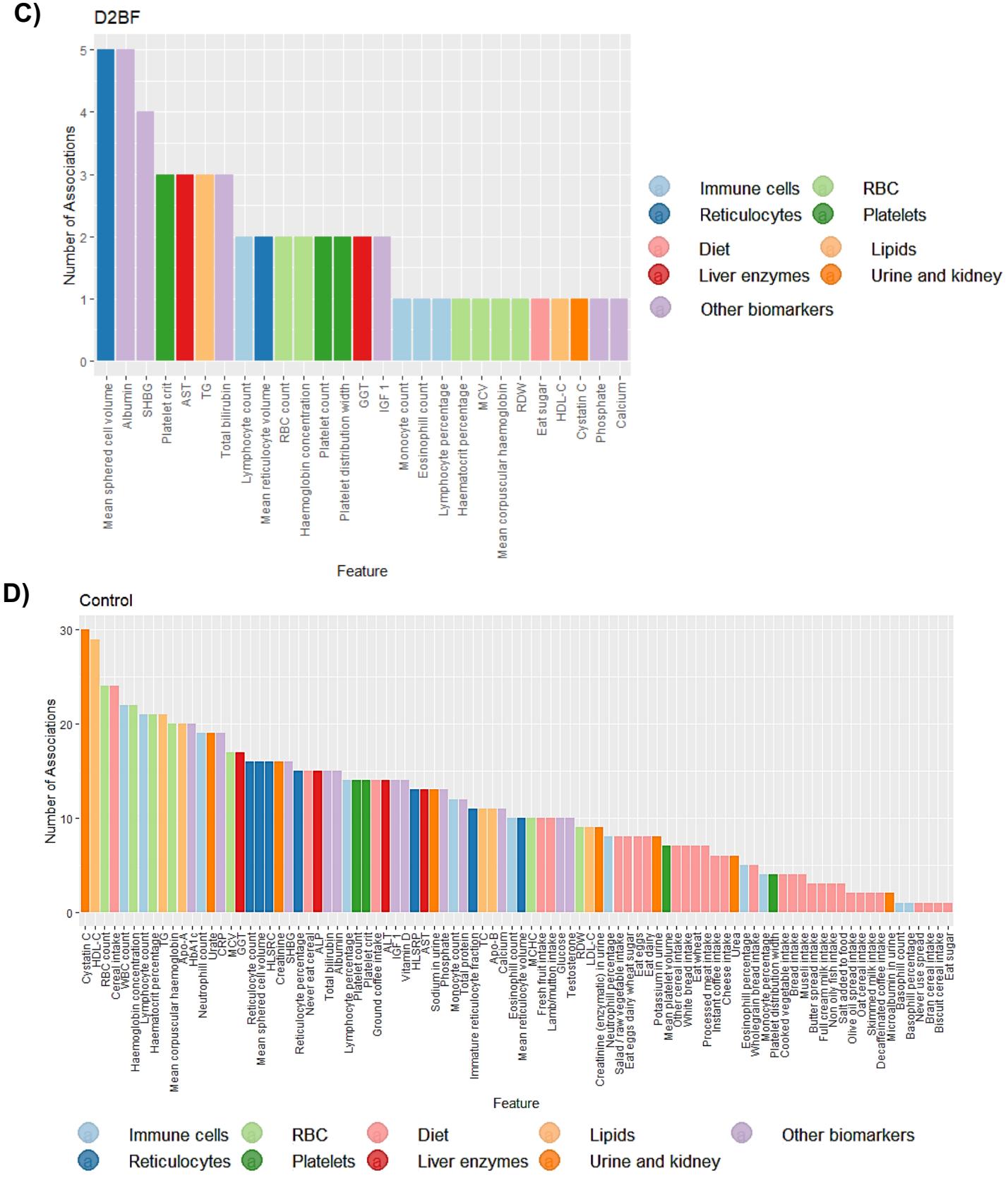
Network graphs for **A)** D2AF, **B)** DW2, **C)** D2BF, and **D)** control cohorts showcasing all associations identified. Blue lines indicate positive associations, pink indicate inverse associations.

A)


D2AF
Number of Associations
30
20
10
0
AST
Cereal intake
IGF 1
WBC count
ALT
Mean sphered cell volume
RBC count
MCV
GGT
HDL-C
Cystatin C
Neutrophil count
Mean corpuscular haemoglobin
SHBG
Haemoglobin concentration
Never eat cereal
Haematocrit percentage
Platelet crit
TG
Total bilirubin
Reticulocyte count
Mean reticulocyte volume
RDW
Platelet count
Sodium in urine
HbA1c
CRP
Reticulocyte percentage
HLSRP
Immature reticulocyte fraction
Fresh fruit intake
Apo-A
Phosphate
Calcium
Lymphocyte percentage
HLSRC
Salt added to food
TC
Urea
Urate
Lymphocyte count
Monocyte count
Platelet distribution width
LDL-C
Apo-B
Total protein
Albumin
Neutrophil percentage
Wholegrain bread intake
MCHC
ALP
Microalbumin in urine
Glucose
Testosterone
Vitamin D
Eosinophil count
Monocyte percentage
Mean platelet volume
White bread intake
Eat eggs dairy wheat sugar
Creatinine (enzymatic) in urine
Oat cereal intake
Other cereal intake
Ground coffee intake
Eat eggs
Eat dairy
Processed meat intake
Creatinine
Basophil count
Cooked vegetable intake
Full cream milk intake
Eat wheat
Eosinophil percentage
Salad / raw vegetable intake
Bread intake
Semi skimmed milk intake
Instant coffee intake
Non oily fish intake
Lamb/mutton intake
Basophil percentage
Muesli intake
Skimmed milk intake
Feature
Immune cells
Reticulocytes
RBC
Platelets
Diet
Liver enzymes
Lipids
Urine and kidney
Other biomarkers

B)


DW2
Number of Associations
15
10
5
0
Mean sphered cell volume
Mean reticulocyte volume
MCV
IGF 1
AST
RBC count
Platelet count
Platelet crit
Total bilirubin
Albumin
Haemoglobin concentration
Haematocrit percentage
SHBG
ALP
TC
LDL-C
TG
RDW
CRP
Mean corpuscular haemoglobin
GGT
Apo-B
Cystatin C
Urea
Calcium
Lymphocyte percentage
Mean platelet volume
HDL-C
HbA1c
WBC count
Monocyte count
Neutrophil count
Monocyte percentage
Neutrophil percentage
Immature reticulocyte fraction
Platelet distribution width
Cereal intake
Feature
Immune cells
Reticulocytes
RBC
Platelets
Diet
Liver enzymes
Lipids
Urine and kidney
Other biomarkers



Appendix Figure 3

Bar charts for **A)** D2AF **B)** DW2 **C)** D2BF and **D)** control cohorts showing the number of associations for each feature studied; features with zero associations are not shown

Whole blood count features	
White blood cell (leukocyte) count	Eosinophil count
Red blood cell (erythrocyte) count	Basophil count
Haemoglobin concentration	Lymphocyte percentage
Haematocrit percentage	Monocyte percentage
Mean corpuscular volume	Neutrophil percentage
Mean corpuscular haemoglobin	Eosinophil percentage
Mean corpuscular haemoglobin concentration	Basophil percentage
Platelet count	Reticulocyte percentage
Platelet crit	Reticulocyte count
Red blood cell (erythrocyte) distribution width	Mean reticulocyte volume
Mean platelet (thrombocyte) volume	Mean sphered cell volume
Platelet distribution width	Immature reticulocyte fraction
Lymphocyte count	High light scatter reticulocyte percentage
Monocyte count	High light scatter reticulocyte count
Neutrophil count	
Dietary features	
Cooked vegetable intake	Lamb/mutton intake
Salad / raw vegetable intake	Cheese intake
Fresh fruit intake	Salt added to food
Bread intake	Cereal type
Cereal intake	Milk type used
Processed meat intake	Spread type
Non oily fish intake	Bread type
Never eat eggs, dairy, wheat, sugar	Coffee type
Biomarker features	
Alanine aminotransferase	Alkaline phosphatase
Phosphate	HDL cholesterol
Calcium	Gamma glutamyl transferase
Vitamin D	Glycated haemoglobin (HbA1c)
Microalbumin in urine	LDL direct
Creatinine (enzymatic) in urine	Aspartate aminotransferase
Potassium in urine	Triglycerides
Sodium in urine	Glucose

Biomarker features	
Direct bilirubin	Total protein
Cholesterol	SHBG
Testosterone	IGF 1
Creatinine	Apolipoprotein A
C reactive protein	Apolipoprotein B
Cystatin C	Lipoprotein A
Albumin	Urea
Urate	Total bilirubin
Clinical features	
First Diagnosis	genetic sex male
Diseases2	Body mass index
Time to Death months	Binary Death
Time to Disease	Current Age Date
Age at Disease Last	Age Now
main death	smoking status coding
date death	age recruitment

Appendix Table 1

Comprehensive list of all ninety-one features utilised in this association analysis study, as they are recorded in the UK Biobank directory. Features are grouped by data type.

Cohort	Class1	Marker1	Class2	Marker2	Coefficient	P value
Control	Biomarker	Sodium in urine	Diet	Cereal intake	-2.13486	7.42E-33
Control	Biomarker	Sodium in urine	Diet	Fresh fruit intake	-3.63513	6.70E-20
Control	Biomarker	Apo-A	Diet	Cereal intake	-0.00921	9.13E-20
Control	Biomarker	Urate	Diet	Cereal intake	-2.26712	2.10E-17
Control	Biomarker	Sodium in urine	Diet	Never eat cereal	0.00550	4.69E-13
Control	Biomarker	HDL-C	Diet	Cereal intake	-0.01024	9.90E-14
Control	Biomarker	IGF 1	Diet	Cereal intake	0.16719	2.41E-13
Control	Biomarker	Vitamin D	Diet	Cereal intake	0.64105	6.75E-13
Control	Biomarker	ALP	Diet	Ground coffee intake	-0.01183	5.56E-12
Control	Biomarker	ALP	Diet	Instant coffee intake	0.00980	8.17E-12
Control	Biomarker	Vitamin D	Diet	Never eat cereal	-0.01147	1.66E-10
Control	Biomarker	Apo-A	Diet	Never eat cereal	0.90366	3.34E-10
Control	Biomarker	Cystatin C	Diet	Lamb/mutton intake	-1.39591	1.12E-09
Control	Biomarker	Creatinine (enzymatic) in urine	Diet	Cereal intake	-162.46998	2.68E-09
Control	Biomarker	Apo-A	Diet	Ground coffee intake	0.81487	2.54E-08
Control	Biomarker	Sodium in urine	Diet	Oat cereal intake	-0.00534	3.68E-08
Control	Biomarker	Sodium in urine	Diet	White bread intake	0.00403	5.86E-08
Control	Biomarker	Cystatin C	Diet	Eat eggs	2.02599	8.33E-08
Control	Biomarker	Urate	Diet	Never eat cereal	0.00304	2.12E-07
Control	Biomarker	Cystatin C	Diet	White bread intake	1.46489	2.36E-07
Control	Biomarker	IGF 1	Diet	Never eat cereal	-0.03732	2.72E-07
Control	Biomarker	Sodium in urine	Diet	Wholegrain bread intake	-0.00357	2.80E-07
Control	Biomarker	IGF 1	Diet	Wholegrain bread intake	0.02972	3.29E-07
Control	Biomarker	Cystatin C	Diet	Wholegrain bread intake	-1.32196	4.24E-07
Control	Biomarker	HDL-C	Diet	Eat eggs	-0.66633	7.70E-07
Control	Biomarker	HDL-C	Diet	Ground coffee intake	0.56532	7.71E-07
Control	Biomarker	Creatinine	Diet	Instant coffee intake	0.01417	8.38E-07
Control	Biomarker	HDL-C	Diet	Never eat cereal	0.57062	1.05E-06
Control	Biomarker	Vitamin D	Diet	Salad / raw vegetable intake	0.92090	1.88E-07
Control	Biomarker	Creatinine	Diet	Processed meat intake	0.01192	1.74E-07
Control	Biomarker	Cystatin C	Diet	Non oily fish intake	-1.19249	2.14E-07
Control	Biomarker	TC	Diet	Ground coffee intake	0.18729	1.70E-06
Control	Biomarker	Cystatin C	Diet	Instant coffee intake	1.33325	2.85E-06
Control	Biomarker	Vitamin D	Diet	Olive oil spread intake	0.00692	6.48E-06
Control	Biomarker	HDL-C	Diet	Eat wheat	-0.61127	7.12E-06
Control	Biomarker	Cystatin C	Diet	Eat wheat	1.74678	7.26E-06
Control	Biomarker	Cystatin C	Diet	Ground coffee intake	-1.53648	7.33E-06
Control	Biomarker	Creatinine	Diet	Ground coffee intake	-0.01547	1.22E-05
Control	Biomarker	Vitamin D	Diet	Full cream milk intake	-0.01227	1.37E-05
Control	Biomarker	Vitamin D	Diet	Butter spread intake	-0.00695	1.38E-05
Control	Biomarker	Cystatin C	Diet	Eat eggs dairy wheat sugar	1.65993	1.57E-05
Control	Biomarker	Cystatin C	Diet	Eat dairy	1.65993	1.57E-05
Control	Biomarker	GGT	Diet	Cereal intake	-0.51911	2.74E-06
Control	Biomarker	HDL-C	Diet	Eat eggs dairy wheat sugar	-0.57990	1.91E-05
Control	Biomarker	HDL-C	Diet	Eat dairy	-0.57990	1.91E-05
Control	Biomarker	Apo-A	Diet	Eat eggs	-0.80675	2.26E-05
Control	Biomarker	Potassium in urine	Diet	Instant coffee intake	-0.00386	3.22E-05
Control	Biomarker	Potassium in urine	Diet	Ground coffee intake	0.00438	4.56E-05
Control	Biomarker	Cystatin C	Diet	Other cereal intake	1.49231	4.99E-05
Control	Biomarker	IGF 1	Diet	Muesli intake	0.03115	5.19E-05
Control	Biomarker	Sodium in urine	Diet	Muesli intake	-0.00391	7.65E-05
Control	Biomarker	Urea	Diet	Processed meat intake	0.09598	1.36E-05
Control	Biomarker	Creatinine (enzymatic) in urine	Diet	Butter spread intake	0.00002	1.50E-04
Control	Biomarker	IGF 1	Diet	Cooked vegetable intake	0.22802	3.12E-05
Control	Biomarker	HDL-C	Diet	Other cereal intake	-0.57018	2.08E-04
Control	Biomarker	Creatinine	Diet	Other cereal intake	0.01435	2.41E-04
Control	Biomarker	Vitamin D	Diet	Skimmed milk intake	0.00819	2.68E-04
Control	Biomarker	CRP	Diet	Cooked vegetable intake	-0.10812	4.80E-05
Control	Biomarker	Creatinine	Diet	Eat eggs	0.01628	2.88E-04
Control	Biomarker	GGT	Diet	Fresh fruit intake	-1.01459	6.23E-05
Control	Biomarker	Potassium in urine	Diet	Muesli intake	0.00410	5.04E-04

Control	Biomarker	IGF 1	Diet	White bread intake	-0.02636	5.59E-04
Control	Biomarker	Apo-A	Diet	Eat wheat	-0.69264	8.46E-04
Control	Biomarker	Creatinine	Diet	White bread intake	0.01150	9.95E-04
Control	Biomarker	Creatinine (enzymatic) in urine	Diet	Fresh fruit intake	-254.92634	1.86E-04
Control	Biomarker	Sodium in urine	Diet	Other cereal intake	0.00354	1.14E-03
Control	Biomarker	CRP	Diet	Instant coffee intake	0.05427	1.18E-03
Control	Biomarker	Creatinine	Diet	Cheese intake	-0.00956	2.22E-04
Control	Biomarker	Apo-A	Diet	Eat eggs dairy wheat sugar	-0.66186	1.44E-03
Control	Biomarker	Apo-A	Diet	Eat dairy	-0.66186	1.44E-03
Control	Biomarker	Vitamin D	Diet	Decaffeinated coffee intake	0.00859	1.62E-03
Control	Biomarker	Apo-A	Diet	Other cereal intake	-0.68433	1.71E-03
Control	Biomarker	Potassium in urine	Diet	Bread intake	-0.25642	3.07E-04
Control	Biomarker	Vitamin D	Diet	Bran cereal intake	0.00818	1.79E-03
Control	Biomarker	Creatinine (enzymatic) in urine	Diet	Cooked vegetable intake	-239.05406	3.40E-04
Control	Biomarker	Cystatin C	Diet	Salad / raw vegetable intake	-0.00404	4.33E-04
Control	Biomarker	LDL-C	Diet	Ground coffee intake	0.18546	2.64E-03
Control	Biomarker	Creatinine	Diet	Salad / raw vegetable intake	-0.39289	4.72E-04
Control	Biomarker	TC	Diet	Cereal intake	-0.01811	4.83E-04
Control	Biomarker	Creatinine	Diet	Eat eggs dairy wheat sugar	0.01392	3.38E-03
Control	Biomarker	Creatinine	Diet	Eat dairy	0.01392	3.38E-03
Control	Biomarker	Sodium in urine	Diet	Bread intake	0.29752	5.89E-04
Control	Biomarker	Cystatin C	Diet	Fresh fruit intake	-0.00455	1.10E-03
Control	Biomarker	CRP	Diet	Ground coffee intake	-0.06309	6.74E-03
Control	Biomarker	Sodium in urine	Diet	Full cream milk intake	0.00415	7.10E-03
Control	Biomarker	Urate	Diet	Cheese intake	-0.00159	1.26E-03
Control	Biomarker	GGT	Diet	Wholegrain bread intake	-0.00469	1.04E-02
Control	Biomarker	Cystatin C	Diet	Never use spread	-1.75966	1.07E-02
Control	Biomarker	HDL-C	Diet	Salad / raw vegetable intake	0.01017	2.34E-03
Control	Biomarker	Cystatin C	Diet	Skimmed milk intake	-1.25966	1.40E-02
Control	Biomarker	Creatinine (enzymatic) in urine	Diet	Wholegrain bread intake	-0.00002	1.46E-02
Control	Biomarker	Creatinine (enzymatic) in urine	Diet	Bread intake	-38.58395	2.58E-03
Control	Biomarker	Albumin	Diet	Ground coffee intake	0.05140	1.60E-02
Control	Biomarker	TC	Diet	Instant coffee intake	-0.11398	1.61E-02
Control	Biomarker	GGT	Diet	Never eat cereal	0.00523	1.65E-02
Control	Biomarker	Vitamin D	Diet	Eat sugar	0.01694	1.73E-02
Control	Biomarker	CRP	Diet	Fresh fruit intake	-0.09378	3.06E-03
Control	Biomarker	Creatinine (enzymatic) in urine	Diet	Never eat cereal	0.00002	2.31E-02
Control	Biomarker	Total protein	Diet	Cheese intake	-0.02326	3.74E-03
Control	Biomarker	CRP	Diet	Salad / raw vegetable intake	-0.07850	4.35E-03
Control	Biomarker	CRP	Diet	White bread intake	0.04208	3.03E-02
Control	Biomarker	HDL-C	Diet	Biscuit cereal intake	-0.47808	3.11E-02
Control	Biomarker	Urate	Diet	Decaffeinated coffee intake	-0.00258	3.12E-02
Control	Biomarker	Creatinine (enzymatic) in urine	Diet	Oat cereal intake	-0.00002	3.42E-02
Control	Biomarker	Total protein	Diet	Muesli intake	-0.03089	4.10E-02
Control	Biomarker	Cystatin C	Diet	Full cream milk intake	1.41812	4.31E-02
Control	Biomarker	CRP	Diet	Other cereal intake	0.04651	4.56E-02
Control	Biomarker	HDL-C	Diet	White bread intake	-0.36061	4.68E-02
Control	Biomarker	AST	Diet	Fresh fruit intake	0.23702	1.07E-02
Control	Biomarker	Apo-A	Diet	Cheese intake	0.37314	1.02E-02
Control	Biomarker	ALP	Diet	Lamb/mutton intake	-0.00439	1.22E-02
Control	Biomarker	Urate	Diet	Non oily fish intake	0.00142	1.50E-02
Control	Biomarker	ALP	Diet	Bread intake	0.12048	1.82E-02
Control	Biomarker	HDL-C	Diet	Lamb/mutton intake	0.28739	1.75E-02

Control	Biomarker	Vitamin D	Diet	Non oily fish intake	0.00430	1.98E-02
Control	Biomarker	Creatinine (enzymatic) in urine	Diet	Salad / raw vegetable intake	-169.85987	2.31E-02
Control	Biomarker	SHBG	Diet	Lamb/mutton intake	-0.00429	3.15E-02
Control	Biomarker	Apo-A	Diet	Lamb/mutton intake	0.36264	3.33E-02
Control	Biomarker	HDL-C	Diet	Cheese intake	0.25975	3.35E-02
Control	Biomarker	Urate	Diet	Processed meat intake	0.00129	3.85E-02
Control	Biomarker	Glucose	Diet	Cooked vegetable intake	0.02281	4.89E-02
Control	Cell count	Mean sphered cell volume	Diet	Never eat cereal	0.04357	6.81E-11
Control	Cell count	MCV	Diet	Never eat cereal	0.04768	2.57E-08
Control	Cell count	Mean corpuscular haemoglobin	Diet	Eat eggs	-0.12630	6.76E-06
Control	Cell count	Mean corpuscular haemoglobin	Diet	Never eat cereal	0.10458	9.76E-06
Control	Cell count	Mean corpuscular haemoglobin	Diet	Ground coffee intake	0.09951	3.41E-05
Control	Cell count	Mean reticulocyte volume	Diet	Never eat cereal	0.02260	3.65E-05
Control	Cell count	MCV	Diet	Eat eggs	-0.04640	3.87E-05
Control	Cell count	Mean platelet volume	Diet	Butter spread intake	-0.12220	1.34E-04
Control	Cell count	Mean corpuscular haemoglobin	Diet	Eat eggs dairy wheat sugar	-0.10674	2.02E-04
Control	Cell count	Mean corpuscular haemoglobin	Diet	Eat dairy	-0.10674	2.02E-04
Control	Cell count	MCV	Diet	Eat eggs dairy wheat sugar	-0.03967	6.47E-04
Control	Cell count	MCV	Diet	Eat dairy	-0.03967	6.47E-04
Control	Cell count	Mean sphered cell volume	Diet	Eat eggs dairy wheat sugar	-0.03191	7.37E-04
Control	Cell count	Mean sphered cell volume	Diet	Eat dairy	-0.03191	7.37E-04
Control	Cell count	Mean sphered cell volume	Diet	Eat eggs	-0.03263	1.56E-03
Control	Cell count	Mean corpuscular haemoglobin	Diet	Eat wheat	-0.10004	1.93E-03
Control	Cell count	HLSRP	Diet	Never eat cereal	0.74408	4.01E-03
Control	Cell count	Mean sphered cell volume	Diet	Eat wheat	-0.03030	4.44E-03
Control	Cell count	RBC count	Diet	Eat eggs	0.46615	5.02E-03
Control	Cell count	MCV	Diet	Eat wheat	-0.03550	1.31E-02
Control	Cell count	MCV	Diet	Other cereal intake	-0.03277	1.42E-02
Control	Cell count	MCV	Diet	Ground coffee intake	0.02995	1.71E-02
Control	Cell count	Immature reticulocyte fraction	Diet	Never eat cereal	2.20985	2.01E-02
Control	Cell count	Mean platelet volume	Diet	Olive oil spread intake	0.09340	3.14E-02
Control	Cell count	RBC count	Diet	Eat wheat	0.41232	3.19E-02
Control	Cell count	HLSRC	Diet	Never eat cereal	14.23224	4.16E-02
Control	Cell count	WBC count	Diet	White bread intake	0.06226	4.39E-02
Control	Cell count	RBC count	Diet	Eat eggs dairy wheat sugar	0.39423	4.42E-02
Control	Cell count	RBC count	Diet	Eat dairy	0.39423	4.42E-02
Control	Cell count	RBC count	Diet	Ground coffee intake	-0.34873	4.64E-02
Control	Cell count	MCHC	Diet	Ground coffee intake	0.12908	4.80E-02
Control	Cell count	Mean sphered cell volume	Diet	Cereal intake	-0.17428	7.56E-15
Control	Cell count	Immature reticulocyte fraction	Diet	Cereal intake	-0.00157	1.51E-09
Control	Cell count	HLSRP	Diet	Cereal intake	-0.00484	2.58E-09
Control	Cell count	MCV	Diet	Cereal intake	-0.11270	7.42E-09
Control	Cell count	HLSRC	Diet	Cereal intake	-0.00021	4.96E-08
Control	Cell count	Mean reticulocyte volume	Diet	Cereal intake	-0.18396	1.12E-07
Control	Cell count	Mean corpuscular haemoglobin	Diet	Cereal intake	-0.04080	1.87E-07
Control	Cell count	WBC count	Diet	Fresh fruit intake	-0.08807	1.02E-06
Control	Cell count	Neutrophil count	Diet	Fresh fruit intake	-0.06864	1.73E-06
Control	Cell count	WBC count	Diet	Cereal intake	-0.03744	1.31E-05
Control	Cell count	Reticulocyte percentage	Diet	Cereal intake	-0.00928	1.35E-04
Control	Cell count	Platelet count	Diet	Cereal intake	-1.08333	2.82E-04
Control	Cell count	Neutrophil count	Diet	Cereal intake	-0.02628	3.11E-04
Control	Cell count	Reticulocyte count	Diet	Cereal intake	-0.00039	1.69E-03
Control	Cell count	Lymphocyte count	Diet	Cereal intake	-0.01053	1.76E-03
Control	Cell count	Platelet crit	Diet	Cereal intake	-0.00078	2.15E-03

Control	Cell count	Lymphocyte percentage	Diet	Salad / raw vegetable intake	0.20819	1.65E-02
Control	Cell count	Monocyte percentage	Diet	Cereal intake	0.03122	2.30E-02
Control	Cell count	Mean reticulocyte volume	Diet	Salad / raw vegetable intake	0.20850	2.46E-02
Control	Cell count	Monocyte count	Diet	Fresh fruit intake	-0.00517	2.98E-02
Control	Cell count	MCV	Diet	Fresh fruit intake	-0.13821	3.07E-02
Control	Cell count	Neutrophil percentage	Diet	Processed meat intake	0.01457	5.96E-06
Control	Cell count	Neutrophil count	Diet	Processed meat intake	0.08054	6.35E-05
Control	Cell count	Lymphocyte percentage	Diet	Processed meat intake	-0.01472	1.82E-04
Control	Cell count	Mean platelet volume	Diet	Cheese intake	-0.09011	2.40E-03
Control	Cell count	Lymphocyte count	Diet	Salt added to food	0.16522	3.20E-03
Control	Cell count	Reticulocyte count	Diet	Lamb/mutton intake	4.50489	3.39E-03
Control	Cell count	HLSRC	Diet	Lamb/mutton intake	11.64008	5.10E-03
Control	Cell count	Reticulocyte percentage	Diet	Lamb/mutton intake	0.20597	5.19E-03
Control	Cell count	WBC count	Diet	Salt added to food	0.05468	6.31E-03
Control	Cell count	HLSRP	Diet	Lamb/mutton intake	0.53052	7.12E-03
Control	Cell count	WBC count	Diet	Processed meat intake	0.04766	2.21E-02
Control	Cell count	Mean sphered cell volume	Diet	Salt added to food	0.01700	2.29E-02
Control	Cell count	Lymphocyte percentage	Diet	Lamb/mutton intake	0.01197	3.31E-02
Control	Biomarker	TG	Cell count	Reticulocyte count	0.00538	6.57E-69
Control	Biomarker	HDL-C	Cell count	Mean sphered cell volume	3.29475	2.58E-64
Control	Biomarker	CRP	Cell count	Neutrophil count	0.10943	1.42E-63
Control	Biomarker	Albumin	Cell count	Haemoglobin concentration	0.07848	5.29E-63
Control	Biomarker	TG	Cell count	Reticulocyte percentage	0.10677	5.00E-59
Control	Biomarker	TG	Cell count	HLSRC	0.00187	1.18E-57
Control	Biomarker	Phosphate	Cell count	Lymphocyte count	0.68847	1.73E-56
Control	Biomarker	Apo-A	Cell count	Mean sphered cell volume	4.04100	6.04E-55
Control	Biomarker	Albumin	Cell count	Haematocrit percentage	0.21402	2.64E-54
Control	Biomarker	HbA1c	Cell count	Mean corpuscular haemoglobin	-0.07515	7.89E-54
Control	Biomarker	CRP	Cell count	WBC count	0.12567	3.67E-52
Control	Biomarker	ALT	Cell count	Haemoglobin concentration	0.01685	4.16E-51
Control	Biomarker	Calcium	Cell count	Haematocrit percentage	5.67364	2.27E-50
Control	Biomarker	Calcium	Cell count	Haemoglobin concentration	1.91315	4.28E-49
Control	Biomarker	TG	Cell count	HLSRP	0.03704	1.74E-48
Control	Biomarker	Calcium	Cell count	RBC count	0.63390	1.05E-42
Control	Biomarker	Total bilirubin	Cell count	Haemoglobin concentration	0.04318	4.66E-42
Control	Biomarker	Apo-B	Cell count	Haemoglobin concentration	0.71807	1.50E-38
Control	Biomarker	TG	Cell count	Platelet distribution width	0.09500	1.66E-38
Control	Biomarker	Urate	Cell count	HLSRC	0.00002	2.97E-37
Control	Biomarker	CRP	Cell count	Lymphocyte percentage	-0.44188	8.18E-37
Control	Biomarker	ALT	Cell count	HLSRC	0.00012	5.89E-36
Control	Biomarker	Albumin	Cell count	RBC count	0.02097	5.07E-35
Control	Biomarker	ALT	Cell count	Haematocrit percentage	0.04114	6.05E-35
Control	Biomarker	ALT	Cell count	Reticulocyte count	0.00031	1.88E-34
Control	Biomarker	Urate	Cell count	HLSRP	0.00043	3.17E-34
Control	Biomarker	Urate	Cell count	Reticulocyte count	0.00005	2.59E-33
Control	Biomarker	HbA1c	Cell count	Lymphocyte count	0.02118	1.03E-32
Control	Biomarker	Cystatin C	Cell count	RBC count	0.42698	1.13E-31
Control	Biomarker	Total bilirubin	Cell count	Haematocrit percentage	0.10944	4.58E-31
Control	Biomarker	GGT	Cell count	HLSRC	0.00005	1.31E-30
Control	Biomarker	HDL-C	Cell count	Mean reticulocyte volume	3.31169	1.60E-30
Control	Biomarker	Apo-A	Cell count	MCV	2.51199	1.81E-30
Control	Biomarker	Total protein	Cell count	RBC count	0.01197	3.01E-30
Control	Biomarker	TG	Cell count	Lymphocyte count	0.09853	3.52E-30
Control	Biomarker	HbA1c	Cell count	RDW	0.02582	4.78E-30
Control	Biomarker	HDL-C	Cell count	RBC count	-0.15285	5.08E-30

Control	Biomarker	HbA1c	Cell count	MCHC	-0.03141	1.01E-29
Control	Biomarker	Apo-B	Cell count	Haematocrit percentage	1.85657	1.13E-29
Control	Biomarker	Urate	Cell count	Reticulocyte percentage	0.00104	1.81E-29
Control	Biomarker	TG	Cell count	Mean reticulocyte volume	-1.19558	1.15E-28
Control	Biomarker	GGT	Cell count	HLSRP	0.00107	1.20E-28
Control	Biomarker	HbA1c	Cell count	MCV	-0.14069	3.88E-28
Control	Biomarker	Total protein	Cell count	Platelet crit	0.00149	4.30E-28
Control	Biomarker	CRP	Cell count	Neutrophil percentage	0.44033	1.06E-27
Control	Biomarker	Apo-B	Cell count	RBC count	0.21709	1.23E-27
Control	Biomarker	HDL-C	Cell count	MCV	1.81036	1.47E-27
Control	Biomarker	LDL-C	Cell count	Haemoglobin concentration	0.17183	4.10E-27
Control	Biomarker	ALP	Cell count	RBC count	0.00217	1.11E-26
Control	Biomarker	SHBG	Cell count	Reticulocyte count	-0.00014	1.26E-26
Control	Biomarker	ALT	Cell count	HLSRP	0.00224	4.00E-26
Control	Biomarker	Cystatin C	Cell count	Haematocrit percentage	3.20616	7.46E-26
Control	Biomarker	Testosterone	Cell count	Haematocrit percentage	0.14086	2.83E-25
Control	Biomarker	TG	Cell count	WBC count	0.26268	4.46E-25
Control	Biomarker	Total protein	Cell count	Haematocrit percentage	0.09049	4.75E-25
Control	Biomarker	SHBG	Cell count	HLSRC	-0.00005	6.94E-25
Control	Biomarker	ALT	Cell count	RBC count	0.00422	1.16E-24
Control	Biomarker	TG	Cell count	Mean sphered cell volume	-0.77338	1.47E-24
Control	Biomarker	Apo-A	Cell count	Mean corpuscular haemoglobin	0.88739	1.68E-24
Control	Biomarker	Cystatin C	Cell count	WBC count	1.90038	1.88E-24
Control	Biomarker	ALP	Cell count	WBC count	0.01039	5.98E-24
Control	Biomarker	Apo-A	Cell count	Mean reticulocyte volume	3.87602	1.50E-23
Control	Biomarker	GGT	Cell count	Reticulocyte count	0.00012	1.95E-23
Control	Biomarker	Total bilirubin	Cell count	Platelet count	-1.89008	2.49E-23
Control	Biomarker	Total protein	Cell count	Platelet count	1.72775	2.98E-23
Control	Biomarker	SHBG	Cell count	Reticulocyte percentage	-0.00286	6.24E-23
Control	Biomarker	CRP	Cell count	Monocyte count	0.00777	6.80E-23
Control	Biomarker	ALT	Cell count	Reticulocyte percentage	0.00552	8.06E-23
Control	Biomarker	Total bilirubin	Cell count	Platelet crit	-0.00147	1.50E-22
Control	Biomarker	Cystatin C	Cell count	Haemoglobin concentration	1.01885	3.36E-22
Control	Biomarker	Phosphate	Cell count	Lymphocyte percentage	5.32820	3.86E-22
Control	Biomarker	GGT	Cell count	Reticulocyte percentage	0.00247	3.94E-22
Control	Biomarker	LDL-C	Cell count	RBC count	0.05472	1.28E-21
Control	Biomarker	Testosterone	Cell count	Haemoglobin concentration	0.04465	1.40E-21
Control	Biomarker	Phosphate	Cell count	Neutrophil percentage	-5.95783	3.33E-21
Control	Biomarker	SHBG	Cell count	HLSRP	-0.00105	4.76E-21
Control	Biomarker	ALP	Cell count	Platelet crit	0.00025	5.18E-21
Control	Biomarker	Total bilirubin	Cell count	Lymphocyte count	-0.01892	9.19E-21
Control	Biomarker	Total protein	Cell count	Lymphocyte count	0.01732	9.38E-21
Control	Biomarker	Calcium	Cell count	Platelet crit	0.05818	9.60E-21
Control	Biomarker	TC	Cell count	Haemoglobin concentration	0.11689	3.66E-20
Control	Biomarker	LDL-C	Cell count	Haematocrit percentage	0.43664	4.94E-20
Control	Biomarker	ALP	Cell count	Neutrophil count	0.00752	9.54E-20
Control	Biomarker	HDL-C	Cell count	WBC count	-0.62295	3.53E-19
Control	Biomarker	Total bilirubin	Cell count	RBC count	0.01040	1.40E-18
Control	Biomarker	HDL-C	Cell count	Mean corpuscular haemoglobin	0.58760	1.86E-18
Control	Biomarker	Total protein	Cell count	Haemoglobin concentration	0.02646	5.61E-18
Control	Biomarker	GGT	Cell count	Haemoglobin concentration	0.00464	1.00E-17
Control	Biomarker	ALP	Cell count	Haematocrit percentage	0.01477	1.55E-17
Control	Biomarker	Cystatin C	Cell count	Eosinophil count	0.10917	1.56E-17
Control	Biomarker	Calcium	Cell count	Lymphocyte count	0.71741	2.05E-17
Control	Biomarker	HbA1c	Cell count	WBC count	0.04564	3.05E-17

Control	Biomarker	ALP	Cell count	Haemoglobin concentration	0.00498	4.80E-17
Control	Biomarker	HDL-C	Cell count	Reticulocyte count	-0.00740	8.74E-17
Control	Biomarker	HbA1c	Cell count	Haemoglobin concentration	-0.02492	2.95E-16
Control	Biomarker	Urate	Cell count	Haemoglobin concentration	0.00163	4.38E-16
Control	Biomarker	HbA1c	Cell count	Reticulocyte percentage	-0.01190	5.47E-16
Control	Biomarker	TG	Cell count	Haemoglobin concentration	0.11918	6.96E-16
Control	Biomarker	Total bilirubin	Cell count	WBC count	-0.04860	7.66E-16
Control	Biomarker	Total protein	Cell count	WBC count	0.04437	9.68E-16
Control	Biomarker	Total bilirubin	Cell count	Monocyte count	-0.00440	1.94E-15
Control	Biomarker	GGT	Cell count	Mean corpuscular haemoglobin	0.00740	3.44E-15
Control	Biomarker	ALP	Cell count	Platelet count	0.27308	3.89E-15
Control	Biomarker	TG	Cell count	MCHC	0.11075	4.03E-15
Control	Biomarker	Phosphate	Cell count	WBC count	1.07217	4.86E-15
Control	Biomarker	IGF 1	Cell count	Haemoglobin concentration	0.01869	1.05E-14
Control	Biomarker	Albumin	Cell count	RDW	-0.03054	1.40E-14
Control	Biomarker	Cystatin C	Cell count	Monocyte count	0.13684	1.46E-14
Control	Biomarker	HbA1c	Cell count	Platelet crit	0.00108	1.98E-14
Control	Biomarker	Cystatin C	Cell count	Neutrophil count	1.17309	2.77E-14
Control	Biomarker	HDL-C	Cell count	Platelet distribution width	-0.15722	1.20E-13
Control	Biomarker	Apo-A	Cell count	RBC count	-0.13744	4.63E-13
Control	Biomarker	TC	Cell count	Haematocrit percentage	0.27996	5.46E-13
Control	Biomarker	Albumin	Cell count	Eosinophil percentage	-0.06291	5.67E-13
Control	Biomarker	IGF 1	Cell count	RBC count	0.00620	1.02E-12
Control	Biomarker	Calcium	Cell count	Platelet count	58.76990	1.07E-12
Control	Biomarker	Urea	Cell count	Haemoglobin concentration	-0.07413	1.37E-12
Control	Biomarker	IGF 1	Cell count	Haematocrit percentage	0.05075	2.01E-12
Control	Biomarker	Cystatin C	Cell count	Lymphocyte count	0.47813	2.03E-12
Control	Biomarker	Phosphate	Cell count	Monocyte count	0.08967	2.62E-12
Control	Biomarker	AST	Cell count	Haemoglobin concentration	0.01304	6.06E-12
Control	Biomarker	HDL-C	Cell count	Neutrophil count	-0.39754	6.09E-12
Control	Biomarker	CRP	Cell count	Platelet crit	0.00162	6.14E-12
Control	Biomarker	GGT	Cell count	Immature reticulocyte fraction	0.00023	8.79E-12
Control	Biomarker	LDL-C	Cell count	Platelet count	6.74783	1.18E-11
Control	Biomarker	TG	Cell count	RBC count	0.03662	1.41E-11
Control	Biomarker	Albumin	Cell count	Eosinophil count	-0.00423	1.59E-11
Control	Biomarker	HbA1c	Cell count	Reticulocyte count	-0.00048	1.81E-11
Control	Biomarker	HDL-C	Cell count	HLSRC	-0.00236	1.82E-11
Control	Biomarker	Urate	Cell count	Immature reticulocyte fraction	0.00008	2.02E-11
Control	Biomarker	Urea	Cell count	Haematocrit percentage	-0.20617	3.33E-11
Control	Biomarker	Calcium	Cell count	WBC count	1.70695	3.80E-11
Control	Biomarker	CRP	Cell count	Platelet count	1.97218	5.62E-11
Control	Biomarker	Phosphate	Cell count	Eosinophil count	0.06235	6.24E-11
Control	Biomarker	TG	Cell count	Neutrophil count	0.14149	8.28E-11
Control	Biomarker	HDL-C	Cell count	Lymphocyte count	-0.16462	9.83E-11
Control	Biomarker	Apo-B	Cell count	Platelet count	22.87279	1.33E-10
Control	Biomarker	Testosterone	Cell count	RBC count	0.01145	1.34E-10
Control	Biomarker	TC	Cell count	Platelet count	5.05183	1.45E-10
Control	Biomarker	HbA1c	Cell count	Eosinophil count	0.00245	1.61E-10
Control	Biomarker	Urate	Cell count	Haematocrit percentage	0.00389	3.59E-10
Control	Biomarker	Phosphate	Cell count	Platelet crit	0.02296	5.42E-10
Control	Biomarker	SHBG	Cell count	RBC count	-0.00142	1.42E-09
Control	Biomarker	Albumin	Cell count	Reticulocyte count	0.00073	1.59E-09
Control	Biomarker	TG	Cell count	Immature reticulocyte fraction	0.00553	3.00E-09
Control	Biomarker	GGT	Cell count	Haematocrit percentage	0.00997	7.48E-09
Control	Biomarker	Glucose	Cell count	Neutrophil percentage	0.86957	1.02E-08

Control	Biomarker	HDL-C	Cell count	Reticulocyte percentage	-0.11797	1.13E-08
Control	Biomarker	Calcium	Cell count	Reticulocyte count	0.01932	1.17E-08
Control	Biomarker	CRP	Cell count	Immature reticulocyte fraction	0.00181	1.39E-08
Control	Biomarker	HbA1c	Cell count	Platelet count	1.08602	1.50E-08
Control	Biomarker	AST	Cell count	HLSRC	0.00010	1.59E-08
Control	Biomarker	HDL-C	Cell count	Eosinophil count	-0.02863	3.32E-08
Control	Biomarker	TC	Cell count	RBC count	0.02760	4.04E-08
Control	Biomarker	Apo-A	Cell count	WBC count	-0.56181	4.17E-08
Control	Biomarker	Urea	Cell count	Mean reticulocyte volume	-0.46505	4.66E-08
Control	Biomarker	ALT	Cell count	Immature reticulocyte fraction	0.00042	6.08E-08
Control	Biomarker	Apo-B	Cell count	Reticulocyte count	0.00782	7.24E-08
Control	Biomarker	SHBG	Cell count	Mean sphered cell volume	0.01930	1.14E-07
Control	Biomarker	SHBG	Cell count	Platelet count	-0.20995	1.39E-07
Control	Biomarker	Phosphate	Cell count	Platelet count	25.71181	1.68E-07
Control	Biomarker	ALT	Cell count	Lymphocyte percentage	0.05172	1.69E-07
Control	Biomarker	ALT	Cell count	MCHC	0.00662	1.82E-07
Control	Biomarker	AST	Cell count	Neutrophil percentage	-0.08976	2.08E-07
Control	Biomarker	HbA1c	Cell count	Monocyte count	0.00288	2.08E-07
Control	Biomarker	Apo-A	Cell count	Platelet distribution width	-0.15886	2.51E-07
Control	Biomarker	Total bilirubin	Cell count	Reticulocyte count	0.00044	2.59E-07
Control	Biomarker	Total bilirubin	Cell count	Eosinophil count	-0.00232	2.68E-07
Control	Biomarker	Creatinine	Cell count	Mean corpuscular haemoglobin	-0.01082	3.57E-07
Control	Biomarker	ALT	Cell count	RDW	-0.00531	4.00E-07
Control	Biomarker	GGT	Cell count	MCV	0.01357	4.12E-07
Control	Biomarker	SHBG	Cell count	Haemoglobin concentration	-0.00349	6.63E-07
Control	Biomarker	AST	Cell count	Haematocrit percentage	0.03012	7.56E-07
Control	Biomarker	SHBG	Cell count	Immature reticulocyte fraction	-0.00020	8.07E-07
Control	Biomarker	Testosterone	Cell count	WBC count	-0.04758	9.64E-07
Control	Biomarker	SHBG	Cell count	Haematocrit percentage	-0.01010	1.01E-06
Control	Biomarker	TG	Cell count	RDW	-0.06363	1.09E-06
Control	Biomarker	CRP	Cell count	HLSRC	0.00023	1.16E-06
Control	Biomarker	CRP	Cell count	HLSRP	0.00505	1.18E-06
Control	Biomarker	Total protein	Cell count	Neutrophil count	0.02398	1.58E-06
Control	Biomarker	Urate	Cell count	WBC count	0.00196	1.60E-06
Control	Biomarker	AST	Cell count	HLSRP	0.00185	1.64E-06
Control	Biomarker	Apo-A	Cell count	Eosinophil count	-0.03433	1.83E-06
Control	Biomarker	AST	Cell count	Lymphocyte percentage	0.07441	1.83E-06
Control	Biomarker	Glucose	Cell count	Monocyte percentage	-0.19014	2.21E-06
Control	Biomarker	Albumin	Cell count	Monocyte percentage	-0.05769	2.49E-06
Control	Biomarker	Testosterone	Cell count	Mean reticulocyte volume	0.19816	2.50E-06
Control	Biomarker	GGT	Cell count	MCHC	0.00281	2.60E-06
Control	Biomarker	HbA1c	Cell count	Lymphocyte percentage	0.12231	2.79E-06
Control	Biomarker	AST	Cell count	Reticulocyte count	0.00022	3.12E-06
Control	Biomarker	Testosterone	Cell count	Neutrophil count	-0.03621	4.11E-06
Control	Biomarker	HDL-C	Cell count	HLSRP	-0.03870	5.12E-06
Control	Biomarker	ALT	Cell count	Neutrophil percentage	-0.05378	5.80E-06
Control	Biomarker	Apo-A	Cell count	Lymphocyte count	-0.16774	9.73E-06
Control	Biomarker	TG	Cell count	Platelet crit	0.00352	9.88E-06
Control	Biomarker	IGF 1	Cell count	Neutrophil percentage	0.10565	9.93E-06
Control	Biomarker	Cystatin C	Cell count	Eosinophil percentage	0.94753	1.12E-05
Control	Biomarker	Testosterone	Cell count	Mean sphered cell volume	0.13075	1.57E-05
Control	Biomarker	IGF 1	Cell count	Lymphocyte percentage	-0.09114	1.68E-05
Control	Biomarker	TC	Cell count	Mean platelet volume	-0.07293	1.71E-05
Control	Biomarker	Urate	Cell count	RBC count	0.00036	1.80E-05
Control	Biomarker	Total protein	Cell count	Monocyte count	0.00258	1.99E-05
Control	Biomarker	HbA1c	Cell count	Neutrophil percentage	-0.13114	2.32E-05

Control	Biomarker	Urea	Cell count	Mean sphered cell volume	-0.27439	2.37E-05
Control	Biomarker	Apo-B	Cell count	Platelet crit	0.01360	2.52E-05
Control	Biomarker	Glucose	Cell count	Monocyte count	-0.01369	2.82E-05
Control	Biomarker	Phosphate	Cell count	Haemoglobin concentration	-0.38560	2.83E-05
Control	Biomarker	LDL-C	Cell count	Platelet crit	0.00383	3.04E-05
Control	Biomarker	Total bilirubin	Cell count	RDW	-0.01324	3.06E-05
Control	Biomarker	TG	Cell count	Haematocrit percentage	0.21728	3.27E-05
Control	Biomarker	Total bilirubin	Cell count	MCHC	0.01609	3.67E-05
Control	Biomarker	Apo-B	Cell count	Lymphocyte count	0.17879	4.34E-05
Control	Biomarker	TG	Cell count	Eosinophil count	0.00877	4.70E-05
Control	Biomarker	LDL-C	Cell count	Mean platelet volume	-0.09032	4.74E-05
Control	Biomarker	Creatinine	Cell count	Platelet count	-0.31725	5.47E-05
Control	Biomarker	GGT	Cell count	Lymphocyte count	0.00165	6.29E-05
Control	Biomarker	Creatinine	Cell count	Mean sphered cell volume	-0.02887	6.41E-05
Control	Biomarker	ALP	Cell count	Lymphocyte count	0.00180	6.45E-05
Control	Biomarker	TC	Cell count	Lymphocyte percentage	0.47153	6.72E-05
Control	Biomarker	Vitamin D	Cell count	WBC count	-0.00538	7.04E-05
Control	Biomarker	HDL-C	Cell count	Haemoglobin concentration	-0.19370	7.94E-05
Control	Biomarker	ALP	Cell count	Monocyte count	0.00047	8.62E-05
Control	Biomarker	Creatinine	Cell count	RBC count	0.00191	8.88E-05
Control	Biomarker	Total bilirubin	Cell count	Mean corpuscular haemoglobin	0.02812	9.03E-05
Control	Biomarker	IGF 1	Cell count	Lymphocyte count	-0.00714	9.09E-05
Control	Biomarker	AST	Cell count	Mean corpuscular haemoglobin	0.01516	1.13E-04
Control	Biomarker	Microalbumin in urine	Cell count	Neutrophil count	0.00329	1.32E-04
Control	Biomarker	TG	Cell count	Monocyte count	0.01153	1.41E-04
Control	Biomarker	AST	Cell count	Reticulocyte percentage	0.00420	1.67E-04
Control	Biomarker	Glucose	Cell count	Reticulocyte percentage	0.03796	1.78E-04
Control	Biomarker	Cystatin C	Cell count	Mean sphered cell volume	-2.70092	2.12E-04
Control	Biomarker	Calcium	Cell count	Neutrophil count	0.92006	2.50E-04
Control	Biomarker	Urate	Cell count	Neutrophil count	0.00133	2.65E-04
Control	Biomarker	ALT	Cell count	Lymphocyte count	0.00344	2.69E-04
Control	Biomarker	TC	Cell count	Platelet crit	0.00277	2.76E-04
Control	Biomarker	CRP	Cell count	RDW	0.01803	2.98E-04
Control	Biomarker	LDL-C	Cell count	Lymphocyte percentage	0.57048	3.26E-04
Control	Biomarker	HDL-C	Cell count	Haematocrit percentage	-0.53795	3.55E-04
Control	Biomarker	Total bilirubin	Cell count	Neutrophil count	-0.02205	3.67E-04
Control	Biomarker	AST	Cell count	Monocyte percentage	0.01782	3.85E-04
Control	Biomarker	Phosphate	Cell count	Basophil count	0.01292	4.66E-04
Control	Biomarker	Albumin	Cell count	Basophil percentage	-0.00904	4.86E-04
Control	Biomarker	Glucose	Cell count	Reticulocyte count	0.00171	4.90E-04
Control	Biomarker	TG	Cell count	MCV	-0.29482	4.92E-04
Control	Biomarker	Phosphate	Cell count	MCHC	-0.33053	5.45E-04
Control	Biomarker	SHBG	Cell count	Lymphocyte count	-0.00172	6.15E-04
Control	Biomarker	Microalbumin in urine	Cell count	WBC count	0.00390	6.96E-04
Control	Biomarker	SHBG	Cell count	Platelet crit	-0.00013	8.35E-04
Control	Biomarker	Apo-B	Cell count	Lymphocyte percentage	1.94014	8.82E-04
Control	Biomarker	Glucose	Cell count	Lymphocyte percentage	-0.54758	9.86E-04
Control	Biomarker	Albumin	Cell count	HLSRC	0.00019	1.17E-03
Control	Biomarker	CRP	Cell count	Eosinophil count	0.00261	1.39E-03
Control	Biomarker	Calcium	Cell count	HLSRC	0.00518	1.49E-03
Control	Biomarker	Apo-B	Cell count	Mean platelet volume	-0.28073	1.54E-03
Control	Biomarker	Creatinine	Cell count	MCV	-0.02109	1.64E-03
Control	Biomarker	Apo-B	Cell count	HLSRC	0.00217	1.68E-03
Control	Biomarker	TG	Cell count	Platelet count	3.71396	1.74E-03
Control	Biomarker	AST	Cell count	Neutrophil count	-0.01125	1.76E-03
Control	Biomarker	Cystatin C	Cell count	Reticulocyte count	0.01050	2.03E-03
Control	Biomarker	Creatinine	Cell count	Platelet crit	-0.00022	2.30E-03
Control	Biomarker	Glucose	Cell count	Lymphocyte count	-0.04364	2.37E-03
Control	Biomarker	Vitamin D	Cell count	Neutrophil count	-0.00372	2.38E-03

Control	Biomarker	Sodium in urine	Cell count	Haemoglobin concentration	-0.00124	2.43E-03
Control	Biomarker	Phosphate	Cell count	Mean corpuscular haemoglobin	-0.55789	2.63E-03
Control	Biomarker	HbA1c	Cell count	Neutrophil count	0.01822	3.23E-03
Control	Biomarker	Urate	Cell count	Lymphocyte count	0.00052	3.61E-03
Control	Biomarker	ALP	Cell count	MCV	-0.01089	4.28E-03
Control	Biomarker	Testosterone	Cell count	Monocyte count	-0.00327	4.37E-03
Control	Biomarker	Urate	Cell count	MCHC	0.00080	4.39E-03
Control	Biomarker	TC	Cell count	MCHC	0.05138	4.44E-03
Control	Biomarker	Apo-A	Cell count	Neutrophil count	-0.30815	4.67E-03
Control	Biomarker	Albumin	Cell count	Lymphocyte count	0.01275	5.93E-03
Control	Biomarker	Sodium in urine	Cell count	Haematocrit percentage	-0.00348	6.35E-03
Control	Biomarker	Vitamin D	Cell count	Lymphocyte count	-0.00154	6.90E-03
Control	Biomarker	GGT	Cell count	Mean sphered cell volume	0.01182	6.91E-03
Control	Biomarker	SHBG	Cell count	WBC count	-0.00449	7.80E-03
Control	Biomarker	ALP	Cell count	Mean corpuscular haemoglobin	-0.00412	8.74E-03
Control	Biomarker	Calcium	Cell count	RDW	-0.43547	9.18E-03
Control	Biomarker	Albumin	Cell count	Reticulocyte percentage	0.00973	1.08E-02
Control	Biomarker	HbA1c	Cell count	HLSRP	-0.00224	1.09E-02
Control	Biomarker	Cystatin C	Cell count	HLSRC	0.00367	1.12E-02
Control	Biomarker	Albumin	Cell count	Monocyte count	-0.00329	1.15E-02
Control	Biomarker	Urea	Cell count	RBC count	-0.01437	1.18E-02
Control	Biomarker	Apo-B	Cell count	Reticulocyte percentage	0.11177	1.28E-02
Control	Biomarker	HbA1c	Cell count	Haematocrit percentage	-0.03523	1.44E-02
Control	Biomarker	Urate	Cell count	Platelet distribution width	0.00041	1.73E-02
Control	Biomarker	Potassium in urine	Cell count	MCV	0.00589	1.77E-02
Control	Biomarker	Total bilirubin	Cell count	Reticulocyte percentage	0.00634	1.78E-02
Control	Biomarker	SHBG	Cell count	Mean reticulocyte volume	0.01841	1.80E-02
Control	Biomarker	Phosphate	Cell count	Eosinophil percentage	0.50153	1.82E-02
Control	Biomarker	GGT	Cell count	RDW	-0.00165	1.86E-02
Control	Biomarker	SHBG	Cell count	Mean platelet volume	0.00264	1.98E-02
Control	Biomarker	Sodium in urine	Cell count	Immature reticulocyte fraction	0.00007	1.99E-02
Control	Biomarker	Potassium in urine	Cell count	WBC count	-0.00238	2.17E-02
Control	Biomarker	Cystatin C	Cell count	Mean corpuscular haemoglobin	-0.71164	2.24E-02
Control	Biomarker	IGF 1	Cell count	Immature reticulocyte fraction	-0.00054	2.25E-02
Control	Biomarker	Potassium in urine	Cell count	Mean corpuscular haemoglobin	0.00228	2.29E-02
Control	Biomarker	Glucose	Cell count	Eosinophil percentage	-0.10761	2.37E-02
Control	Biomarker	IGF 1	Cell count	Eosinophil count	-0.00110	2.53E-02
Control	Biomarker	HbA1c	Cell count	Eosinophil percentage	0.01970	2.55E-02
Control	Biomarker	Testosterone	Cell count	HLSRP	-0.00334	2.95E-02
Control	Biomarker	AST	Cell count	MCHC	0.00662	2.95E-02
Control	Biomarker	Glucose	Cell count	HLSRC	0.00054	3.09E-02
Control	Biomarker	Cystatin C	Cell count	RDW	0.31707	3.20E-02
Control	Biomarker	LDL-C	Cell count	Mean reticulocyte volume	-0.47514	3.40E-02
Control	Biomarker	Total protein	Cell count	Mean corpuscular haemoglobin	-0.01991	3.45E-02
Control	Biomarker	HbA1c	Cell count	Immature reticulocyte fraction	0.00067	3.81E-02
Control	Biomarker	Potassium in urine	Cell count	Neutrophil count	-0.00182	3.87E-02
Control	Biomarker	Testosterone	Cell count	Reticulocyte percentage	-0.00857	4.28E-02
Control	Biomarker	ALT	Cell count	Mean corpuscular haemoglobin	0.00760	4.63E-02
Control	Biomarker	Albumin	Cell count	Lymphocyte percentage	0.13800	4.95E-02
D2AF	Biomarker	Sodium in urine	Diet	Fresh fruit intake	-4.18853	7.97E-18
D2AF	Biomarker	Sodium in urine	Diet	Cereal intake	-2.13206	1.29E-17
D2AF	Biomarker	Sodium in urine	Diet	Salt added to food	0.00555	2.36E-14
D2AF	Biomarker	Apo-A	Diet	Cereal intake	-0.01080	7.59E-14

D2AF	Biomarker	Urate	Diet	Cereal intake	-2.77218	3.53E-10
D2AF	Biomarker	Cystatin C	Diet	White bread intake	1.51668	1.10E-08
D2AF	Biomarker	Cystatin C	Diet	Wholegrain bread intake	-1.49993	1.69E-08
D2AF	Biomarker	HDL-C	Diet	Cereal intake	-0.01149	8.62E-09
D2AF	Biomarker	Sodium in urine	Diet	Bread intake	0.54637	1.08E-08
D2AF	Biomarker	Apo-A	Diet	Never eat cereal	0.96949	1.18E-07
D2AF	Biomarker	IGF 1	Diet	Salt added to food	-0.03132	1.54E-07
D2AF	Biomarker	TC	Diet	Fresh fruit intake	-0.07335	3.04E-07
D2AF	Biomarker	Vitamin D	Diet	Cereal intake	0.62944	4.70E-07
D2AF	Biomarker	Sodium in urine	Diet	Processed meat intake	0.00394	4.60E-07
D2AF	Biomarker	Sodium in urine	Diet	White bread intake	0.00448	2.87E-06
D2AF	Biomarker	Creatinine (enzymatic) in urine	Diet	Fresh fruit intake	-349.27164	5.23E-07
D2AF	Biomarker	HDL-C	Diet	Never eat cereal	0.67073	3.89E-06
D2AF	Biomarker	IGF 1	Diet	Cereal intake	0.18256	7.71E-07
D2AF	Biomarker	TC	Diet	Cereal intake	-0.03577	1.70E-06
D2AF	Biomarker	Urate	Diet	Never eat cereal	0.00289	1.80E-05
D2AF	Biomarker	GGT	Diet	Cereal intake	-2.00823	4.42E-06
D2AF	Biomarker	Sodium in urine	Diet	Wholegrain bread intake	-0.00407	5.68E-05
D2AF	Biomarker	Cystatin C	Diet	Full cream milk intake	1.74247	7.67E-05
D2AF	Biomarker	Vitamin D	Diet	Full cream milk intake	-0.01529	8.34E-05
D2AF	Biomarker	IGF 1	Diet	Wholegrain bread intake	0.02979	9.34E-05
D2AF	Biomarker	Sodium in urine	Diet	Never eat cereal	0.00447	1.11E-04
D2AF	Biomarker	AST	Diet	Eat wheat	-0.01259	3.48E-04
D2AF	Biomarker	LDL-C	Diet	Fresh fruit intake	-0.04559	1.02E-04
D2AF	Biomarker	IGF 1	Diet	Never eat cereal	-0.03324	6.21E-04
D2AF	Biomarker	Cystatin C	Diet	Skimmed milk intake	-1.43994	9.03E-04
D2AF	Biomarker	Creatinine (enzymatic) in urine	Diet	Cooked vegetable intake	-260.66561	2.29E-04
D2AF	Biomarker	Creatinine (enzymatic) in urine	Diet	Other cereal intake	0.00003	1.54E-03
D2AF	Biomarker	ALP	Diet	Instant coffee intake	0.00633	2.08E-03
D2AF	Biomarker	Sodium in urine	Diet	Muesli intake	-0.00576	2.19E-03
D2AF	Biomarker	Vitamin D	Diet	Never eat cereal	-0.00945	2.58E-03
D2AF	Biomarker	AST	Diet	Eat eggs dairy wheat sugar	-0.01131	3.34E-03
D2AF	Biomarker	Cystatin C	Diet	Non oily fish intake	-0.91877	6.72E-04
D2AF	Biomarker	AST	Diet	Eat eggs	-0.01158	4.10E-03
D2AF	Biomarker	IGF 1	Diet	White bread intake	-0.02639	5.06E-03
D2AF	Biomarker	Cystatin C	Diet	Cooked vegetable intake	-0.00653	8.83E-04
D2AF	Biomarker	GGT	Diet	Never eat cereal	0.00233	5.28E-03
D2AF	Biomarker	ALP	Diet	Ground coffee intake	-0.00862	5.68E-03
D2AF	Biomarker	Cystatin C	Diet	Salt added to food	0.86481	1.12E-03
D2AF	Biomarker	IGF 1	Diet	Oat cereal intake	0.03217	6.76E-03
D2AF	Biomarker	CRP	Diet	Cereal intake	-0.08687	1.20E-03
D2AF	Biomarker	HDL-C	Diet	Other cereal intake	-0.58454	7.90E-03
D2AF	Biomarker	Creatinine (enzymatic) in urine	Diet	Wholegrain bread intake	-0.00003	1.03E-02
D2AF	Biomarker	Sodium in urine	Diet	Oat cereal intake	-0.00448	1.05E-02
D2AF	Biomarker	Urea	Diet	Cereal intake	0.03198	2.11E-03
D2AF	Biomarker	Urea	Diet	Bread intake	0.01138	2.28E-03
D2AF	Biomarker	AST	Diet	Eat dairy	-0.01068	1.54E-02
D2AF	Biomarker	Creatinine (enzymatic) in urine	Diet	Cereal intake	-128.58148	3.35E-03
D2AF	Biomarker	HDL-C	Diet	Eat eggs	-0.49859	2.09E-02
D2AF	Biomarker	TC	Diet	Never eat cereal	0.13840	2.20E-02
D2AF	Biomarker	Apo-B	Diet	Fresh fruit intake	-0.01074	3.85E-03
D2AF	Biomarker	Apo-B	Diet	Cereal intake	-0.00545	4.51E-03
D2AF	Biomarker	Albumin	Diet	Salad / raw vegetable intake	0.10546	4.54E-03
D2AF	Biomarker	Vitamin D	Diet	Fresh fruit intake	0.85584	5.80E-03
D2AF	Biomarker	Sodium in urine	Diet	Other cereal intake	0.00364	3.60E-02
D2AF	Biomarker	Cystatin C	Diet	Ground coffee intake	-1.41241	3.69E-02
D2AF	Biomarker	Creatinine	Diet	Semi skimmed milk intake	0.01061	3.71E-02
D2AF	Biomarker	Cystatin C	Diet	Other cereal intake	1.05906	3.91E-02
D2AF	Biomarker	Cystatin C	Diet	Eat wheat	1.11025	3.96E-02

D2AF	Biomarker	Vitamin D	Diet	Salad / raw vegetable intake	0.77758	8.48E-03
D2AF	Biomarker	TG	Diet	Processed meat intake	0.11056	1.07E-02
D2AF	Biomarker	HDL-C	Diet	Non oily fish intake	0.34921	2.06E-02
D2AF	Biomarker	Urea	Diet	Lamb/mutton intake	0.08359	2.67E-02
D2AF	Biomarker	GGT	Diet	Lamb/mutton intake	0.00170	2.71E-02
D2AF	Biomarker	Apo-A	Diet	Fresh fruit intake	-0.00963	3.43E-02
D2AF	Biomarker	LDL-C	Diet	Cereal intake	-0.01691	4.92E-02
D2AF	Cell count	Mean sphered cell volume	Diet	Never eat cereal	0.05429	2.05E-14
D2AF	Cell count	MCV	Diet	Never eat cereal	0.05706	4.72E-11
D2AF	Cell count	Mean corpuscular haemoglobin	Diet	Never eat cereal	0.14481	9.99E-11
D2AF	Cell count	Mean reticulocyte volume	Diet	Never eat cereal	0.03299	8.54E-09
D2AF	Cell count	WBC count	Diet	White bread intake	0.11303	2.89E-07
D2AF	Cell count	RBC count	Diet	Eat eggs dairy wheat sugar	0.57114	1.03E-06
D2AF	Cell count	Neutrophil count	Diet	White bread intake	0.14187	1.12E-06
D2AF	Cell count	RBC count	Diet	Eat dairy	0.56925	2.26E-06
D2AF	Cell count	RBC count	Diet	Eat eggs	0.58217	2.60E-06
D2AF	Cell count	RBC count	Diet	Eat wheat	0.53699	2.14E-05
D2AF	Cell count	Neutrophil count	Diet	Ground coffee intake	-0.17542	6.80E-04
D2AF	Cell count	RBC count	Diet	Never eat cereal	-0.46556	7.72E-04
D2AF	Cell count	WBC count	Diet	Wholegrain bread intake	-0.08651	8.68E-04
D2AF	Cell count	WBC count	Diet	Ground coffee intake	-0.13163	1.30E-03
D2AF	Cell count	HLSRP	Diet	Never eat cereal	0.80868	2.24E-03
D2AF	Cell count	RDW	Diet	Full cream milk intake	0.24535	2.32E-03
D2AF	Cell count	Neutrophil count	Diet	Wholegrain bread intake	-0.10667	2.89E-03
D2AF	Cell count	MCV	Diet	Eat eggs	-0.03581	3.87E-03
D2AF	Cell count	Mean corpuscular haemoglobin	Diet	Oat cereal intake	-0.09758	4.14E-03
D2AF	Cell count	Reticulocyte percentage	Diet	Never eat cereal	0.30203	8.64E-03
D2AF	Cell count	Immature reticulocyte fraction	Diet	Never eat cereal	2.79242	1.09E-02
D2AF	Cell count	Mean sphered cell volume	Diet	Semi skimmed milk intake	-0.02412	1.52E-02
D2AF	Cell count	WBC count	Diet	Instant coffee intake	0.08405	1.62E-02
D2AF	Cell count	Haemoglobin concentration	Diet	Eat eggs dairy wheat sugar	0.14448	1.86E-02
D2AF	Cell count	Haemoglobin concentration	Diet	Eat dairy	0.14645	2.01E-02
D2AF	Cell count	Haematocrit percentage	Diet	Eat dairy	0.04964	2.61E-02
D2AF	Cell count	Haematocrit percentage	Diet	Eat eggs dairy wheat sugar	0.04837	2.99E-02
D2AF	Cell count	MCV	Diet	Oat cereal intake	-0.03418	3.59E-02
D2AF	Cell count	RDW	Diet	Wholegrain bread intake	-0.14506	3.98E-02
D2AF	Cell count	MCV	Diet	Eat eggs dairy wheat sugar	-0.03010	4.63E-02
D2AF	Cell count	Mean sphered cell volume	Diet	Cereal intake	-0.27703	3.34E-16
D2AF	Cell count	MCV	Diet	Cereal intake	-0.20144	1.91E-11
D2AF	Cell count	Mean corpuscular haemoglobin	Diet	Cereal intake	-0.07645	9.64E-11
D2AF	Cell count	HLSRP	Diet	Cereal intake	-0.00708	1.42E-07
D2AF	Cell count	Mean reticulocyte volume	Diet	Cereal intake	-0.25439	5.63E-07
D2AF	Cell count	Immature reticulocyte fraction	Diet	Cereal intake	-0.00190	6.35E-07
D2AF	Cell count	Reticulocyte percentage	Diet	Cereal intake	-0.01617	4.97E-06
D2AF	Cell count	HLSRC	Diet	Cereal intake	-0.00029	8.31E-06
D2AF	Cell count	Neutrophil count	Diet	Fresh fruit intake	-0.07042	3.73E-04
D2AF	Cell count	Neutrophil count	Diet	Cereal intake	-0.03580	4.37E-04
D2AF	Cell count	WBC count	Diet	Cereal intake	-0.04636	4.74E-04
D2AF	Cell count	Reticulocyte count	Diet	Cereal intake	-0.00062	4.99E-04
D2AF	Cell count	WBC count	Diet	Fresh fruit intake	-0.08745	1.02E-03
D2AF	Cell count	RDW	Diet	Cooked vegetable intake	-0.03632	5.97E-03
D2AF	Cell count	Mean sphered cell volume	Diet	Fresh fruit intake	-0.22874	3.37E-02
D2AF	Cell count	Mean sphered cell volume	Diet	Salt added to food	0.02658	1.28E-05
D2AF	Cell count	WBC count	Diet	Salt added to food	0.07993	7.64E-05
D2AF	Cell count	MCV	Diet	Salt added to food	0.02762	2.56E-04

D2AF	Cell count	WBC count	Diet	Processed meat intake	0.07296	4.60E-04
D2AF	Cell count	Neutrophil count	Diet	Processed meat intake	0.08477	4.55E-03
D2AF	Cell count	Neutrophil count	Diet	Salt added to food	0.08354	8.61E-03
D2AF	Cell count	Mean corpuscular haemoglobin	Diet	Salt added to food	0.05842	1.15E-02
D2AF	Cell count	Mean reticulocyte volume	Diet	Salt added to food	0.01325	4.76E-02
D2AF	Biomarker	HDL-C	Cell count	Mean sphered cell volume	4.37776	5.18E-54
D2AF	Biomarker	AST	Cell count	Mean corpuscular haemoglobin	0.03235	1.12E-49
D2AF	Biomarker	Apo-A	Cell count	Mean sphered cell volume	5.44134	6.11E-48
D2AF	Biomarker	AST	Cell count	MCV	0.07879	8.02E-45
D2AF	Biomarker	GGT	Cell count	Mean corpuscular haemoglobin	0.00632	3.57E-42
D2AF	Biomarker	GGT	Cell count	Mean sphered cell volume	0.01826	2.47E-39
D2AF	Biomarker	GGT	Cell count	MCV	0.01507	2.57E-36
D2AF	Biomarker	AST	Cell count	Mean sphered cell volume	0.08234	1.50E-35
D2AF	Biomarker	AST	Cell count	Platelet count	-0.82835	2.24E-32
D2AF	Biomarker	TG	Cell count	Reticulocyte count	0.00444	8.79E-31
D2AF	Biomarker	Apo-A	Cell count	MCV	3.73931	4.59E-30
D2AF	Biomarker	AST	Cell count	Platelet crit	-0.00064	6.40E-30
D2AF	Biomarker	HDL-C	Cell count	MCV	2.78799	7.51E-29
D2AF	Biomarker	ALT	Cell count	Mean corpuscular haemoglobin	0.01960	1.88E-28
D2AF	Biomarker	TG	Cell count	Reticulocyte percentage	0.09228	1.30E-27
D2AF	Biomarker	ALT	Cell count	Haemoglobin concentration	0.01068	1.42E-27
D2AF	Biomarker	Phosphate	Cell count	Lymphocyte count	0.75635	7.17E-27
D2AF	Biomarker	ALT	Cell count	Reticulocyte count	0.00024	1.24E-26
D2AF	Biomarker	SHBG	Cell count	Mean sphered cell volume	0.04417	3.35E-26
D2AF	Biomarker	Apo-B	Cell count	Haemoglobin concentration	0.75321	1.17E-24
D2AF	Biomarker	CRP	Cell count	Neutrophil count	0.06538	2.61E-24
D2AF	Biomarker	Total bilirubin	Cell count	Platelet crit	-0.00208	3.01E-24
D2AF	Biomarker	Total bilirubin	Cell count	Platelet count	-2.57919	3.30E-24
D2AF	Biomarker	ALT	Cell count	Reticulocyte percentage	0.00500	3.36E-24
D2AF	Biomarker	Apo-B	Cell count	Haematocrit percentage	2.12824	6.05E-23
D2AF	Biomarker	Apo-A	Cell count	Mean corpuscular haemoglobin	1.27772	1.45E-22
D2AF	Biomarker	TG	Cell count	Platelet distribution width	0.08202	2.06E-22
D2AF	Biomarker	Total bilirubin	Cell count	Haemoglobin concentration	0.04380	2.26E-22
D2AF	Biomarker	TG	Cell count	HLSRC	0.00151	2.32E-22
D2AF	Biomarker	IGF 1	Cell count	Mean sphered cell volume	-0.16186	2.81E-22
D2AF	Biomarker	CRP	Cell count	WBC count	0.08151	3.08E-22
D2AF	Biomarker	Calcium	Cell count	Haemoglobin concentration	1.85724	2.04E-21
D2AF	Biomarker	LDL-C	Cell count	Haematocrit percentage	0.56496	8.46E-21
D2AF	Biomarker	LDL-C	Cell count	Haemoglobin concentration	0.19328	8.65E-21
D2AF	Biomarker	TG	Cell count	HLSRP	0.03179	1.40E-20
D2AF	Biomarker	TG	Cell count	Lymphocyte count	0.09966	1.67E-20
D2AF	Biomarker	Calcium	Cell count	Haematocrit percentage	5.25944	5.18E-20
D2AF	Biomarker	IGF 1	Cell count	RBC count	0.01042	1.03E-19
D2AF	Biomarker	Cystatin C	Cell count	RDW	0.92052	1.66E-19
D2AF	Biomarker	HbA1c	Cell count	Lymphocyte count	0.01776	1.89E-19
D2AF	Biomarker	ALT	Cell count	HLSRC	0.00008	3.84E-19
D2AF	Biomarker	IGF 1	Cell count	Mean corpuscular haemoglobin	-0.05013	8.72E-19
D2AF	Biomarker	HDL-C	Cell count	Mean reticulocyte volume	3.61415	9.97E-19

D2AF	Biomarker	Calcium	Cell count	RBC count	0.63659	2.12E-18
D2AF	Biomarker	CRP	Cell count	Lymphocyte percentage	-0.30618	2.58E-18
D2AF	Biomarker	ALT	Cell count	MCV	0.04066	3.78E-18
D2AF	Biomarker	Cystatin C	Cell count	WBC count	1.78329	6.04E-18
D2AF	Biomarker	IGF 1	Cell count	MCV	-0.12531	7.08E-18
D2AF	Biomarker	ALT	Cell count	HLSRP	0.00170	1.53E-17
D2AF	Biomarker	Urate	Cell count	Reticulocyte percentage	0.00105	2.01E-17
D2AF	Biomarker	Urate	Cell count	HLSRP	0.00041	3.54E-17
D2AF	Biomarker	Calcium	Cell count	Platelet crit	0.07669	3.55E-17
D2AF	Biomarker	HDL-C	Cell count	Mean corpuscular haemoglobin	0.85178	5.53E-17
D2AF	Biomarker	Urate	Cell count	HLSRC	0.00002	7.18E-17
D2AF	Biomarker	HbA1c	Cell count	Mean corpuscular haemoglobin	-0.04862	8.67E-17
D2AF	Biomarker	Albumin	Cell count	Haemoglobin concentration	0.05893	1.09E-16
D2AF	Biomarker	TC	Cell count	Haemoglobin concentration	0.13089	1.11E-16
D2AF	Biomarker	Albumin	Cell count	RDW	-0.04884	3.35E-16
D2AF	Biomarker	LDL-C	Cell count	RBC count	0.06331	3.58E-16
D2AF	Biomarker	Apo-B	Cell count	RBC count	0.22752	4.01E-16
D2AF	Biomarker	TG	Cell count	MCHC	0.12674	8.88E-16
D2AF	Biomarker	SHBG	Cell count	Platelet count	-0.36603	1.15E-15
D2AF	Biomarker	TC	Cell count	Haematocrit percentage	0.36936	1.90E-15
D2AF	Biomarker	GGT	Cell count	Platelet count	-0.12365	2.49E-15
D2AF	Biomarker	Calcium	Cell count	Lymphocyte count	0.97219	6.07E-15
D2AF	Biomarker	Urate	Cell count	Reticulocyte count	0.00004	1.05E-14
D2AF	Biomarker	ALT	Cell count	Haematocrit percentage	0.02329	1.20E-14
D2AF	Biomarker	Apo-A	Cell count	Mean reticulocyte volume	4.24004	1.41E-14
D2AF	Biomarker	GGT	Cell count	Mean reticulocyte volume	0.01562	1.46E-14
D2AF	Biomarker	HbA1c	Cell count	WBC count	0.04243	1.99E-14
D2AF	Biomarker	Total bilirubin	Cell count	Haematocrit percentage	0.10418	2.75E-14
D2AF	Biomarker	HDL-C	Cell count	RBC count	-0.15631	1.49E-13
D2AF	Biomarker	AST	Cell count	Reticulocyte percentage	0.00485	1.51E-13
D2AF	Biomarker	GGT	Cell count	Platelet crit	-0.00009	2.89E-13
D2AF	Biomarker	Urea	Cell count	Mean sphered cell volume	-0.53854	4.75E-13
D2AF	Biomarker	AST	Cell count	Monocyte percentage	0.01996	7.11E-13
D2AF	Biomarker	Calcium	Cell count	Platelet count	83.13003	9.88E-13
D2AF	Biomarker	Creatinine	Cell count	Mean corpuscular haemoglobin	-0.01826	1.16E-12
D2AF	Biomarker	Albumin	Cell count	Haematocrit percentage	0.15137	1.53E-12
D2AF	Biomarker	Total bilirubin	Cell count	Mean corpuscular haemoglobin	0.06057	1.99E-12
D2AF	Biomarker	ALT	Cell count	Platelet count	-0.41996	2.04E-12
D2AF	Biomarker	Cystatin C	Cell count	Neutrophil count	1.15709	2.58E-12
D2AF	Biomarker	CRP	Cell count	Neutrophil percentage	0.28741	5.01E-12
D2AF	Biomarker	ALT	Cell count	Platelet crit	-0.00033	6.70E-12
D2AF	Biomarker	IGF 1	Cell count	Platelet count	1.26717	1.18E-11
D2AF	Biomarker	HDL-C	Cell count	Reticulocyte count	-0.00902	1.29E-11
D2AF	Biomarker	TG	Cell count	Mean reticulocyte volume	-0.91431	1.35E-11
D2AF	Biomarker	Total bilirubin	Cell count	Reticulocyte count	0.00073	2.23E-11
D2AF	Biomarker	IGF 1	Cell count	Platelet crit	0.00100	3.07E-11
D2AF	Biomarker	TC	Cell count	Platelet count	6.19290	4.01E-11
D2AF	Biomarker	CRP	Cell count	RDW	0.02950	4.46E-11
D2AF	Biomarker	SHBG	Cell count	Platelet crit	-0.00025	8.88E-11
D2AF	Biomarker	AST	Cell count	Reticulocyte count	0.00020	1.53E-10
D2AF	Biomarker	Urea	Cell count	Mean corpuscular haemoglobin	-0.16398	1.68E-10
D2AF	Biomarker	Creatinine	Cell count	MCV	-0.04294	1.83E-10
D2AF	Biomarker	CRP	Cell count	Monocyte count	0.00567	1.87E-10
D2AF	Biomarker	LDL-C	Cell count	Platelet count	8.03884	1.90E-10
D2AF	Biomarker	Urea	Cell count	MCV	-0.41820	1.98E-10

D2AF	Biomarker	AST	Cell count	Mean reticulocyte volume	0.06335	2.72E-10
D2AF	Biomarker	Albumin	Cell count	RBC count	0.01734	3.97E-10
D2AF	Biomarker	GGT	Cell count	RBC count	-0.00067	4.40E-10
D2AF	Biomarker	Creatinine	Cell count	Mean sphered cell volume	-0.04926	4.52E-10
D2AF	Biomarker	CRP	Cell count	Platelet count	1.90774	5.09E-10
D2AF	Biomarker	CRP	Cell count	Platelet crit	0.00153	5.44E-10
D2AF	Biomarker	AST	Cell count	Neutrophil percentage	-0.06624	1.22E-09
D2AF	Biomarker	ALT	Cell count	MCHC	0.00584	1.55E-09
D2AF	Biomarker	SHBG	Cell count	MCV	0.02371	2.54E-09
D2AF	Biomarker	SHBG	Cell count	Mean reticulocyte volume	0.03759	2.57E-09
D2AF	Biomarker	SHBG	Cell count	HLSRC	-0.00005	3.36E-09
D2AF	Biomarker	Microalbumin in urine	Cell count	Neutrophil count	0.00611	3.75E-09
D2AF	Biomarker	HDL-C	Cell count	HLSRC	-0.00322	3.77E-09
D2AF	Biomarker	Phosphate	Cell count	Platelet crit	0.03433	3.99E-09
D2AF	Biomarker	Urea	Cell count	Haemoglobin concentration	-0.08505	4.66E-09
D2AF	Biomarker	Phosphate	Cell count	WBC count	1.25969	5.95E-09
D2AF	Biomarker	Calcium	Cell count	WBC count	2.12168	6.29E-09
D2AF	Biomarker	TG	Cell count	Haemoglobin concentration	0.11065	8.07E-09
D2AF	Biomarker	HbA1c	Cell count	MCV	-0.09291	1.09E-08
D2AF	Biomarker	TG	Cell count	WBC count	0.18582	1.13E-08
D2AF	Biomarker	Phosphate	Cell count	Platelet count	41.60185	1.27E-08
D2AF	Biomarker	HDL-C	Cell count	Platelet distribution width	-0.16950	1.50E-08
D2AF	Biomarker	GGT	Cell count	Platelet distribution width	0.00082	1.51E-08
D2AF	Biomarker	ALT	Cell count	Neutrophil percentage	-0.04940	1.91E-08
D2AF	Biomarker	Apo-A	Cell count	RBC count	-0.16683	2.14E-08
D2AF	Biomarker	GGT	Cell count	HLSRP	0.00033	2.20E-08
D2AF	Biomarker	SHBG	Cell count	RBC count	-0.00180	2.31E-08
D2AF	Biomarker	Phosphate	Cell count	Neutrophil percentage	-5.69246	2.68E-08
D2AF	Biomarker	AST	Cell count	Neutrophil count	-0.01024	2.69E-08
D2AF	Biomarker	GGT	Cell count	Reticulocyte percentage	0.00083	2.71E-08
D2AF	Biomarker	AST	Cell count	HLSRP	0.00154	3.15E-08
D2AF	Biomarker	Phosphate	Cell count	Eosinophil count	0.07979	3.72E-08
D2AF	Biomarker	Urate	Cell count	Immature reticulocyte fraction	0.00008	4.51E-08
D2AF	Biomarker	ALT	Cell count	RDW	-0.00515	4.95E-08
D2AF	Biomarker	Total bilirubin	Cell count	WBC count	-0.04686	5.08E-08
D2AF	Biomarker	ALT	Cell count	Lymphocyte percentage	0.04257	5.12E-08
D2AF	Biomarker	Microalbumin in urine	Cell count	WBC count	0.00749	5.72E-08
D2AF	Biomarker	TG	Cell count	Mean sphered cell volume	-0.57161	5.97E-08
D2AF	Biomarker	Apo-B	Cell count	Platelet count	25.73464	6.07E-08
D2AF	Biomarker	AST	Cell count	Platelet distribution width	0.00376	6.97E-08
D2AF	Biomarker	Cystatin C	Cell count	Eosinophil count	0.08160	7.44E-08
D2AF	Biomarker	GGT	Cell count	Monocyte percentage	0.00340	7.84E-08
D2AF	Biomarker	Phosphate	Cell count	Monocyte count	0.11646	8.11E-08
D2AF	Biomarker	Total protein	Cell count	Monocyte count	0.00430	1.05E-07
D2AF	Biomarker	Total bilirubin	Cell count	Reticulocyte percentage	0.01355	1.07E-07
D2AF	Biomarker	IGF 1	Cell count	Neutrophil percentage	0.14682	1.29E-07
D2AF	Biomarker	SHBG	Cell count	Reticulocyte count	-0.00011	1.44E-07
D2AF	Biomarker	HbA1c	Cell count	RDW	0.01581	1.95E-07
D2AF	Biomarker	ALT	Cell count	Platelet distribution width	0.00288	2.27E-07
D2AF	Biomarker	AST	Cell count	HLSRC	0.00007	3.49E-07
D2AF	Biomarker	ALT	Cell count	Monocyte percentage	0.01224	3.75E-07
D2AF	Biomarker	AST	Cell count	Haemoglobin concentration	0.00740	3.77E-07
D2AF	Biomarker	SHBG	Cell count	HLSRP	-0.00090	4.83E-07
D2AF	Biomarker	Total bilirubin	Cell count	Lymphocyte count	-0.01619	5.18E-07
D2AF	Biomarker	TG	Cell count	RDW	-0.08415	6.03E-07

D2AF	Biomarker	Urea	Cell count	Haematocrit percentage	-0.22188	6.69E-07
D2AF	Biomarker	LDL-C	Cell count	Platelet crit	0.00541	8.22E-07
D2AF	Biomarker	Phosphate	Cell count	Lymphocyte percentage	4.59689	1.00E-06
D2AF	Biomarker	Cystatin C	Cell count	Monocyte count	0.11368	1.02E-06
D2AF	Biomarker	Total bilirubin	Cell count	MCV	0.11848	1.28E-06
D2AF	Biomarker	Albumin	Cell count	Mean reticulocyte volume	-0.28941	1.28E-06
D2AF	Biomarker	Testosterone	Cell count	Mean sphered cell volume	0.15510	1.56E-06
D2AF	Biomarker	Albumin	Cell count	Reticulocyte count	0.00088	2.06E-06
D2AF	Biomarker	HDL-C	Cell count	Reticulocyte percentage	-0.15348	2.12E-06
D2AF	Biomarker	SHBG	Cell count	Mean corpuscular haemoglobin	0.00793	2.12E-06
D2AF	Biomarker	ALT	Cell count	Neutrophil count	-0.00722	2.63E-06
D2AF	Biomarker	HbA1c	Cell count	MCHC	-0.01552	3.12E-06
D2AF	Biomarker	Total protein	Cell count	Lymphocyte count	0.01446	5.63E-06
D2AF	Biomarker	ALP	Cell count	Platelet distribution width	0.00169	5.75E-06
D2AF	Biomarker	HDL-C	Cell count	Lymphocyte count	-0.18489	6.15E-06
D2AF	Biomarker	Total protein	Cell count	RBC count	0.00852	7.95E-06
D2AF	Biomarker	TC	Cell count	Platelet crit	0.00378	9.76E-06
D2AF	Biomarker	IGF 1	Cell count	Mean reticulocyte volume	-0.12175	9.80E-06
D2AF	Biomarker	Apo-B	Cell count	Platelet crit	0.01815	1.02E-05
D2AF	Biomarker	SHBG	Cell count	WBC count	-0.00707	1.13E-05
D2AF	Biomarker	ALP	Cell count	WBC count	0.00572	1.25E-05
D2AF	Biomarker	SHBG	Cell count	Immature reticulocyte fraction	-0.00023	1.29E-05
D2AF	Biomarker	GGT	Cell count	HLSRC	0.00001	1.67E-05
D2AF	Biomarker	TC	Cell count	RBC count	0.02980	1.97E-05
D2AF	Biomarker	HbA1c	Cell count	Neutrophil count	0.02157	2.03E-05
D2AF	Biomarker	AST	Cell count	Lymphocyte percentage	0.04602	2.09E-05
D2AF	Biomarker	IGF 1	Cell count	Lymphocyte percentage	-0.11269	2.11E-05
D2AF	Biomarker	Apo-A	Cell count	Platelet distribution width	-0.18334	2.86E-05
D2AF	Biomarker	SHBG	Cell count	Mean platelet volume	0.00414	3.19E-05
D2AF	Biomarker	TG	Cell count	Immature reticulocyte fraction	0.00491	3.23E-05
D2AF	Biomarker	Total protein	Cell count	Haematocrit percentage	0.06465	3.63E-05
D2AF	Biomarker	HDL-C	Cell count	HLSRP	-0.05548	4.30E-05
D2AF	Biomarker	Albumin	Cell count	Mean sphered cell volume	-0.19188	4.43E-05
D2AF	Biomarker	IGF 1	Cell count	Monocyte percentage	-0.03309	4.78E-05
D2AF	Biomarker	ALP	Cell count	RDW	0.00271	5.34E-05
D2AF	Biomarker	TC	Cell count	Mean platelet volume	-0.08026	5.50E-05
D2AF	Biomarker	HbA1c	Cell count	Platelet crit	0.00073	5.81E-05
D2AF	Biomarker	ALT	Cell count	Mean sphered cell volume	0.02722	6.31E-05
D2AF	Biomarker	Phosphate	Cell count	Basophil count	0.02138	6.51E-05
D2AF	Biomarker	Testosterone	Cell count	Haematocrit percentage	0.07465	7.95E-05
D2AF	Biomarker	ALP	Cell count	Neutrophil count	0.00413	9.11E-05
D2AF	Biomarker	SHBG	Cell count	Neutrophil count	-0.00508	9.60E-05
D2AF	Biomarker	Testosterone	Cell count	Haemoglobin concentration	0.02526	1.11E-04
D2AF	Biomarker	Urate	Cell count	MCHC	0.00107	1.29E-04
D2AF	Biomarker	Testosterone	Cell count	MCV	0.11593	1.39E-04
D2AF	Biomarker	GGT	Cell count	Immature reticulocyte fraction	0.00007	1.45E-04
D2AF	Biomarker	Cystatin C	Cell count	Lymphocyte percentage	-4.08643	1.70E-04
D2AF	Biomarker	Glucose	Cell count	Reticulocyte percentage	0.04311	2.28E-04
D2AF	Biomarker	Microalbumin in urine	Cell count	RDW	0.00294	2.29E-04
D2AF	Biomarker	SHBG	Cell count	Reticulocyte percentage	-0.00187	2.72E-04
D2AF	Biomarker	Total bilirubin	Cell count	Neutrophil count	-0.02806	3.14E-04
D2AF	Biomarker	Apo-B	Cell count	Lymphocyte count	0.21718	3.85E-04
D2AF	Biomarker	AST	Cell count	Mean platelet volume	0.00613	3.90E-04
D2AF	Biomarker	HDL-C	Cell count	WBC count	-0.44031	4.26E-04
D2AF	Biomarker	SHBG	Cell count	RDW	0.00311	4.28E-04

D2AF	Biomarker	SHBG	Cell count	Monocyte percentage	0.00774	4.48E-04
D2AF	Biomarker	Testosterone	Cell count	Mean reticulocyte volume	0.17613	4.52E-04
D2AF	Biomarker	IGF 1	Cell count	Haematocrit percentage	0.04229	5.64E-04
D2AF	Biomarker	Total protein	Cell count	Haemoglobin concentration	0.02008	5.69E-04
D2AF	Biomarker	Total bilirubin	Cell count	MCHC	0.01866	5.87E-04
D2AF	Biomarker	GGT	Cell count	Monocyte count	0.00021	5.88E-04
D2AF	Biomarker	AST	Cell count	MCHC	0.00520	6.75E-04
D2AF	Biomarker	HDL-C	Cell count	Eosinophil count	-0.02845	7.06E-04
D2AF	Biomarker	CRP	Cell count	Basophil count	0.00084	7.92E-04
D2AF	Biomarker	IGF 1	Cell count	Basophil percentage	-0.00522	7.93E-04
D2AF	Biomarker	Calcium	Cell count	RDW	-0.74671	1.10E-03
D2AF	Biomarker	Sodium in urine	Cell count	Immature reticulocyte fraction	0.00010	1.21E-03
D2AF	Biomarker	HbA1c	Cell count	Eosinophil count	0.00159	1.37E-03
D2AF	Biomarker	Urate	Cell count	Mean corpuscular haemoglobin	0.00199	1.48E-03
D2AF	Biomarker	Apo-A	Cell count	Mean platelet volume	-0.33333	1.50E-03
D2AF	Biomarker	AST	Cell count	RBC count	-0.00209	1.54E-03
D2AF	Biomarker	GGT	Cell count	Reticulocyte count	0.00003	1.98E-03
D2AF	Biomarker	Testosterone	Cell count	Mean corpuscular haemoglobin	0.04062	2.33E-03
D2AF	Biomarker	TG	Cell count	Lymphocyte percentage	0.52932	2.48E-03
D2AF	Biomarker	Total protein	Cell count	Neutrophil percentage	-0.14911	2.51E-03
D2AF	Biomarker	Glucose	Cell count	HLSRP	0.01543	2.76E-03
D2AF	Biomarker	Total protein	Cell count	Platelet crit	0.00085	3.33E-03
D2AF	Biomarker	Total bilirubin	Cell count	HLSRC	0.00018	3.57E-03
D2AF	Biomarker	Cystatin C	Cell count	Eosinophil percentage	0.75798	4.01E-03
D2AF	Biomarker	CRP	Cell count	Haemoglobin concentration	-0.02120	5.12E-03
D2AF	Biomarker	LDL-C	Cell count	Mean platelet volume	-0.08968	5.23E-03
D2AF	Biomarker	IGF 1	Cell count	Platelet distribution width	-0.00641	5.36E-03
D2AF	Biomarker	IGF 1	Cell count	Neutrophil count	0.01703	5.99E-03
D2AF	Biomarker	CRP	Cell count	Immature reticulocyte fraction	0.00115	6.66E-03
D2AF	Biomarker	Calcium	Cell count	Reticulocyte count	0.01856	7.44E-03
D2AF	Biomarker	Total protein	Cell count	Platelet count	1.01978	7.51E-03
D2AF	Biomarker	Phosphate	Cell count	Eosinophil percentage	0.70745	7.75E-03
D2AF	Biomarker	HbA1c	Cell count	Immature reticulocyte fraction	0.00072	9.35E-03
D2AF	Biomarker	Microalbumin in urine	Cell count	Lymphocyte percentage	-0.02070	1.05E-02
D2AF	Biomarker	Microalbumin in urine	Cell count	Mean sphered cell volume	0.01607	1.05E-02
D2AF	Biomarker	Total bilirubin	Cell count	Eosinophil count	-0.00207	1.08E-02
D2AF	Biomarker	HbA1c	Cell count	Monocyte count	0.00213	1.28E-02
D2AF	Biomarker	Vitamin D	Cell count	WBC count	-0.00626	1.30E-02
D2AF	Biomarker	AST	Cell count	WBC count	-0.00869	1.32E-02
D2AF	Biomarker	Glucose	Cell count	Reticulocyte count	0.00166	1.33E-02
D2AF	Biomarker	Calcium	Cell count	Lymphocyte percentage	5.45991	1.63E-02
D2AF	Biomarker	Glucose	Cell count	Haematocrit percentage	-0.21020	1.70E-02
D2AF	Biomarker	Microalbumin in urine	Cell count	Neutrophil percentage	0.02285	1.81E-02
D2AF	Biomarker	Cystatin C	Cell count	Reticulocyte count	0.01087	2.05E-02
D2AF	Biomarker	ALT	Cell count	Immature reticulocyte fraction	0.00022	2.14E-02
D2AF	Biomarker	AST	Cell count	Haematocrit percentage	0.01463	2.19E-02
D2AF	Biomarker	Glucose	Cell count	Platelet crit	-0.00324	2.28E-02
D2AF	Biomarker	HbA1c	Cell count	Basophil count	0.00045	2.92E-02
D2AF	Biomarker	AST	Cell count	Monocyte count	0.00082	2.95E-02
D2AF	Biomarker	IGF 1	Cell count	HLSRP	-0.00237	3.07E-02
D2AF	Biomarker	Glucose	Cell count	Monocyte count	-0.01146	3.47E-02
D2AF	Biomarker	Urea	Cell count	Mean reticulocyte volume	-0.36818	4.35E-02
D2AF	Biomarker	IGF 1	Cell count	Haemoglobin concentration	0.01174	4.82E-02
D2AF	Biomarker	Total bilirubin	Cell count	RBC count	0.00628	4.90E-02

DW2	Cell count	Mean sphered cell volume	Diet	Cereal intake	-0.41529	3.10E-02
DW2	Biomarker	Albumin	Cell count	RBC count	0.05841	1.62E-14
DW2	Biomarker	Total bilirubin	Cell count	Platelet crit	-0.00349	6.21E-12
DW2	Biomarker	Total bilirubin	Cell count	Platelet count	-4.33439	9.39E-12
DW2	Biomarker	Albumin	Cell count	Haemoglobin concentration	0.15228	2.83E-11
DW2	Biomarker	Total bilirubin	Cell count	Mean sphered cell volume	0.40959	1.08E-10
DW2	Biomarker	Total bilirubin	Cell count	MCV	0.33276	1.12E-09
DW2	Biomarker	AST	Cell count	Mean sphered cell volume	0.10381	1.48E-09
DW2	Biomarker	IGF 1	Cell count	Platelet count	3.65058	2.40E-09
DW2	Biomarker	SHBG	Cell count	Mean sphered cell volume	0.08550	4.90E-09
DW2	Biomarker	IGF 1	Cell count	Mean sphered cell volume	-0.35245	6.01E-09
DW2	Biomarker	Albumin	Cell count	Mean reticulocyte volume	-0.81501	7.60E-09
DW2	Biomarker	Albumin	Cell count	Haematocrit percentage	0.38552	1.02E-08
DW2	Biomarker	Total bilirubin	Cell count	Mean corpuscular haemoglobin	0.12303	1.36E-08
DW2	Biomarker	SHBG	Cell count	Mean reticulocyte volume	0.09641	1.85E-08
DW2	Biomarker	Albumin	Cell count	Mean sphered cell volume	-0.69025	1.97E-08
DW2	Biomarker	AST	Cell count	Platelet count	-0.98075	5.90E-08
DW2	Biomarker	IGF 1	Cell count	Platelet crit	0.00269	1.31E-07
DW2	Biomarker	SHBG	Cell count	RBC count	-0.00517	7.80E-07
DW2	Biomarker	LDL-C	Cell count	Platelet count	23.29152	9.83E-07
DW2	Biomarker	CRP	Cell count	Neutrophil count	0.10005	1.03E-06
DW2	Biomarker	ALP	Cell count	Mean sphered cell volume	0.04317	1.51E-06
DW2	Biomarker	AST	Cell count	Platelet crit	-0.00072	1.71E-06
DW2	Biomarker	Total bilirubin	Cell count	Mean reticulocyte volume	0.38791	1.75E-06
DW2	Biomarker	TC	Cell count	Platelet count	17.40191	2.26E-06
DW2	Biomarker	IGF 1	Cell count	Mean reticulocyte volume	-0.34787	5.71E-06
DW2	Biomarker	IGF 1	Cell count	RBC count	0.02011	8.01E-06
DW2	Biomarker	HbA1c	Cell count	Mean corpuscular haemoglobin	-0.10774	9.88E-06
DW2	Biomarker	Urea	Cell count	Mean sphered cell volume	-1.33486	1.08E-05
DW2	Biomarker	Apo-B	Cell count	Platelet count	78.56808	1.41E-05
DW2	Biomarker	ALP	Cell count	Mean reticulocyte volume	0.04682	1.62E-05
DW2	Biomarker	Albumin	Cell count	RDW	-0.10486	1.87E-05
DW2	Biomarker	AST	Cell count	MCV	0.06894	3.84E-05
DW2	Biomarker	LDL-C	Cell count	Platelet crit	0.01630	9.98E-05
DW2	Biomarker	SHBG	Cell count	MCV	0.05582	1.31E-04
DW2	Biomarker	GGT	Cell count	Mean sphered cell volume	0.01619	1.57E-04
DW2	Biomarker	CRP	Cell count	WBC count	0.11370	1.59E-04
DW2	Biomarker	IGF 1	Cell count	RDW	-0.04881	1.60E-04
DW2	Biomarker	TG	Cell count	Haemoglobin concentration	0.31159	1.72E-04
DW2	Biomarker	IGF 1	Cell count	Mean platelet volume	-0.04660	2.31E-04
DW2	Biomarker	TC	Cell count	Platelet crit	0.01209	2.35E-04
DW2	Biomarker	TC	Cell count	Haematocrit percentage	0.81639	2.51E-04
DW2	Biomarker	SHBG	Cell count	Platelet count	-0.65506	2.61E-04
DW2	Biomarker	TG	Cell count	Mean reticulocyte volume	-1.79160	3.78E-04
DW2	Biomarker	HbA1c	Cell count	MCV	-0.24579	4.22E-04
DW2	Biomarker	AST	Cell count	Mean corpuscular haemoglobin	0.02442	5.51E-04
DW2	Biomarker	Total bilirubin	Cell count	Platelet distribution width	0.02451	5.52E-04

DW2	Biomarker	TG	Cell count	RBC count	0.10292	5.79E-04
DW2	Biomarker	LDL-C	Cell count	Haematocrit percentage	1.02368	8.15E-04
DW2	Biomarker	Apo-B	Cell count	Platelet crit	0.05473	8.55E-04
DW2	Biomarker	TG	Cell count	Haematocrit percentage	0.81749	1.13E-03
DW2	Biomarker	Cystatin C	Cell count	Lymphocyte percentage	-8.24615	2.05E-03
DW2	Biomarker	AST	Cell count	Mean reticulocyte volume	0.08157	2.14E-03
DW2	Biomarker	TC	Cell count	Haemoglobin concentration	0.26708	2.15E-03
DW2	Biomarker	IGF 1	Cell count	MCV	-0.20724	2.32E-03
DW2	Biomarker	ALP	Cell count	MCV	0.02873	2.34E-03
DW2	Biomarker	CRP	Cell count	Immature reticulocyte fraction	0.00301	2.74E-03
DW2	Biomarker	Calcium	Cell count	RBC count	1.09643	3.22E-03
DW2	Biomarker	Urea	Cell count	Mean reticulocyte volume	-1.26255	3.35E-03
DW2	Biomarker	HDL-C	Cell count	Mean sphered cell volume	4.86054	3.50E-03
DW2	Biomarker	ALP	Cell count	RDW	0.00597	4.04E-03
DW2	Biomarker	SHBG	Cell count	Platelet crit	-0.00047	4.24E-03
DW2	Biomarker	Calcium	Cell count	Haemoglobin concentration	3.13754	4.28E-03
DW2	Biomarker	AST	Cell count	Mean platelet volume	0.01189	4.41E-03
DW2	Biomarker	TG	Cell count	Mean sphered cell volume	-1.39798	4.61E-03
DW2	Biomarker	LDL-C	Cell count	Haemoglobin concentration	0.33793	4.77E-03
DW2	Biomarker	Albumin	Cell count	MCV	-0.40195	5.00E-03
DW2	Biomarker	LDL-C	Cell count	RBC count	0.11429	7.38E-03
DW2	Biomarker	Cystatin C	Cell count	RDW	1.32090	9.75E-03
DW2	Biomarker	Cystatin C	Cell count	Neutrophil percentage	8.59078	1.23E-02
DW2	Biomarker	Calcium	Cell count	Haematocrit percentage	8.27618	1.65E-02
DW2	Biomarker	HDL-C	Cell count	Mean reticulocyte volume	5.14318	2.29E-02
DW2	Biomarker	CRP	Cell count	Lymphocyte percentage	-0.33776	2.36E-02
DW2	Biomarker	IGF 1	Cell count	Haemoglobin concentration	0.04287	2.47E-02
DW2	Biomarker	Apo-B	Cell count	Haematocrit percentage	3.17709	2.53E-02
DW2	Biomarker	GGT	Cell count	Mean reticulocyte volume	0.01506	2.69E-02
DW2	Biomarker	GGT	Cell count	MCV	0.01090	3.26E-02
DW2	Biomarker	AST	Cell count	Monocyte percentage	0.02136	3.34E-02
DW2	Biomarker	TC	Cell count	RBC count	0.08033	3.73E-02
DW2	Biomarker	Urea	Cell count	MCV	-0.81979	3.74E-02
DW2	Biomarker	ALP	Cell count	Monocyte count	0.00080	4.08E-02
D2BF	Cell count	RBC count	Diet	Eat sugar	2.41031	3.57E-02
D2BF	Biomarker	Albumin	Cell count	RBC count	0.03775	3.69E-05
D2BF	Biomarker	Total bilirubin	Cell count	Platelet crit	-0.00370	5.15E-05
D2BF	Biomarker	SHBG	Cell count	Mean reticulocyte volume	0.10338	7.31E-05
D2BF	Biomarker	TG	Cell count	Lymphocyte count	0.18706	2.75E-04
D2BF	Biomarker	GGT	Cell count	Monocyte count	0.00070	3.80E-04
D2BF	Biomarker	SHBG	Cell count	Mean sphered cell volume	0.07030	3.92E-04
D2BF	Biomarker	Total bilirubin	Cell count	Platelet distribution width	0.03156	4.69E-04
D2BF	Biomarker	Total bilirubin	Cell count	Platelet count	-3.94588	7.57E-04
D2BF	Biomarker	Cystatin C	Cell count	Lymphocyte percentage	-10.66387	1.18E-03
D2BF	Biomarker	Albumin	Cell count	Mean sphered cell volume	-0.50827	1.61E-03
D2BF	Biomarker	Albumin	Cell count	Haematocrit percentage	0.25718	2.88E-03
D2BF	Biomarker	IGF 1	Cell count	Mean sphered cell volume	-0.25207	4.44E-03
D2BF	Biomarker	AST	Cell count	Mean corpuscular haemoglobin	0.02900	4.49E-03
D2BF	Biomarker	Albumin	Cell count	RDW	-0.08987	5.62E-03
D2BF	Biomarker	SHBG	Cell count	Platelet crit	-0.00061	6.65E-03

D2BF	Biomarker	Albumin	Cell count	Haemoglobin concentration	0.08370	6.92E-03
D2BF	Biomarker	SHBG	Cell count	Platelet count	-0.69373	1.04E-02
D2BF	Biomarker	GGT	Cell count	Platelet distribution width	0.00154	1.14E-02
D2BF	Biomarker	TG	Cell count	Haemoglobin concentration	0.26713	1.42E-02
D2BF	Biomarker	AST	Cell count	MCV	0.07043	1.90E-02
D2BF	Biomarker	HDL-C	Cell count	Eosinophil count	-0.07608	2.15E-02
D2BF	Biomarker	AST	Cell count	Mean sphered cell volume	0.07607	3.42E-02
D2BF	Biomarker	TG	Cell count	Mean sphered cell volume	-1.45112	3.57E-02
D2BF	Biomarker	Phosphate	Cell count	Lymphocyte count	0.87362	3.65E-02
D2BF	Biomarker	Calcium	Cell count	Platelet crit	0.13524	4.39E-02
D2BF	Biomarker	IGF 1	Cell count	Mean reticulocyte volume	-0.31098	4.71E-02

Appendix Table 2

List of all associations identified across all cohorts. Associations are grouped by cohort and ordered by P value.

Feature	*P value*	*Cohorts*
ALT	4.20E-03	Before-After
	1.11E-06	Within 2-After
	<1.00E-12	Within 2-Control
	<1.00E-12	Before-Control
	<1.00E-12	After-Control
Albumin	2.05E-03	Within 2-After
	<1.00E-12	Within 2-Control
	<1.00E-12	Before-Control
	<1.00E-12	After-Control
ALP	<1.00E-12	Within 2-Control
	<1.00E-12	Before-Control
	<1.00E-12	Within 2-After
	<1.00E-12	Before-After
	<1.00E-12	After-Control
Apo-A	3.70E-04	Before-Control
	2.47E-04	Within 2-After
	<1.00E-12	Within 2-Control
	<1.00E-12	After-Control
Apo-B	4.97E-02	Before-Control
	2.65E-02	Within 2-Control
	1.58E-02	After-Control
AST	4.95E-02	Before-Within 2
	3.17E-08	Before-After
	<1.00E-12	Within 2-Control
	<1.00E-12	Before-Control
	<1.00E-12	Within 2-After
	<1.00E-12	After-Control
AST:ALT	<1.00E-12	After-Control
Basophil count	7.58E-04	Within 2-Control
	2.42E-05	After-Control
Basophil percentage	1.74E-02	Within 2-Control
CRP	2.51E-02	Before-After
	3.80E-04	Within 2-After
	<1.00E-12	Within 2-Control
	<1.00E-12	Before-Control
	<1.00E-12	After-Control
TC	1.38E-02	Before-After
	3.80E-03	Within 2-After
	1.60E-07	Before-Control
	<1.00E-12	Within 2-Control
	<1.00E-12	After-Control
Creatinine (enzymatic) in urine	<1.00E-12	Within 2-Control
	<1.00E-12	Before-Control

Creatinine (enzymatic) in urine	<1.00E-12	After-Control
Cystatin C	4.09E-06	Within 2-After
	<1.00E-12	Within 2-Control
	<1.00E-12	Before-Control
	<1.00E-12	Before-After
	<1.00E-12	After-Control
Eosinophil count	6.04E-03	Within 2-Control
	2.95E-08	After-Control
Eosinophil percentage	2.22E-02	After-Control
GGT	<1.00E-12	Within 2-Control
	<1.00E-12	Before-Control
	<1.00E-12	Within 2-After
	<1.00E-12	Before-After
	<1.00E-12	After-Control
Glucose	1.51E-02	Before-Control
	7.01E-04	Within 2-Control
	<1.00E-12	After-Control
HbA1c	3.71E-04	Before-Within 2
	2.10E-04	Before-After
	1.18E-06	Within 2-Control
	<1.00E-12	Before-Control
	<1.00E-12	After-Control
Haemoglobin concentration	2.26E-02	Within 2-After
HDL-C	5.92E-05	Before-Control
	<1.00E-12	Within 2-Control
	<1.00E-12	After-Control
HLSRC	<1.00E-12	Within 2-Control
	<1.00E-12	Before-Control
	<1.00E-12	After-Control
HLSRP	1.27E-02	Within 2-After
	<1.00E-12	Within 2-Control
	<1.00E-12	Before-Control
	<1.00E-12	After-Control
IGF 1	6.75E-05	Before-After
	9.41E-09	Within 2-After
	<1.00E-12	Within 2-Control
	<1.00E-12	Before-Control
	<1.00E-12	After-Control
Immature reticulocyte fraction	<1.00E-12	Within 2-Control
	<1.00E-12	Before-Control
	<1.00E-12	After-Control
LDL-C	4.53E-02	Before-After
	8.82E-07	Before-Control
	1.73E-07	Within 2-Control

LDL-C	<1.00E-12	After-Control
Lymphocyte count	6.75E-05	After-Control
Mean corpuscular haemoglobin	1.50E-02	Within 2-Control
	<1.00E-12	After-Control
MCV	1.12E-03	Within 2-Control
	<1.00E-12	After-Control
Mean reticulocyte volume	5.12E-03	Before-Control
	2.75E-05	Within 2-Control
	<1.00E-12	After-Control
Mean sphered cell volume	6.51E-03	Within 2-Control
	1.77E-11	After-Control
Microalbumin in urine	3.67E-04	Within 2-After
	<1.00E-12	After-Control
Monocyte count	2.06E-03	Before-Control
	7.50E-08	Within 2-Control
	<1.00E-12	After-Control
Monocyte percentage	1.92E-03	Within 2-Control
	1.00E-06	After-Control
Neutrophil count	5.33E-03	Within 2-Control
	8.78E-05	After-Control
Phosphate	8.01E-03	After-Control
Platelet count	2.96E-05	Within 2-After
	2.52E-05	Before-Control
	3.30E-08	After-Control
	<1.00E-12	Within 2-Control
Platelet crit	4.09E-03	Before-After
	3.28E-07	Before-Control
	2.50E-08	After-Control
	1.95E-08	Within 2-After
	<1.00E-12	Within 2-Control
Platelet distribution width	3.78E-02	Within 2-After
	3.48E-02	Before-Control
	4.65E-07	Within 2-Control
	<1.00E-12	After-Control
Potassium in urine	<1.00E-12	Within 2-Control
	<1.00E-12	Before-Control
	<1.00E-12	After-Control
RBC count	2.82E-02	Before-Control
	2.54E-03	Within 2-Control
RDW	4.66E-06	Before-After
	<1.00E-12	Within 2-Control
	<1.00E-12	Before-Control
	<1.00E-12	Within 2-After
	<1.00E-12	After-Control

Reticulocyte count	1.78E-10	Before-Control
	<1.00E-12	Within 2-Control
	<1.00E-12	After-Control
Reticulocyte percentage	<1.00E-12	Within 2-Control
	<1.00E-12	Before-Control
	<1.00E-12	After-Control
SHBG	4.95E-02	Before-Control
	2.47E-02	Within 2-Control
	2.20E-02	Before-After
	9.09E-03	Within 2-After
Sodium in urine	2.23E-06	Within 2-Control
	1.26E-06	Before-Control
	<1.00E-12	After-Control
Testosterone	3.98E-02	Within 2-After
	6.60E-03	Before-After
	7.39E-05	Before-Within 2
	1.61E-07	Within 2-Control
	<1.00E-12	After-Control
Total bilirubin	4.84E-04	Before-After
	1.50E-05	Before-Control
	<1.00E-12	Within 2-After
	<1.00E-12	Within 2-Control
Total protein	1.59E-02	After-Control
TG	1.10E-04	Before-Control
	2.43E-07	Within 2-Control
	<1.00E-12	After-Control
Urate	1.57E-02	Before-Control
	<1.00E-12	Within 2-Control
	<1.00E-12	After-Control
Urea	1.29E-03	After-Control
WBC count	3.57E-04	Within 2-Control
	<1.00E-12	After-Control

Appendix Table 3

Table of significant (correct P value < 0.05) associations between cohorts, identified through ANOVA and Tukey HSD tests. Associations are grouped by feature and ordered by P value. P values smaller than 1.00E-12 are recorded as <1.00E-12.