University of East Anglia

The inter-individual variability of blueberry anthocyanin metabolism and its impact on cardiovascular risk factors

Laura Annemarie Haag

A thesis submitted for the degree of Doctor of Philosophy

University of East Anglia

Norwich Medical School

December 2021

Abstract

Anthocyanins, a subclass of flavonoids, are plant metabolites found in commonly consumed red-, blue-, and purple-coloured fruits and vegetables. In population-based studies, habitual anthocyanin intakes have been associated with reduced risk for cardiovascular diseases. However, there is contradictory evidence gained from randomised controlled trials as wide inter-individual variability in response to anthocyanin intake has been repeatedly observed. A contributing factor is likely to be variability in absorption and metabolisation of anthocyanins, which suggests that individuals are exposed to different levels of potentially clinically bioactive compounds, and which may underpin classification of individuals as 'responders' and 'non-responders'. In contrast to other flavonoid subclasses (i.e. isoflavones, ellagitannins) with unique, metabotype-defining catabolites (i.e. equol and urolithins, respectively), anthocyanins are metabolised to a set of metabolites common to many flavonoids. To date, this commonality in metabolism end-products has added complexity to the identification of a specific metabotype for anthocyanins.

This research gap is addressed in this thesis by using a combination of factor analysis and univariate methods to identify a group of seven urinary metabolites proposed to describe a responder or 'high metaboliser' of anthocyanin intake: *4-hydroxyhippuric acid, 3-hydroxyhippuric acid, hippuric acid, syringic acid, homovanillic acid, dihydroferulic acid, and 3,5-dihydroxyphenylpropionic acid*. In addition, this thesis tested the relationship between the anthocyanin 'responder' metabolite profile and flow-mediated dilation (FMD) and confirmed a strong association ($\beta=0.79$, $p=0.02$) with each doubling in excretion of the panel metabolites associated with a 0.8% increase in FMD.

The results suggest that individuals with a high metaboliser profile may experience greater vascular benefits from the consumption of anthocyanin-rich blueberries than lower metaboliser profiles. To confirm its usefulness as a screening tool, the identified metabolite panel was applied to an ongoing dietary intervention study for the prospective recruitment of individuals classified as high or low metabolisers following a single dose of blueberries.

Table of Contents

Table of Figures

Table of Tables

List of Abbreviations

AIx	augmentation index
AMPK	adenosine monophosphate-activated protein kinase
AOAC	American Association of Analytical Chemists
AUC	area under the curve
BMI	body mass index
CE	cholesteryl ester
cGMP	cyclic guanosine monophosphate
CI	confidence interval
C_{max}	maximum plasma or serum concentrations
CRP	C reactive protein
CV	coefficient of variation
CVD	cardiovascular disease
DBP	diastolic blood pressure
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
EFA	exploratory factor analysis
ESI	electrospray ionization
eNOS	endothelial nitric oxide synthase
ER	endoplasmatic reticulum
ET-1	endothelin-1
FFA	free fatty acids
FFQ	food frequency questionnaire
∆FMD	change in FMD
FC	fold change
FMD	flow-mediated dilation
FPR	false positive rate
GBD	Global Burden of Disease
GIT	gastrointestinal tract
GLUT2	glucose transporter 2
GTP	guanosine triphosphate
HDL	high-density lipoprotein
HDLC	high-density lipoprotein cholesterol
HPFS	Health Professional Follow-Up Study
HR	hazard ratio
HSE	Health Survey for England
ICAM-1	intercellular adhesion molecule 1
IFN-γ	interferon-γ
IKK-β	inhibitor of nuclear factor kappa B kinase
IL-1β	interleukin-1β
IL-2	interleukin-2
IL-6	interleukin-6
IQR	interquartile range
JNK	c-Jun N-terminal kinase
KMO	Kaiser-Meyer-Olkin measure of sample adequacy
LCMS	liquid-chromatography mass spectrometry
LDL	low-density lipoprotein
LDLC	low-density lipoprotein cholesterol
Log_2 FC	logarithm of fold change to the base 2
LPL	lipoprotein lipase
MAP	minimum average partial test

MetS	metabolic syndrome
MI	myocardial infarction
MMP	matrix metalloprotease
mo	month
MRM	multiple reaction monitoring
NADPH	nicotinamide adenine dinucleotide phosphate
NCD	non-communicable disease
NHS	Nurses' Health Study
NHSII	Nurses' Health Study II
NO	nitric oxide
NOS	nitric oxide synthase
oxLDL	oxidised low-density lipoprotein
QC	quality control
PCA	principal component analysis
PKC	protein kinase C
PWV	pulse-wave velocity
RBP4	retinol binding protein-4
RCT	randomised controlled trial
RNS	reactive nitrogen species
ROC	receiver-operating characteristic
ROS	reactive oxygen species
SBP	systolic blood pressure
SCFA	short-chain fatty acids
SD	standard deviation
sdLDL	small dense low-density lipoprotein
SE	standard error
SEM	standard error of the mean
SES	socioeconomic status
sGC	soluble guanylate cyclase
SGLT1	sodium-dependent glucose transporter
SMC	smooth muscle cells
sMRM	scheduled multiple reaction monitoring
SOCS	suppressors of cytokine signalling
sICAM-1	soluble intercellular adhesion molecule 1
sVCAM-1	soluble vascular cell adhesion molecule 1
T2DM	type 2 diabetes mellitus
TG	triglycerides
TLR4	toll-like receptor 4
TMA	trimethylamine
TNF-α	tumour necrosis factor α
TPR	true positive rate
VCAM-1	vascular cell adhesion molecule 1
VLDL	very-low-density lipoproteins
VSMC	vascular smooth muscle cells
WHO	World Health Organisation

Acknowledgements

First and foremost, I would like to thank my supervisors Dr Peter Curtis, Prof Aedin Cassidy, and Dr Amy Jennings for their support and guidance. I would like to especially acknowledge Dr Peter Curtis for his encouragement and support over the years and particularly in the final stages of my thesis.

I extend my gratitude and thanks to the University of East Anglia for the FMH studentship and numerous opportunities to grow personally and professionally.

I would like to thank Dr Vera van der Velpen and Dr Colin Kay for their training in mass spectrometry. I would also like to thank Dr Jon Tang for letting me use their LCMS when ours encountered mechanical difficulties.

I would like to thank Dr Peter Curtis, Caoimhe Flynn, Dr Nicky Bondonno, Faith Ndungwani, Chris Benwell and the research nurses at the Clinical Research Facility for the excellent team work on the AMP study. Sincere thanks to the people of Norfolk who have generously given their time to participate in our study.

Thank you to all my family and all my friends, old and new, who have become like family and have made my time in Norwich an adventure to remember.

Thank you to my parents. I owe my success to your endless support and continuous encouragement.

Finally, a special thank you to Derrick for your unwavering patience, understanding, and support throughout my entire PhD journey.

1 Introduction

Nutrition has long been considered a crucial factor in health and disease, as illustrated by the quote ‘Let thy food be thy medicine, and medicine be thy food’ often credited to Hippocrates (400 BC).[1]

Today this statement still holds to be true, however since the days of Hippocrates our understanding of nutrition and its link to health as well as our diets have changed substantially.

One of the most striking characteristics of our modern diets and lifestyle is the overconsumption of unhealthy foods, such as highly processed and packaged foods, high in free sugars and salt, and the underconsumption of healthy foods including fruits, vegetables, legumes, whole grains, nuts, and seeds. In 2017, the Global Burden of Disease Study (GBD)[2] estimated that globally, the consumption of fruits was less than 100 g/day, or about 40% of what was considered an optimal level of intake of 250 g/d. The optimal intake level was defined by GBD as the amount of that minimised the risk from all causes of death and was based on results from published literature. Further examples of suboptimal consumption were whole grains (30 g/d, 24% of the optimal level of 125 g/d) and nuts and seeds (3 g/d, 12% of the optimal level of 21 g/day). In contrast, the amount of consumed sugar-sweetened beverages (49 g/d), processed meats (4 g/d), salt (6 g/d), and red meats (27 g/d) greatly exceeded the optimal levels globally by over 1500%, 90%, 86%, and 18%.

The GBD study further reported that, globally, the diet lacked in beneficial nutrients (with exception to select regions), such as fibre (11 g/d, 46% of optimal intake), omega-3-fatty acids (0.1 g/d, 4% of optimal intake), calcium (0.4 g/d, 32% of optimal intake), and polyunsaturated fatty acids (4% of total energy intake, 36% of optimal intake).

Dietary patterns characterised by high intakes of red and processed meats, refined grains, sugar, saturated fats and salt, and reduced intakes of fruit, vegetables, whole grains, fibre and other beneficial nutrients have been recognised by the World Health Organisation (WHO) as a modifiable dietary risk for the development of non-communicable diseases (NCD)[2,3]. In 2019, poor diet quality was one of the leading behavioural risk factors for NCDs, only second to smoking and followed by alcohol use and low physical activity.[4,5] It was estimated that 74% of all deaths globally (42 million people annually) are related to NCDs, of which over 80% of these deaths were caused by cardiovascular diseases (CVD) (18.6 million globally), cancers (10 million globally), chronic respiratory diseases (4 million globally), and type 2 diabetes (T2DM) (1.5 million globally). Every one in seven NCD deaths was attributable to dietary risks (about 8 million deaths world-wide, 14% of total global deaths), of which the majority were caused by CVD (6.9 million, about 86% of diet-related global deaths) followed by T2DM (0.4 million, about

5% of diet-related global deaths). In the UK, dietary risks were responsible for every eighth death (78 500, 12.6% of total deaths in UK). Over 80% of these were due to CVDs.

Given the scope and magnitude of diet-related diseases, there has been a growing interest in functional foods as therapeutics,[6] alongside an increasing awareness that changes in lifestyle and diet may alleviate or manage NCDs.[1] Implementing nutritional guidelines to improve diet quality and tackle excess energy intake is one key strategy in controlling and reducing the burden of these diseases.

This introduction will first give an overview of cardiovascular disease and cardiometabolic risk factors. It will then move on to the significance of a diet rich in fruit and polyphenols as potential mediators of health. More specifically, anthocyanins will be introduced. Finally, the metabolism of flavonoids and the variability in the bioavailability of flavonoids in individuals will be discussed, and at the end a summary of the aims of this thesis will be presented.

1.1 Cardiovascular diseases and cardiometabolic risk factors

Cardiovascular diseases are disorders of the heart and blood vessels at the centre of which often lies atherosclerosis, a build-up of fatty deposits in the artery walls called plaque which is associated with hardening of the arterial walls and inflammation. The clinical symptoms are determined by the location of the atherosclerotic site as well as the severity of the plaque. The lipid depositions lead to a narrowing of the artery which impedes the blood flow to organs and limbs and eventually may culminate in the occlusion of the artery stopping blood flow altogether. Manifestations of atherosclerosis include coronary artery disease, angina, heart attacks, ischemia, peripheral vascular disease, renal artery stenosis and stroke. Inflammation can also damage and weaken the vessel wall and may cause it to dilate leading to an aneurysm.[7]

Pathophysiology of atherosclerosis

Key players in the formation of early atherosclerotic lesions are endothelial dysfunction, dyslipidaemia, and inflammation.

The arterial wall consists of three layers. The tunica intima is the innermost layer consisting of a thin membrane lining the inside of blood vessels called the endothelium, and a subendothelial layer of connective tissue. The endothelium has important functions including barrier forming properties, preventing platelet aggregation, maintaining low oxidative stress, inflammation signalling, and regulation of the constriction (vasoconstriction) and enlargement (vasodilation) of the blood vessel, subsequently controlling blood pressure. The tunica media is a layer of smooth muscle cells and elastic tissue which regulates the internal diameter of the vessel. The

tunica adventitia is a layer of connective tissue anchoring the blood vessel into the surrounding tissue.

Early stages of atherosclerosis appear as a fatty streak in the vascular intima.[7] Damage to the endothelium can lead to an activation of endothelial cells initiating pro-atherogenic processes, including increased endothelial permeability, platelet aggregation, production of reactive oxygen species (ROS), and expression of pro-inflammatory molecules.[8] The overproduction of highly reactive ROS, such as the free radical superoxide anion radical, and a lack of enzymatic and non-enzymatic antioxidants causes oxidative stress and can incur damage to cellular lipids, proteins, and DNA bases by changing their structure and inhibiting their function. Increased endothelial permeability along with increased levels of low-density lipoproteins (LDL) allows for LDL particles to enter the subendothelial space where they become oxidised to oxLDL by ROS. oxLDL provokes the expression of adhesion molecules, such as ICAM-1 and VCAM-1, in endothelial cells which attract inflammatory cells to the site. Monocytes adhere to the endothelium and migrate into the intima where they become activated and turn into macrophages. The oxLDL binds to the macrophage scavenger receptors and becomes internalised. However, this modified form of LDL cannot be digested fully, which results in an accumulation of lipids in the macrophages and the formation of foam cells. The abnormal lipid accumulation is perceived as a danger signal by the cells and in an attempt to remove the lipids, foam cells release chemokines and cytokines to recruit additional inflammatory cells, including T-helper cells and further monocytes. Pro-inflammatory cytokines produced by the T-helper cells (such as interferon-γ (IFN-γ) and interleukin-2 (IL-2)) stimulate the differentiation of monocytes to macrophages: These in turn release more ROS, cytokines (such as TNF-α, IL-1 and IL-6), and matrix metalloproteases (MMPs) and become foam cells through uptake of oxLDL. The secreted cytokines further stimulate proliferation of smooth muscle cells (SMC) and fibroblasts and the activation of endothelial cells. MMPs degrade the extracellular matrix and allow the migration of SMC into the intima where they synthesise a protective fibrous cap over the plaque. In advanced plaques, recruited fibroblasts produce large amounts collagen which leads to fibrosis and scarring. The abnormal lipid loading in foam cells can cause the cells to die and as the plaque grows bigger, dead foam cells in the centre form a necrotic core containing lipids and cell debris. The growth of the atherosclerotic plaque is driven by a positive feedback loop of lipid accumulation, activation of endothelial cells, inflammation and SMC proliferation until it may rupture and break through the endothelium. This disruption and spilling out of the necrotic core rapidly causes a thrombus to form which may partially or entirely occlude the vessel lumen and can have catastrophic consequences, such as a stroke or a fatal heart attack.

Cardiometabolic risk factors and the metabolic syndrome

Cardiometabolic risk summarises a set of clinical conditions which increase the risk for cardiovascular disease and T2DM. These include being overweight/obese or abdominal obesity, hyperglycaemia, hyperinsulinemia, hypertension, and dyslipidemia.[9] The increasing prevalence of overweight and obesity gives rise to a clustering of these co-morbidities, termed metabolic syndrome (MetS).[10] Although the exact pathophysiology of MetS is unknown, insulin resistance and excessive abundance of free fatty acids (FFA) appear to be centre of the development.

The metabolic abnormalities in metabolic syndrome are briefly described in the following.

Obesity and central adiposity

In times when energy is abundant, fat cells (adipocytes) grow to store the surplus energy as fat. More specifically, they store fat as triglycerides (TG) which consist of three fatty acids and one glycerol molecule. The cells are located in adipose tissue mainly beneath the skin (subcutaneous fat) and intra-abdominally surrounding internal organs (visceral fat). The prolonged imbalance of energy intake and energy expenditure eventually leads to weight gain.

Overweight and obesity describe a state of excessive accumulation of fat and has been associated with chronic low-grade inflammation[11] and metabolic changes including dyslipidaemia, hyperglycaemia, and hyperinsulinemia. It is commonly measured using the body mass index (BMI), defined as the weight of an individual divided by the square their height. Overweight is classified as a BMI ≥ 25 kg/m^2 and obese as a BMI ≥ 30 kg/m^2. By this definition, according to the Health Survey for England (HSE) of 2019[12], 64% of adults in England are overweight and obese, while a striking 30% of children between the ages of 2 and 15 are overweight or obese. The prevalence of overweight and obesity in adults has increased by 11% since 1993, and the Global Burden of Disease Study 2019 has reported that over the previous ten years, the prevalence of a high BMI has been steadily increasing with an annual increase of 1.86%.[4]

Surprisingly, not all overweight and obese subjects are affected by metabolic changes such as insulin resistance and dyslipidaemia and conversely, non-obese individuals may suffer from metabolic disorders. The amount of abdominal fat has been suggested as the potential link to this observation. Despite having the same amount of total body fat, those individuals with obesity-related metabolic abnormalities were characterized by an excess amount of visceral fat, whereas metabolically healthy obese had more subcutaneous fat[13]. Visceral adiposity has been associated with metabolic complications such as insulin resistance, dyslipidaemia and inflammation and many studies investigating the relationship between fat

distribution and disease risk seem to confirm that the distribution of fat is a more significant risk factor for diseases such as CVD and mortality than excess weight and fat per se[14–20].

Insulin resistance

Overweight and obesity increase the risk for insulin resistance, which means that cells are less responsive to normal or elevated levels of insulin. Insulin is secreted by pancreatic β-cells when blood sugar levels are high and stimulates glucose and FFA uptake in fat and muscle cells, increases fatty acid synthesis and decreases lipolysis, and suppresses gluconeogenesis in the liver. However, when cells are insulin resistant, plasma levels of glucose and FFA are increased. In the insulin resistant liver, reduced suppression of gluconeogenesis augments the already high glucose levels and contribute to hyperglycaemia. This in turn results in more insulin being synthesized by the pancreatic β-cells causing hyperinsulinemia. The increased production of insulin eventually causes endoplasmic reticulum (ER) stress and failure to alleviate this stress leads to death of the β-cells and progression to type II diabetes.

Multiple factors promote insulin resistance, including elevated levels FFA, dysregulation of adipokine secretion from adipose tissue, and inflammation (**Figure 1-1**).

In most obese individuals, FFA levels are increased.[21] This is partly due to increased lipolysis of excess adipose tissue. Once increased, this is further enhanced by the inhibition of insulin action, which under normal conditions would stimulate the removal of FFA from the blood into cells. The mechanism by which FFA promote insulin resistance is thought to involve the formation of lipid intermediates, which activate several kinases that interfere in insulin signalling including protein kinase C (PKC), inhibitor of nuclear factor kappa B kinase (IKK-β) and c-Jun NH2-terminal kinase (JNK).[22] In addition, FFA also induce oxidative stress, which was shown to play a role in insulin resistance.[22] Conversely, in obese subjects, the reduction of plasma FFA using anti-lipolytic drugs by 60-70% in comparison to placebo was demonstrated to restore insulin sensitivity and increase insulin-stimulated glucose uptake by more than two-fold.[23]

Next to storing fat, adipose tissue also functions as a secretory organ releasing cell-signalling proteins called adipokines. The dysregulation of some of the adipokines, such as adiponectin, leptin, and retinol binding protein-4 (RBP4), is associated with insulin resistance. Most of these mechanisms are mediated through the AMP-activated protein kinase (AMPK) which regulates energy homeostasis by promoting fatty acid oxidation, glucose uptake, and glycolysis, while inhibiting fatty acid synthesis, gluconeogenesis and glycogen synthesis. Adiponectin is an activator of AMPK and has an insulin-sensitising effect[24]. A lack of adiponectin, as the case in obesity, thereby supports insulin resistance. Leptin has also been

shown to stimulate fatty acid oxidation through AMPK signalling[25] and additionally acts as an appetite suppressant in the hypothalamus.

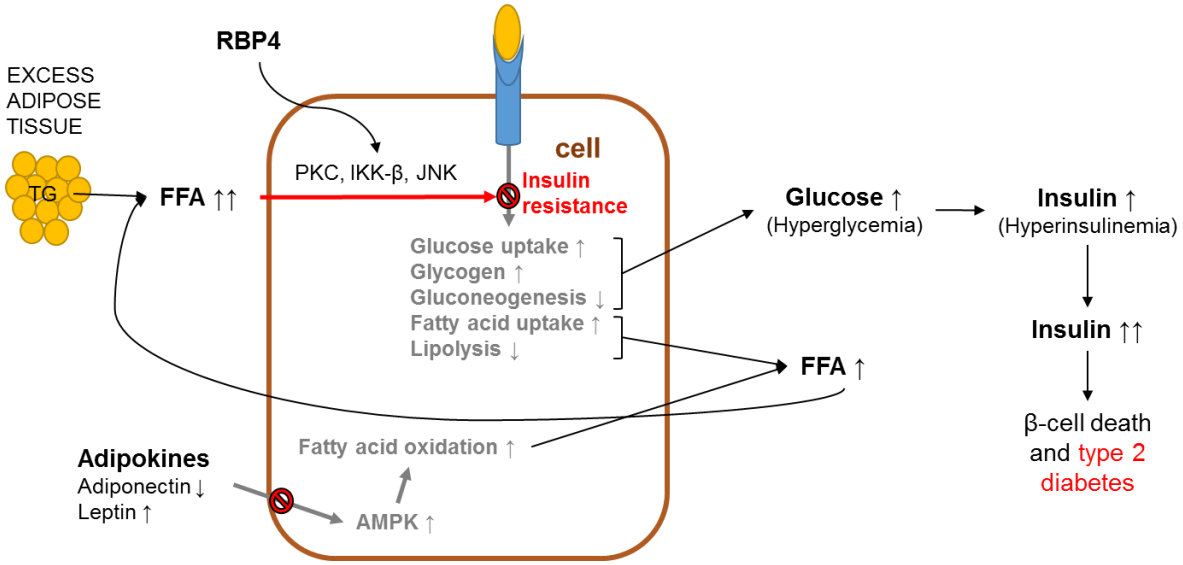


Figure 1-1. Mechanisms of insulin resistance.

Elevated levels FFA, dysregulation of adipokine secretion from adipose tissue, and inflammation can promote insulin resistance, which means that cells are less responsive to normal or elevated levels of insulin. This results in even more increased plasma levels of FFA and glucose. Reduced suppression of gluconeogenesis in the insulin-resistant liver further augments the hyperglycaemia. This in turn results in more insulin being synthesized by pancreatic β-cells causing hyperinsulinemia. The increased production of insulin eventually causes endoplasmic reticulum (ER) stress and failure to alleviate this stress leads to death of the β-cells and progression to type II diabetes.
Greyed out writing denotes original function, which is inhibited under insulin resistance. TG, triglycerides; FFA, free fatty acids; RBP4, retinol-binding protein-4; AMPK, AMP-activated protein kinase; PKC, protein kinase C; IKK-β, inhibitor of nuclear factor kappa B kinase; JNK, c-Jun NH2-terminal kinase.

However, in contrast to adiponectin, leptin levels are increased in obesity and the failure to respond to leptin signalling suggests leptin resistance. It has been demonstrated that the upregulation of suppressors of cytokine signalling (SOCS) plays a role in inhibiting the leptin activation of AMPK.[25] Increased plasma concentrations of RBP4 also induce insulin resistance, however the mechanism by which this is mediated is not fully understood. Recent evidence suggests that RBP4 indirectly promotes insulin resistance by promoting the inflammatory state in adipose tissue through activation of JNK and toll-like-receptor 4 (TLR4) pathways[26,27] (see below).

Inflammation

Overweight and obesity is associated with a chronic low-grade proinflammatory state of the adipose tissue as well as other tissues like the liver, pancreas, brain and skeletal muscle.[11] This state involves increased infiltration of adipose tissue by macrophages and the increased production of inflammatory markers in adipose tissue as well as in the liver (**Figure 1-2)**.

The inflammatory cytokines, which include tumour necrosis factor-α (TNF-α), interleukin-6 (IL-6), monocyte chemotactic protein-1 (MCP-1) and interleukin-1β (IL-1β), interfere in insulin signalling like FFA through intracellular activation of JNK and IKK kinases. TNF-α further stimulates lipolysis, contributing to increased levels of FFA. In addition to the pathways named above, FFA can also activate TLR-4 pathways in adipocytes and macrophages, thereby contributing to the inflammatory state by stimulating the expression of inflammatory cytokines[28]. MCP-1 supports the infiltration of tissue by macrophages, which not only augment the secretion of inflammatory cytokines, but also can induce tissue injury through phagocytosis and degradation of the extracellular matrix by secreting matrix metalloproteinases. In contrast, IL-6 is a pro- and anti-inflammatory cytokine. The opposing roles of IL-6 depend on IL-6 binding to the membrane-bound form of the IL-6 receptor (anti-inflammatory action) or the soluble form, which interacts as a complex with the signalling receptor gp130 (pro-inflammatory action).[29] It is certain that increased plasma concentrations are associated with obesity and insulin resistance. However, it is unclear if IL-6 has a disease promoting or rather a protective role in obesity. An animal study by Mauer et al showed that IL-6 led to increased insulin sensitivity and limits inflammation[30], while acute injection of IL-6 in mice led to insulin resistance in a different study[31].

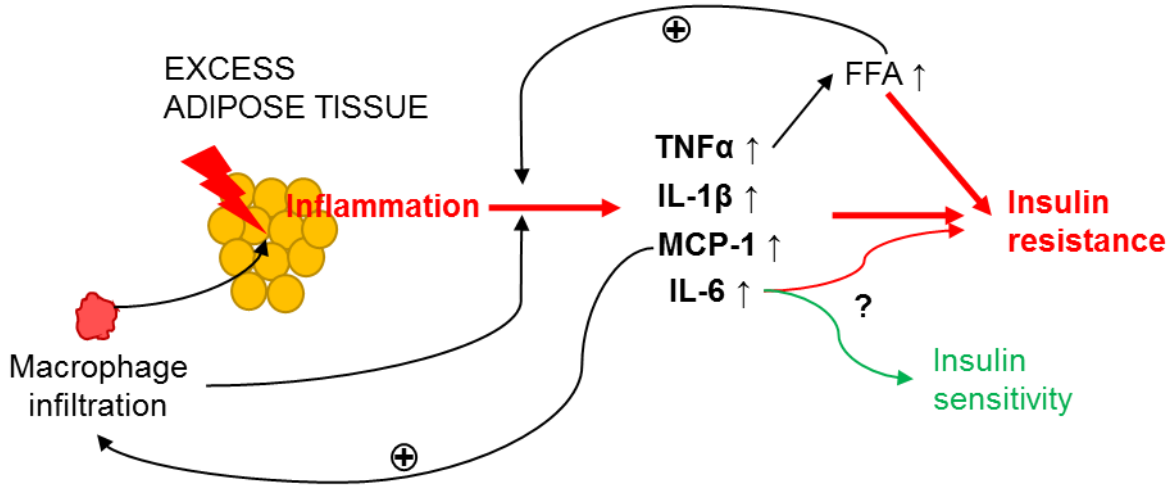


Figure 1-2. Inflammation of adipose tissue in obesity.

Infiltration of macrophages leads to the mild inflammation of adipose tissue and increased secretion of cytokines which promote insulin resistance. TNF-α, tumour necrosis factor α; IL-1β, interleukin-1β; MCP-1, monocyte chemotactic protein-1; IL-6, interleukin-6.

Importantly, inflammation of the hypothalamus appears to play a significant role in leptin and insulin resistance and occurs before weight gain in contrast to inflammation in peripheral tissues as described above.[32] Excess nutrients and high-fat diets, especially those which are high in saturated fats (particularly palmitic acid), cause inflammation in the hypothalamus through activation of JNK and IKK. This leads to the inhibition of leptin and insulin signalling

and thus suppresses the satiety evoking effect of those hormones. It has further been suggested that sustained high-fat feeding may lead to neuron injury though the activation of TLR-4, ceramide biosynthesis and ER stress pathway. Reactive proliferation of neighbouring glial cells may impact neuronal function and eventually lead to the death of neurons responsible for satiety signalling[33].

Liver inflammation is of further significance for the inflammatory state in obesity[11] (Error! Reference source not found.). The liver is a major site of metabolism, regulating gluconeogenesis and glycogen storage, lipogenesis and cholesterol synthesis and secretion. Inflammation in the liver increases the pool of inflammatory cytokines which causes inflammation-induced insulin resistance in the liver by the same mechanisms described above. This results in increased glucose output and contributes to hyperglycaemia. TNF-α and IL-6 have also been shown to activate hepatic lipogenesis which leads to the accumulation of hepatic TG and subsequently increased output of very-low-density lipoproteins (VLDL) and plasma TG levels[34]. Furthermore, acute-phase proteins produced by the liver such as C-reactive protein (CRP), plasminogen activator inhibitor-1, serum amyloid A, and IL-6, as a response to the inflammation are increased in obese individuals[11] which could influence peripheral tissues, e.g. by inducing insulin resistance. Finally, the overabundance of FFA and increased lipogenesis promote the development of a fatty liver, which may culminate in fatty liver disease (steatohepatitis).

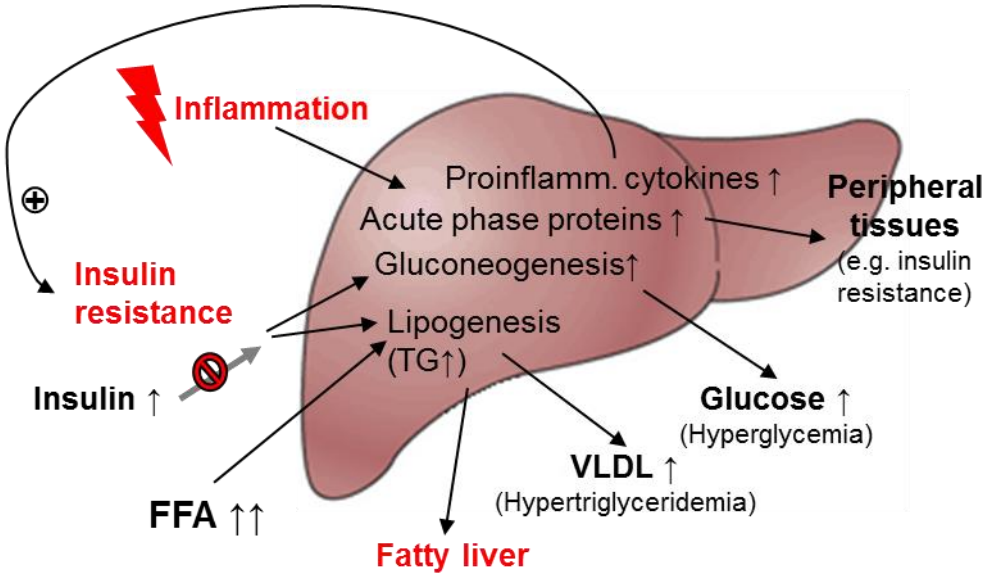


Figure 1-3. Obesity-related inflammation of the liver.

Increased availability of FFA, inflammation and insulin resistance lead to the accumulation of TG in the liver, which promotes hypertriglyceridemia, hyperglycaemia, metabolic disorders in peripheral tissues, and a fatty liver. FFA, free fatty acids; VLDL, very low-density lipoprotein; TG, triglycerides.

Systemic chronic inflammation can amplify the inflammation in atherosclerosis, and non-specific inflammatory markers like CRP are used as a measure in the assessment of atherosclerotic risk.[7]

Dyslipidaemia

A further common change in obese patients are abnormal lipid levels. The key attributes of dyslipidaemia are increased plasma levels of TG (hypertriglyceridemia), increased small density LDL and low HDL-cholesterol (HDL-C). Error! Reference source not found. illustrates the adaptation of the lipid metabolism in obesity.

Dyslipidaemia is likely the result of an interplay between excess nutrients, insulin resistance, and inflammation. Increased release of FFA from adipose tissue due to chronic low-grade inflammation, reduced clearance of FFA and *de novo* fatty acid synthesis in the liver due to insulin resistance may all contribute to the increased availability of FFA. This leads to the overproduction of TG-rich VLDL and subsequently hypertriglyceridemia. In addition, the postprandial clearance of TG from dietary fats packaged into particles called chylomicrons is impaired, further augmenting TG levels in circulation. Hypertriglyceridemia also stimulates the exchange of TG for cholesteryl esters between HDL and LDL particles causing a decrease in HDLC and an increase in small dense LDL particles, which have minimal amounts of TG left. Small dense LDL (sdLDL) are a subtype of LDL and sdLDL-cholesterol has been found to be a major contributor of the total LDLC associated with risk of coronary heart disease[35]. In contrast, HDLC has a protective effect on vascular function, therefore a decrease in HDLC is associated with greater CVD risk [36].

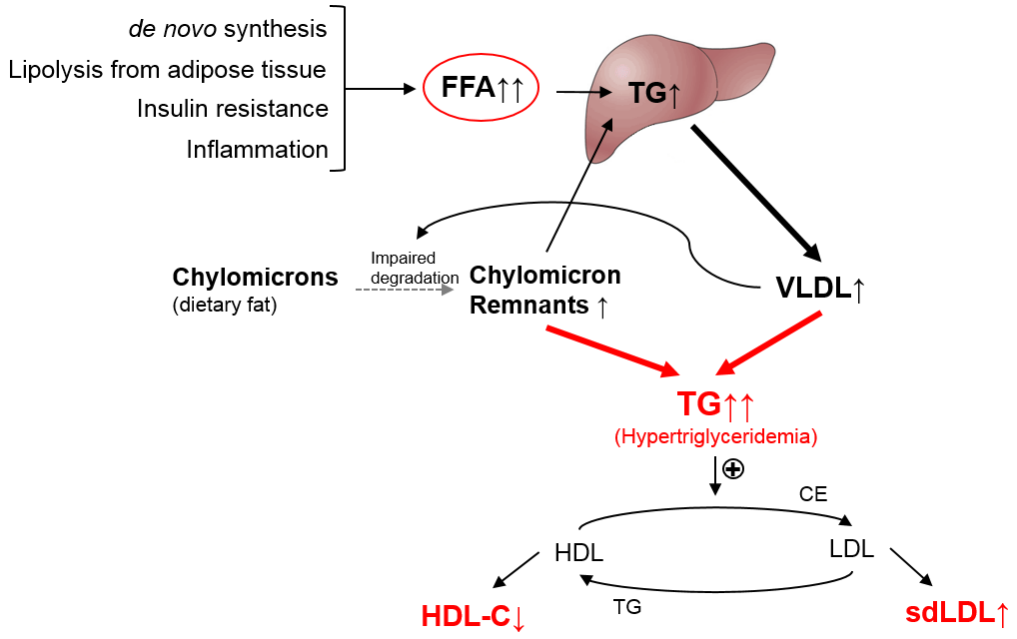


Figure 1-4. Metabolic adaptations of the lipid metabolism in obesity.

Dyslipidaemia in obesity is characterized through increase of FFA in circulation, hepatic overproduction of VLDL, increased blood levels of TG and LDL-C, the formation of sdLDL, and decreased concentration of HDL-C (key aspects highlighted in red). TG, triglycerides; CE, cholesteryl ester; FFA, free fatty acid; VLDL, very low-density lipoprotein; LDL, low density lipoprotein; sdLDL: small dense low-density lipoprotein; HDL, high density lipoprotein; HDL-C, HDL-cholesterol.

Endothelial dysfunction

The endothelium is a thin membrane lining the inside of blood vessels. The functions of a healthy endothelium include providing a barrier between blood and tissues, providing an anti-coagulant surface, and regulation of the vascular tone.

Endothelial cells respond to changes in blood flow, circulating substances and inflammatory mediators to produce different molecules to dilate (nitric oxide, prostacyclin) or constrict (endothelin-1) the vessel.[37] A major regulator of blood flow is nitric oxide (NO), a potent vasodilator as well as inhibitor of platelet adhesion and aggregation, which is synthesised by the endothelial nitric oxide synthase (eNOS). Its counterpart, the very potent vasoconstrictor endothelin-1 (ET-1), is also produced by endothelial cells. NO inhibits the release of ET-1 and the interaction of NO and ET-1 plays an important role in the regulation of vascular tone.[38]

Various triggers can stimulate increased expression of eNOS and NO-release. One of them is shear stress generated by a high blood velocity. The endothelial cell membranes contain ion channels that respond to shear stress and activate the eNOS enzyme, which subsequently generates NO that diffuses to neighbouring smooth muscle cells. Its downstream effects are

mediated by cGMP within the vascular smooth muscle cells which include smooth muscle relaxation and thereby blood vessel dilation.[39] As long as sufficient shear stress persists, the endothelium will release a proportional amount of NO to ensure constant vasodilation, thereby preventing injury and damage from the shear forces.[40] This process is called 'endothelial-dependent vasodilation'.

Endothelial-dependent vasodilation can be measured by ultrasound imaging. This method of measuring is called flow-mediated dilation (FMD) and is the method we used in our study. Shear stress in the brachial artery is artificially induced by inflating a pneumatic cuff for 5 minutes followed by a rapid deflation. In response to the sudden increase in flow, the brachial artery changes its diameter through vasodilation, which can be measured by ultrasound imaging.[41]

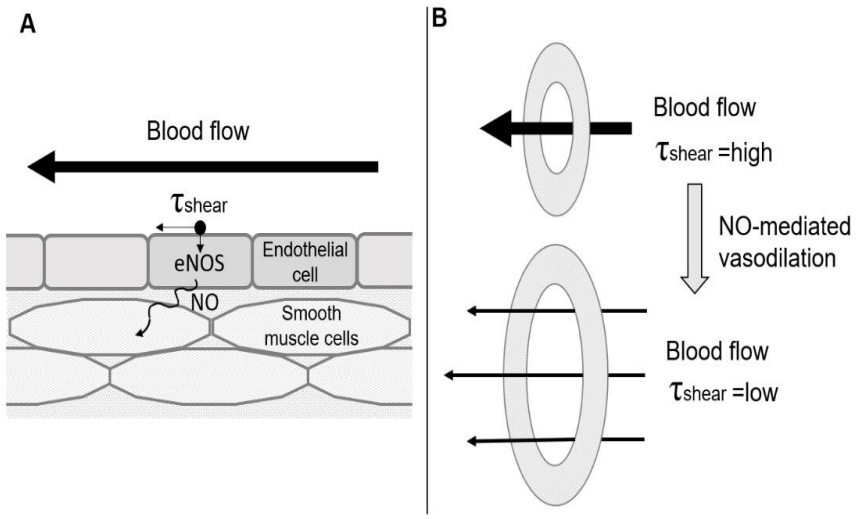


Figure 1-5. Endothelial-dependent vasodilation.

(A) Shear stress is a frictional force generated by high blood velocity. It stimulates ion channels in the endothelial cells. eNOS is activated and generates NO that diffuses into the smooth muscle cells and causes relaxation, resulting in vasodilatation. τ_{shear} = shear stress.[294]

(B) Functional endothelial cells trigger vasodilatation to reduce potentially damaging forces generated by high blood flow velocities. The vasodilation can be measured by the 'flow-mediated dilation'.

Illustration is based on the description from Korkmaz (2009).[41]

A dysfunctional endothelium is one of the first landmarks in atherogenesis and by itself often regarded as a predictor of cardiovascular disease.[42] Damage to the endothelium is characterised by a dysregulation of vascular homeostasis and a reduced bioavailability of NO. The resulting imbalance between NO and ET-1 favours vasoconstriction and thus contributes to hypertension through impaired endothelium-dependent vasodilation and increased vasoconstriction. In addition, endothelial dysfunction is associated with increased synthesis of pro-inflammatory cytokines and adhesion molecules, increased oxidative stress, and increased platelet aggregation.[43]

In the context of the metabolic syndrome, contributors to endothelial dysfunction include insulin resistance and oxidative stress. In a healthy endothelium, insulin stimulates both ET-1 and NO production. However, in an insulin resistant state and compensatory hyperinsulinemia, stimulation of NO release is impaired, while the production of ET-1 is not.[38] Increased oxidative stress can reduce NO bioavailability. NO is a highly reactive radical which rapidly reacts with ROS to form peroxynitrite, thereby reducing the NO-concentration before it reaches the smooth muscle cells to induce vasodilation. Under normal physiological conditions, ROS are balanced by a system of antioxidants,[44] but in an inflammatory state this balance is disrupted, leading to increased oxidative stress. Chronic inflammation[45], hyperglycemia[40] and dyslipidemia[46] can all enhance oxidative stress.

1.2 The case for fruits

Globally, in 2019, a diet low in fruits was the fifth highest behavioural risk factor for CVD and T2DM, surpassed by smoking, and diets high in sodium, low in whole grains, and low in legumes.[5] The total deaths among CVD and T2DM attributable to a diet low in fruits amounted to 1.6% or 0.92 million. The role of fruits in disease prevention is highlighted by a prospective epidemiological study involving three large US cohorts, the Nurses' Health Study (NHS), the Health Professional Follow-Up Study (HPFS), and the Nurses' Health Study II (NHSII).

In this study by Muraki et al[47], a total of 187,382 men and women of white ethnicity and in health professions were involved. Fruit consumption was recorded every four years between 1984 and 2008. The chosen methodology was a self-administered, 118-item, semi-quantitative food frequency questionnaire (FFQ), which consistently asked how often a standard portion size of ten individual fruits (see **Figure 1-6**) and fruit juices (including apple, orange, grapefruit, and other juices) was consumed. FFQs are tools used to estimate habitual dietary intake of a defined number of food items over a period of time. In comparison to 24h dietary recalls or food records, which aim to capture the exact dietary intake on a certain number of days, they are relatively simple and inexpensive to evaluate. However, FFQs are subject to a systematic measurement error — participants may under-report true food intake or change in diet during the period of the survey, which may lead to over- or underestimation of true intake.[48] Therefore, testing the validity and reproducibility is required to ensure accurate intake data. The validity is the ability to record true intake and can be tested by comparing the results with intake data derived from a reference and presumably less-biased method, such as 24h dietary recalls. The reproducibility is the ability to produce the same result when the FFQ is repeated. The FFQ used in the three cohorts was tested for validity and reproducibility for each cohort[49–51] and correlation coefficients specific for individual food items was applied to reduce bias in the intake estimates.

Another source of systematic error is confounding. Confounding is caused by a factor (i.e., a confounder) that is associated with the exposure being investigated (in this case fruit intake) and that is also a risk factor for the health outcome (in this case T2DM) and could thereby distort the relationship of interest between the exposure and the outcome.[52] To reduce the effect of confounders, they are commonly adjusted for by using various techniques, such as stratification for confounding variables or including confounders in multivariate analysis. In the study by Muraki et al, a number of confounders were accounted for. The authors stratified their model by age and calendar year, and adjusted for BMI, ethnicity, physical activity, smoking, multivitamin use, family history of diabetes, menopausal status, oral contraceptive use, total energy intake, fruit juice consumption, and the modified alternate healthy eating index score. Although a range of confounders were considered, there is still a potential of residual or unmeasured confounding. For example, a possible confounder that was not adjusted for is socioeconomic status (SES). It is plausible that SES may influence diabetes and fruit intake through access to health-care services, information, or healthy food. A meta-analysis investigating the association between T2DM incidence and SES found that low levels of educational level, occupation, and income were associated with an increased risk of T2DM by 30 – 40%.[53] In the study by Muraki et al, the statistical model was not adjusted for SES likely because the study population itself consisted entirely of health professionals of similar SES, mitigating SES as a confounding factor.

To measure the effect of specific fruit on health, the authors determined the hazard ratio of T2DM (the health event) for consuming different fruits (the exposure). A hazard ratio (HR) is a time-to-event analysis and is expressed as a ratio of probabilities between two groups for an individual to have an event at any given time of follow-up. If the HR = 1, the event rate in both groups is the same, i.e. fruit intake did not have an impact on the development of T2DM. If the HR > 1, this means that at any given time it is more likely for people in the exposed group to experience an event in comparison to the reference group and vice versa if HR < 1, i.e. fruit intake increases or decreases the probability of T2DM. To calculate the hazard ratio, the authors assessed each participant's person years from the date of the baseline FFQ to the date of one of the following occurrences: T2DM diagnosis, death, date of the last return of a valid questionnaire, or end of study. Participants were divided by the amount of fruit intake into five groups ranging from <1 serving/month to >5 servings/week in increments of 3 servings. Finally, they pooled the results from all three cohorts.

During the examined time period, 6.5% participants developed diabetes (12,198 of 187,382). The study found that greater intakes of specific fruits, including blueberries, grapes, and prunes were associated with lower chance of developing type II diabetes than some other fruits (cantaloupe, strawberries or oranges) and fruit juice (see Figure 1-6). Specifically, this meant that with every three servings of fruit per week, the probability to be diagnosed with T2DM at

any given time during the period of follow-up was reduced. The strongest significant association was reported for blueberries, with a HR (95% confidence interval (CI)) of 0.74 (0.66 – 0.83), i.e. the probability of T2DM was reduced by 26%. In contrast, those who consumed three servings of cantaloupe or fruit juice per week increased their risk to develop T2DM by 10% and 8%, respectively. Interestingly, total fruit intake only had a small impact on T2DM (HR 0.98 (0.96 – 0.99)).

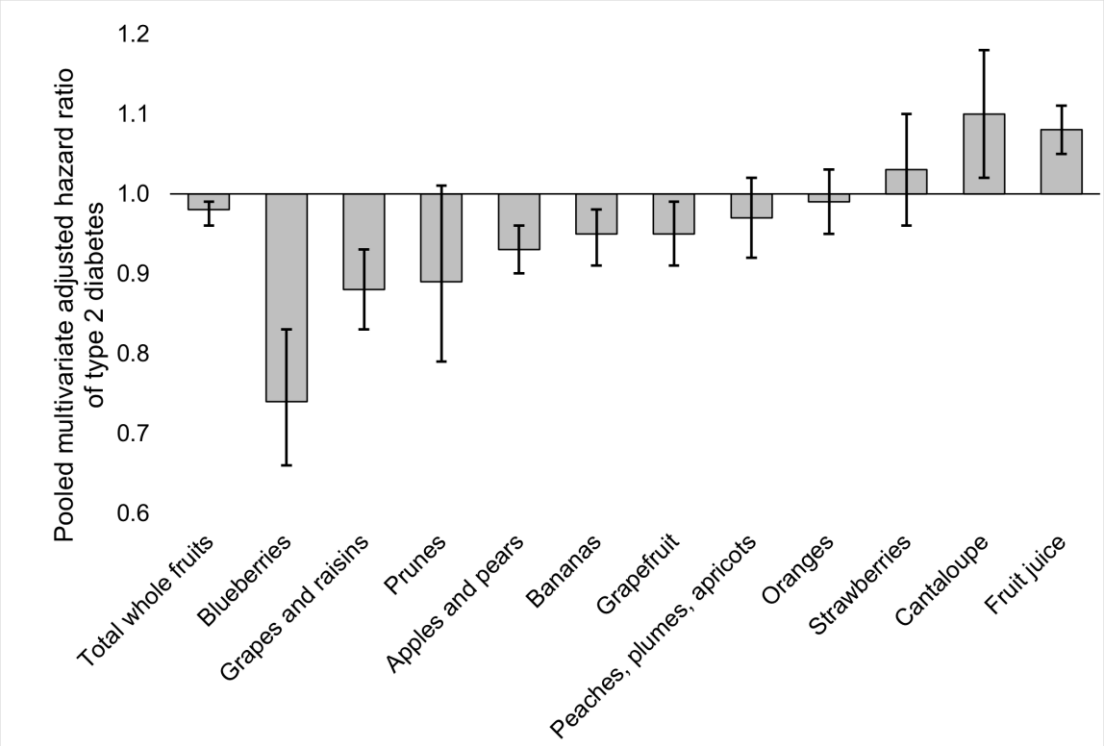


Figure 1-6. Intake of three servings/week of most fruits associated with lower risk of developing T2DM.

Pooled multivariate adjusted hazard ratio of T2DM for every three servings/week of total whole fruit, individual fruits, or fruit juice. Error bars represent 95% confidence intervals (CI). The estimates were pooled from three cohorts totalling 187,382 participants and adjusted for age, ethnicity, body mass index, smoking status, multivitamin use, physical activity, family history of diabetes, menopausal status or postmenopausal hormone use, oral contraceptive use, total energy intake, fruit juice consumption, alternate healthy eating index score. Individual fruit consumption was mutually adjusted. Figure created from data presented in Tables 3 and 4 in Muraki et al.[47]

In total, the evidence suggested that certain fruits were more beneficial in lowering the risk T2DM than other types of fruit. Due to the largely white and educated American study population, these results may not be generalisable to all adults. However, it is probable that the biological mechanisms underpinning these results remain the same across different populations.

A further prospective study involving the same three study cohorts (133,468 US men and women, followed up for 24 years) by Bertoia et al investigated the change in fruit and vegetable intake and weight change.[54] The study population and dietary assessment every four years

were as described above, while adjustments for covariates also included additional lifestyle and dietary factors, such as hours of sleep, hours spent watching the TV, change of intake in fried foods, whole grain, nuts, sweets, processed meats and other food groups. Results pooled across all three cohorts showed that increased intake of fruits and non-starchy vegetables (cruciferous and green leafy vegetables) was inversely associated with weight gain over time and provided more evidence that the strength of the association depended on the type of fruit or vegetable. Total fruit consumption was associated with a weight change of -0.24 kg per daily serving over four years. Among fruits, particularly blueberries were most strongly inversely associated with weight change with -0.63 kg per daily serving over a four-year interval, followed by apples and pears with -0.56 kg per daily serving. In comparison, a daily serving of peaches, plums and apricots had no impact on weight change, and total vegetables, cruciferous vegetables and green leafy vegetables were inversely associated with weight change with -0.11 kg, -031 kg, and -0.24 kg per daily serving over a four-year interval (**Figure 1-7)**.

One explanation for the heterogeneity in the effectives of different types of fruit could be that the health benefits may be mediated by their nutritional or chemical components. Fruits and vegetables are good sources of antioxidants, vitamins, and minerals, including folate, vitamin C, and potassium, which may explain their protective effects. In addition, fruits and vegetables contain dietary fibre, which promotes gut health and helps regulate metabolic processes like blood sugar levels, satiety, and inflammation. High fibre intake has been linked to increased insulin-sensitivity and lower diet-induced obesity.[55] Fibre are carbohydrates from plants which are indigestible for humans, and can traditionally be classified by their solubility.[56] Insoluble fibre, such as cellulose, does not dissolve in water and helps food move through the digestive system and prevent constipation. Soluble fibre dissolves in water and incudes pectins and beta glucans which can be fermented by the gut microbiota to a variety of metabolites, such as short chain fatty acids (SCFA). The health benefits of fibre intake appear to be mediated by microbial fermentation products like SCFA and the protective effects exerted on the gut barrier. The recommended daily intake for adults in the UK is around 30 g per day. Generally, all plant foods contain soluble and insoluble fibre, but not all food sources contain all types of fibre. Soluble fibre such as pectins are mainly found in fruit, while beta glucans can be found in oats and cereals. Good sources of insoluble fibre are beans, wholegrains, fruits and vegetables. One cup of boiled black beans can provide about 15 g of fibre (around half of the recommended daily amount), while one apple or a pear can already provide around 5 g of fibre. Among fruits, raspberries are particularly high in fibre with 8 g per cup.[57]

Notably, the study examining the association of fruit and vegetable consumption with weight change, reported that adjustment for fibre intake did not attenuate the impact of fruit and vegetable intake on weight change, meaning the effect of greater fruit intake went beyond the impact of fibre on weight loss. Similarly, adjusting for total energy intake did not greatly change

the relationship. Because the dietary assessment did not allow for accurate measurements of total energy intake, this finding should be considered cautiously. However, taken together these results could give an indication that a mechanism other than fibre intake and reduction in calories are responsible for the association.

In addition to vitamins, minerals, and fibre, fruits and vegetables also contain phytochemicals. These are polyphenolic chemical compounds found almost ubiquitously in plants[58] which may have additional health-promoting benefits. In the last few decades, evidence has emerged that polyphenols in plants may be key contributing factors to the protective effects of a fruit-rich diet. A follow-up study using the same three cohorts by the same authors highlighted that the inverse association between fruit and vegetable intake and weight loss may be due to increased intake of flavonoids, a class of polyphenols.[59] Flavonoid intake was assessed by linking the FFQ data with a nutrient composition database with data on flavonoid content for foods. After multivariate adjustment, two flavonoid subclasses which are found in blueberries were negatively associated with weight change over each four-year interval: anthocyanins (-0.1 kg per additional standard deviation (SD, 10 mg)/day) and proanthocyanins (-0.24 kg per

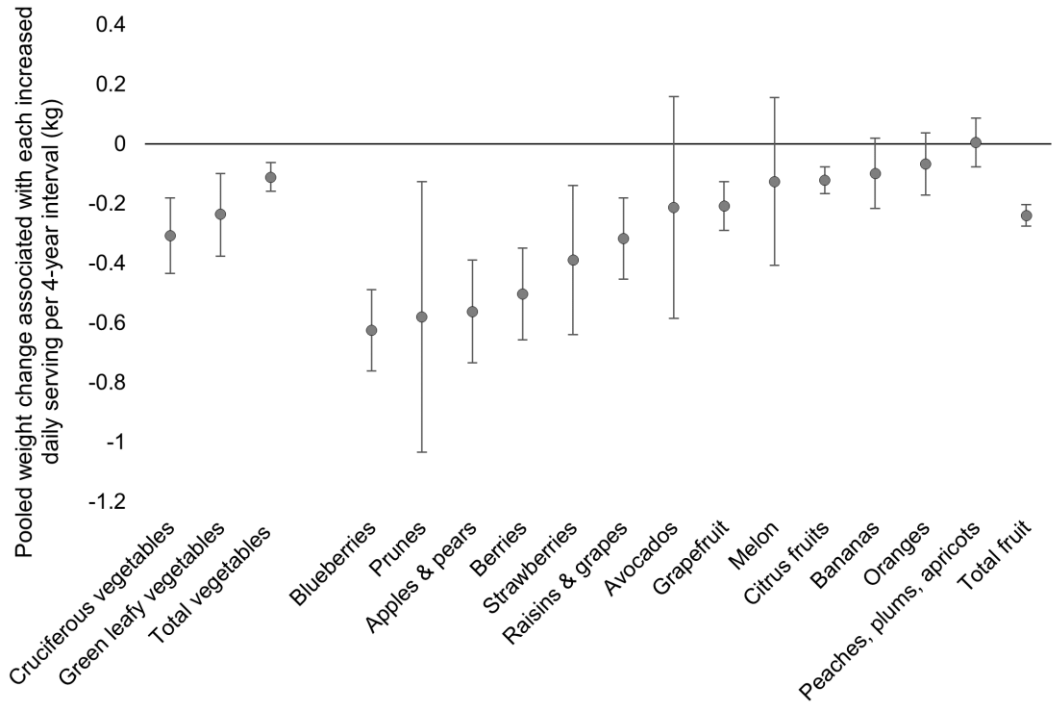


Figure 1-7. Increased fruit and vegetable intake associated with less weight gain over a 4-year interval

Figure shows weight change per 4-year interval associated with each increased daily serving of fruit or vegetable. Among fruits, particularly blueberries, apples, and pears were inversely associated with weight change. Error bars represent 95% CI. The estimates were pooled from three cohorts totalling 133 468 participants and adjusted for baseline age, baseline BMI, lifestyle variables including smoking, physical activity, sleep, television watching, and foods considered to affect weight, such as fried foods and whole grains.
Figure is adapted from Figures 1, 2, and 3 in Bertoia et al[54], copyright 2015 Bertoia et al used under the Creative Commons CC BY license (https://creativecommons.org/licenses/by/4.0/)

additional SD (72 mg)/day). Although the magnitude of association per SD may be small, a single serving of fruit or vegetables often offers more than one SD, e.g. a serving of blueberries equals 120 mg of anthocyanins or 12 standard deviations. Importantly, similarly to the second study mentioned above, a significant inverse association between anthocyanins and weight change remained after adjustment for fibre content. This suggests that food sources high in these flavonoids may prevent weight gain through mechanisms alternative to fibre.

Collectively, these three studies indicate that fruits may not be equally beneficial for health and that polyphenols in fruits may be key contributors to the underlying mechanisms. As Figure 1-6 and Figure 1-7 show, this is particularly true for blueberries.

1.3 Overview of dietary polyphenols

Polyphenols

Polyphenols are a large group of phytochemicals with many subclasses, of which some are toxic, and others show significant health benefits. In most cases, plant tissue contains a complex mixtures of polyphenols found within the call walls and inside the plant cell vacuoles, usually displaying a higher concentrations towards in the skin and seeds of the plant.[60] In nature, polyphenols perform a very large variety of important functions which include antioxidant properties, defence against pathogens, protection from UV radiation, structural support, or determining features such as colour, flavour and odour.[61]

Plant polyphenols can occur as simple monomers or highly polymerized. Predominantly they are also conjugated with one or more sugar residues in the form of monosaccharides, disaccharides or oligosaccharides[62] up to a size larger than 30 kDa[63]. Therefore, their molecular weight, water-solubility and their digestibility and intestinal absorption can differ to quite some extent.

Chemically, polyphenols are compounds with at least one aromatic ring which contain one or more hydroxyl groups bonded directly to an aromatic hydrocarbon. They are a large and diverse group, which can be divided up into four different classes: flavonoids, phenolic acids, stilbenes, and lignans (**Figure 1-8**).[64]

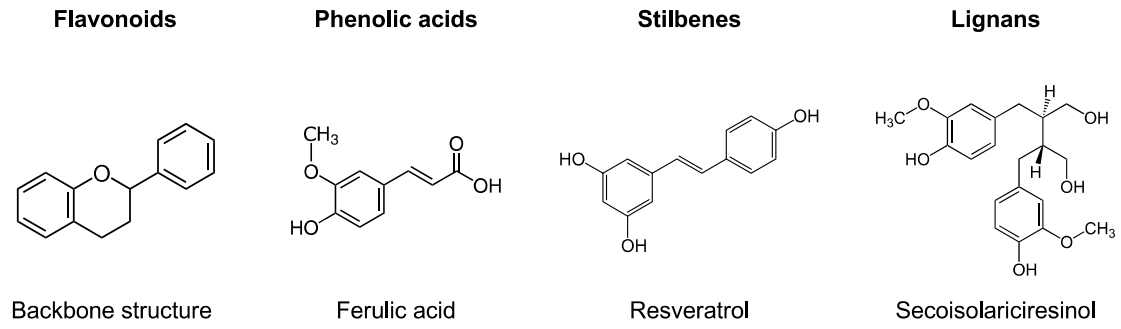


Figure 1-8. The four main polyphenol classes and exemplary molecules thereof

Polyphenols are a large family of compounds which contain at least one aromatic ring with one or more hydroxyl groups. They can be subdivided into the four main classes depicted in the figure: flavonoids, phenolic acids, stilbenes, and lignans.

Flavonoids

The chemical backbone of flavonoids consists of two rings of phenolic nature A and B and a heterocyclic ring C forming a three-carbon bridge between rings A and B. The 15-carbon structure is often abbreviated as C_6-C_3-C_6. The position of the B ring in relation to ring C and the addition of functional groups define multiple subclasses of flavonoids (see **Figure 1-9**). There are six main subclasses present in the habitual diet: flavan-3-ols (also called catechins) and their oligo- and polymers (proanthocyanidins), anthocyanins, flavonols, flavones, isoflavones, and flavanones. In plants, they most commonly occur as glycosides (with the addition of one or more sugar molecules), though flavan-3ols are an exception and are most often found as aglycones. Next to the addition of a sugar moiety, they can also be modified through acylation and methylation after synthesis. The structural differences alter their biological function in plants[65] and bioavailability and bioactivity in humans.[66,67] The main dietary sources of the main flavonoid subclasses are shown in **Table 1-1**.

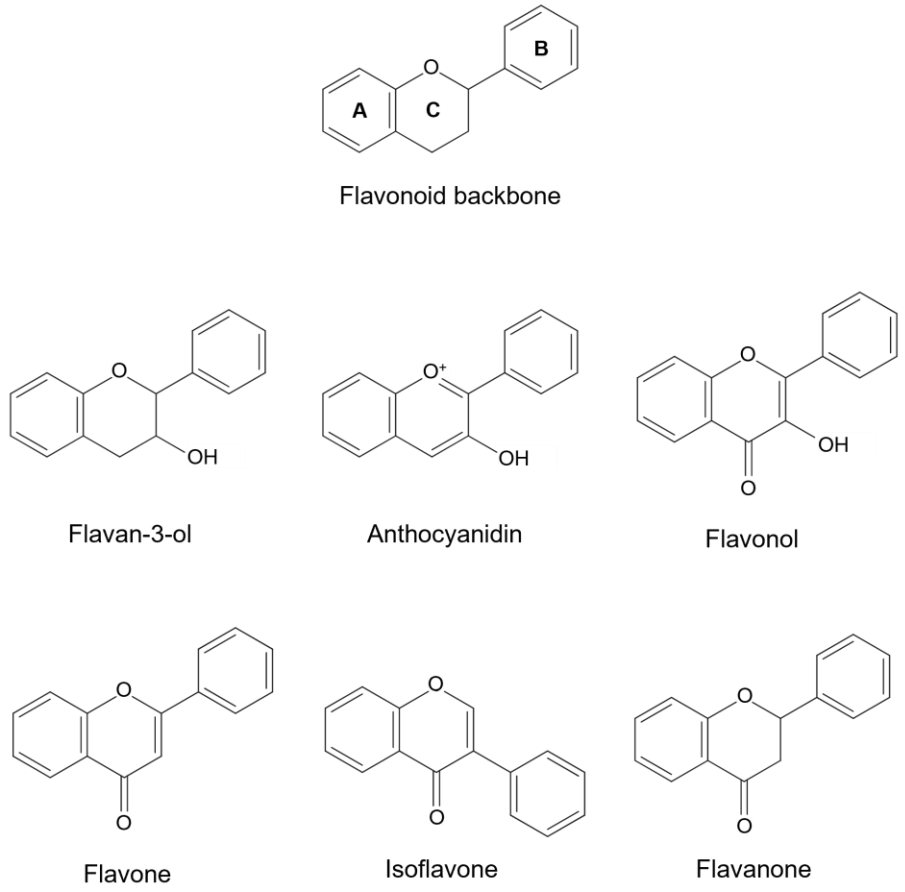


Figure 1-9. The chemical structures of the flavonoid backbone and six main flavonoid subclasses.

Flavonoids are plant secondary metabolites widespread throughout the plant kingdom and can be found in many commonly consumed fruits, vegetables, and beverages. The backbone consists of two rings of phenolic nature A and B and a heterocyclic ring C forming a three-carbon bridge between rings A and B. The subclasses shown in the figure are the six main subclasses present in the habitual diet.

Table 1-1. Examples of frequently consumed dietary sources of flavonoids in the UK diet.

Flavonoid subclass	Examples of dietary sources
Flavonols	Onion, green and black tea, leafy vegetables
Flavones	Celery, tea leaves and herbs (such as chamomile, parsley, mint)
Flavan-3-ols	Green and black tea, chocolate, red wine, apple, pear
Anthocyanidins	Red/blue/purple fruits (such as blueberries, blackberries, cherries, grapes) and vegetables (such as aubergine, red onion)
Flavanones	Orange, grapefruit and other citrus fruits and juices
Isoflavones	Soy products

In population-based studies, the habitual intake of anthocyanins has been shown to be among the most effective of the six subclasses in reducing the risk of cardiovascular diseases.

One prospective study investigated the incidence of myocardial infarction (MI) in 93,600 American women aged 25 to 45 years from the NHSII over 18 years of follow-up.[68] The authors found that the risk for MI was reduced by 32% (HR 0.68, 95% CI [0.49–0.96]) in the quintile of highest anthocyanin intake (25 mg/d) compared to the quintile with lowest intake (2.5 mg/d) after multivariate adjustment. A significant reduction in risk was only observed for anthocyanins, but not any other flavonoid subclass.

A similar study was performed in American men by the same authors. 43,800 men aged 32 to 81 years from the HPFS were observed over 24 years.[69] During this time, 4,046 cases of MI and 1572 cases of stroke were recorded. The authors only investigated the impact of anthocyanins and one other subclass, flavanones, on MI and stroke due to previous information regarding the associations between anthocyanins and MI risk, and flavanones and stroke risk in women. In agreement with the previous study, the results showed that anthocyanins reduced the risk for non-fatal MI by 13% (HR 0.87, 95% CI [0.75, 1.00]). They further showed that flavanones reduced risk for ischemic stroke by 22% (HR 0.78, 95% CI [0.62, 0.97]), but anthocyanins had no effect on stroke risk. The authors postulated that this difference in impact on CVD risk might be due to the fact that flavanones are better absorbed and may cross the blood-brain barrier to a greater extent.

A further prospective study found that anthocyanins reduced the risk of developing hypertension, as diagnosed by a doctor.[68] In total, 156,957 American participants from the NHS, NHSII, and the HPFS were followed up for 14 years. Anthocyanins were shown to reduce risk of developing hypertension by 8% (HR 0.92, 95% CI [0.86, 0.98]) in adults in the highest quintile of anthocyanin consumption in comparison to adults in the lowest quintile of anthocyanin consumption. The authors did not observe significant impact of total flavonoid or the other flavonoids subclasses on incidence hypertension, supporting the suggestion that anthocyanins are one of the most effective flavonoid subclasses in protecting the cardiovascular system.

A cross-sectional study provided further support for an inverse association of anthocyanins with blood pressure.[70] Measurements of arterial stiffness, as well as central and peripheral blood pressure were collected from 1898 women from the TwinsUK cohort. A comparison between participants with the highest and lowest intakes of flavonoids revealed that of total flavonoids and all flavonoid subclasses, only anthocyanins were associated with lower blood pressure (mean ± SE: central blood pressure -3.4 ± 1.41 mmHg; mean arterial pressure -2.31 ± 1.16 mmHg). Furthermore, anthocyanins and flavones were associated with lower arterial

stiffness, a marker of structural and cellular changes of the vessel wall (mean ± SE: 0.4 ±0.22 m/s and 0.4 ±0.25 m/s, respectively).

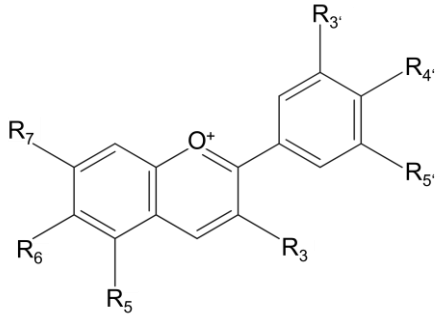
All four of the studies mentioned above have some limitations, mainly due to residual or unmeasured confounding and the dietary assessment. Dietary intake was assessed via FFQs and flavonoid intake estimated using the USDA flavonoid database.[71,72] This method will by its nature include measurement errors through over- or underestimation of food servings, by not capturing all sources of flavonoids, and by not capturing seasonal differences of flavonoid content in foods due to varying growing conditions or food processing steps. The cross-sectional study design does not allow for any inferences on causality. In addition, the results may not be generalisable to study population which are not predominantly white with European ancestry. However, it is likely that underlying biological mechanisms remain the same across population. Taken together these studies suggest that anthocyanins may exhibit a stronger protective role in comparison to other flavonoid subclasses and that further research is warranted.

In the US American cohorts, the main source of anthocyanins was from blueberries and strawberries, while in the UK cohort, the main contributors were grapes, pears, wine, and berries. This thesis will focus on anthocyanins, particularly those derived from blueberries, and investigate differences in how people metabolise these compounds after eating blueberries and whether these differences affect any cardiovascular risk markers.

Anthocyanins

Anthocyanins are the glycosides of anthocyanidins, i.e. they are made up of an aglycone (anthocyanidin), a sugar moiety, and often acyl groups which can impact the colour and function in plants.[73] Over 700 anthocyanins have been identified to date, of which 90% are based on the six common anthocyanidins cyanidin, delphinidin, pelargonidin, malvidin, peonidin, and petunidin (**Table 1-2**). Anthocyanidins are all derived from 2-phenylbenzopyrilium (flavylium cation) and differ in the hydroxyl and methoxy groups at the B-ring. While most species contain anthocyanins based on only one anthocyanidin, in some species, including blueberries, they are based on two or more. The most widely distributed anthocyanidin is cyanidin.

Table 1-2. Substitution pattern of the six most common anthocyanidins (adapted from [74,75])



	3	5	6	7	3'	4'	5'	Colour	Distribution in edible plant parts
Pelargonidin	OH	OH	H	OH	H	OH	H	Orange	12%
Cyanidin					OH	OH	H	Orange-Red	50%
Delphinidin					OH	OH	OH	Red	12%
Peonidin					OMe	OH	H	Orange-Red	12%
Petunidin					OMe	OH	OH	Red	7%
Malvidin					OMe	OH	OMe	Red-Blue	7%

Anthocyanidins are generally found as glycosides in plants. Anthocyanins contain one or two units of sugar molecules which make anthocyanins more stable and water-soluble than anthocyanidins. Glucose is the most frequently found sugar in anthocyanins, further sugars contained in anthocyanins are galactose, rhamnose, arabinose, xylose and glucuronic acid.[76]

The sugar molecule is normally bound via O-linkage to the 3-hydroxyl group and a sugar molecule located at the 3-position is present in almost all anthocyanins. Other possible positions for sugar attachments are 5, 7, 3', 4', and 5'. The anthocyanins found in varying degrees in blueberries depending on the variety are delphinidin, malvidin, petunidin, peonidin, and cyanidin as 3-glucosides, 3-galactosides, and 3-arabinosides.[76] The most common anthocyanin in edible plants is cyanidin-3-glucoside.[75] Furthermore, acylation to the sugar group often occurs with molecules such as organic acids (e.g. benzoic acid, hydroxycinnamic acids) and other flavonoids. This lends further stability to the anthocyanin.

General structural characteristics of anthocyanins can be described by an electron deficiency, which makes them extremely reactive towards free radicals like ROS and reactive nitrogen species (RNS), thus resulting in unique metabolic patterns and biological activities.[77] Anthocyanins may act as antioxidants and reduce oxidative stress by scavenging ROS and RNS, by chelating metal ions which can catalyse the formation of ROS, by stimulating the expression of detoxifying antioxidant enzymes such as the superoxide dismutase and glutathione-S-transferases, and by inhibiting pro-oxidant enzymes which produce ROS and RNS, such as the inducible nitric oxide synthase (produces NO) and NADPH oxidase

(produces the superoxide radical). This effective antioxidant property can prevent the oxidation of other molecules such as polyunsaturated lipids, proteins, and DNA. The degree of their antioxidant efficacy differs between different anthocyanins. It depends on the structure and extent of methoxylation and hydroxylation of the B-ring. Thus, delphinidin with three hydroxyl groups at the B-ring is the strongest antioxidant of the common anthocyanidins for scavenging the superoxide anion, followed by cyanidin and pelargonidin.

Multiple in vitro assays exist to measure antioxidant capacity by assessing either the amount of lipid peroxidation or the ability to scavenge free radicals. Using different in vitro methods, 3-glucosides of delphinidin, petunidin, and malvidin were shown to be powerful antioxidants, with a 2-times higher iron reducing capacity than the reference vitamin C and 3- to 6-times greater capacity to scavenge free radicals than a vitamin E analog (Trolox).[78] The antioxidant efficacy may also be enhanced or inhibited by the presence of other phytochemicals or vitamins. For example, in an in vitro setting, the flavanol catechin and anthocyanin malvidin-3-glucoside were observed to synergistically inhibit the peroxidation of linoleic acid, meaning their combined effect was stronger than the sum of their individual effects and indicating that catechin regenerated malvidin-3-glucoside.[79] Interestingly, the authors found that only malvidin-3-glucoside and to a lesser extent peonidin-3-glucoside were recycled by catechin, but did not observe this for the glucosides of delphinidin, cyanidin, and petunidin.

Cell models revealed that anthocyanins also demonstrate intracellular antioxidant activity. A cell study using human vascular endothelial cells observed that cyanidin-3-glucoside was transported across the cell membrane via the transporter bilitranslocase and demonstrated intracellular antioxidant activity even at very low, physiologically relevant concentrations (concentration at half maximal effect = 0.9 nM).[80]

In humans, antioxidant activity of anthocyanins and other flavonoids has been shown by assessing the antioxidant capacity of plasma or serum following the consumption of polyphenol-rich foods. As fruits also contain compounds other than polyphenols which exert antioxidant effects, a study tested whether blueberry increased serum antioxidant capacity in comparison to a control containing the same amount of fructose and ascorbic acid found in the equivalent amount of blueberries, each of which may contribute to the antioxidant properties of blueberries.[81] The authors reported that an easily achievable dose of 75 g blueberries significantly increased the serum antioxidant capacity over two-fold in comparison to the control in the first two hours postprandially, suggesting that blueberries were effective in providing acute protection from oxidation and that this was likely due to the phenolic component in blueberries. In a different study, Mazza et al investigated the postprandial serum antioxidant capacity in human subjects following the consumption of blueberry powder. Although the authors estimated that only a very small fraction of ingested anthocyanins were

absorbed (0.002 to 0.003%), they observed that the time-dependent appearance of anthocyanins in serum was correlated with the increase in the antioxidant capacity of the serum in comparison to a control supplement without blueberries.

As already implied by the study by Mazza et al, anthocyanins are absorbed very poorly as intact molecules (see details on absorption and metabolization in Section 1.5). Due to the rapid metabolization upon ingestion, anthocyanins likely only contribute to the circulation of antioxidant compounds for a short amount of time postprandially. Instead, degradation products and metabolites may play an important role in sustaining the antioxidant activity of anthocyanins. Anthocyanin metabolites such as hydroxybenzoic acids and hydroxycinnamic acids, which reach much higher concentrations in plasma, also exhibit antioxidant properties against different types of free radicals and could prevent or decrease overproduction of reactive species.[82,83]

The beneficial effect of anthocyanin intake on health has been studied intensively in epidemiological, clinical, and mechanistic studies for their impact on the cardiovascular system. Details are listed in **Table 1-3**.

- Intermediate markers of cardiovascular risk[84–91]: which include lowering blood pressure, improving endothelial function, reducing arterial stiffness, inhibiting vascular [68–70,92–94] and general[95,96] inflammation, improving dyslipidaemia[86,97], positively impacting bodyweight regulation and body composition[54,59,98].
- Disease incidence: decrease the risk of heart attack, stroke[68–70,92–94] and type II diabetes.[99,100]

Additional beneficial findings include associations with a reduction of colon cancer[101,102] and neurodegenerative diseases as well as improving cognitive function,[103–105] even with as little as a single dose of fresh blueberries.[106] Anthocyanins are also effective antioxidants and can therefore prevent the oxidation of other molecules such as polyunsaturated fatty acids, proteins, and DNA.[107]

Therefore, it is not surprising that the scientific interest and number of publications regarding the association of anthocyanins and health has risen drastically over the past two decades (**Figure 1-10**).

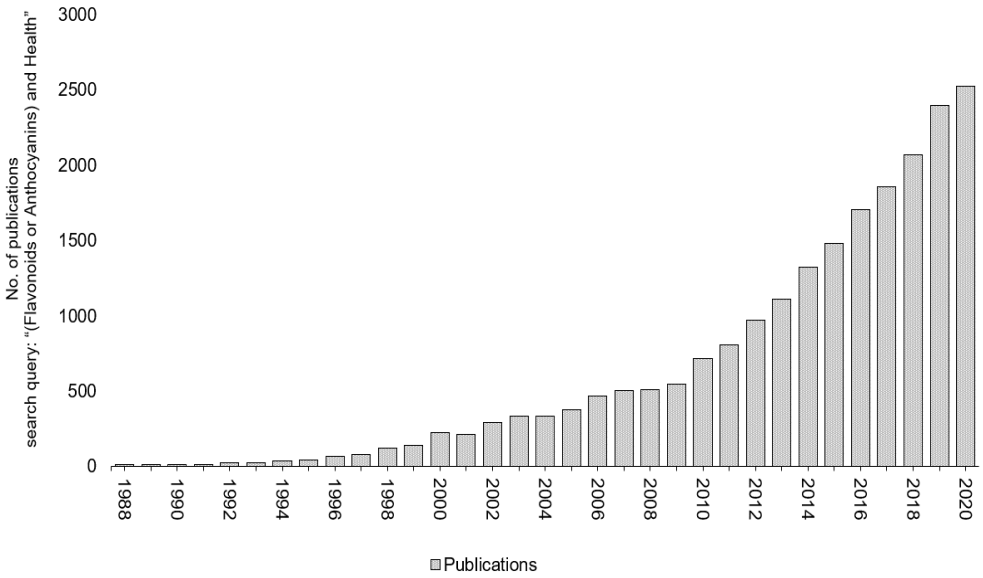


Figure 1-10. Flavonoid and polyphenol publications since 1988

The number of annual publications relating to flavonoids or anthocyanins and health since 1998. Data from PubMed[295] for the search query "(flavonoids or anthocyanins) and health".

Table 1-3. Summary table of studies demonstrating health benefits of anthocyanins.

Study aim	Study type	Study population	Country (cohort)	Study product	Outcome assessment	Effect estimate	Conclusion	Ref
Effects of habitual intake of dietary flavonoids on blood pressure and risk of hypertension in population-based studies	population-based prospective, 14 years follow-up	156,957 health professionals	US (NHS; NHSII; HPFS)	FFQ; USDA	Incident hypertension	HR [95% CI]: 0.92 [0.86, 0.98] (highest vs lowest quintile after multivariate adjustment)	Anthocyanins compounds may contribute to the prevention of hypertension and have vasodilatory properties	108
Examine the relation between habitual anthocyanin intake and coronary artery disease and stroke.	population-based prospective, 24 years follow-up	43,880 men	US (HPFS)	FFQ; USDA	nonfatal myocardial infarction (men)	HR: 0.87 [0.75, 1.00]	Higher intakes of fruit-based anthocyanins were associated with a 13% lower risk of nonfatal MI in men.	69
Examine the relationship between habitual anthocyanin intake and risk of myocardial infarction in young and middle-aged women.	population-based prospective, 18 years follow-up	93,600 young and middle-aged women	US (NHS II)	FFQ; USDA	myocardial infarction (women)	HR: 0.68; [0.49-0.96]	A high intake of anthocyanins may reduce myocardial infarction risk by 32% in predominantly young women.	68
Examine the association between change in intake of specific fruits and vegetables and change in weight.	population-based prospective, 24 years follow-up	133,468 Health Professionals	US NHS; NHSII; HPFS	FFQ; USDA	Intake of fruits was inversely associated with 4-year weight change	For each increased daily serving: Blueberries: ↓ 1.38 lbs Apples/pears: ↓ 1.24 lbs Berries: ↓ 1.11 lbs Bananas: ↓ 0.22 lbs Peaches, plums, apricots: ↑ 0.01 lbs	Increased consumption of fruits was inversely associated with weight change, with important differences by type suggesting that other characteristics of these foods influence the magnitude of their association with weight change.	54
Examine whether dietary intake of specific flavonoid subclasses is associated with weight change over time.	population-based prospective, 24 years follow-up	124,086 men (Health Professionals)	US (NHS; NHSII; HPFS)	FFQ; USDA	Weight change per four-year interval	↓ −0.23 lbs per additional 10 mg of anthocyanins per day	Higher intake of foods rich in anthocyanins may contribute to weight maintenance in adulthood	59
Examine the associations between flavonoid intakes and fat mass	population-based cross-sectional	2,734 women	UK (TwinsUK)	FFQ; USDA	Limb-to-trunk fat mass ratio (FMR)	difference between extreme quintiles of anthocyanin intake: ↓ 0.03 ± 0.02	Higher habitual intake of anthocyanin is associated with lower fat mass	98

Table 1-3. Continued

Study aim	Study type	Study population	Country (cohort)	Study product	Outcome assessment	Effect estimate	Conclusion	Ref
Examine associations between habitual flavonoid intakes and direct measures of arterial stiffness, central blood pressure, and atherosclerosis	population-based cross-sectional	1898 women	UK (TwinsUK)	FFQ; USDA	central SBP MAP PWV DBP	highest vs lowest quintile of intake: ↓ 3.04 mmHg ↓ 2.31 mmHg ↓ 0.4 m/s ↓ 1.86 mmHg	Higher intake of anthocyanins is inversely associated with lower arterial stiffness and central blood pressure	70
Examine associations between habitual intake of flavonoid subclasses, insulin resistance, and related inflammatory biomarkers.	population-based cross-sectional	1997 women	UK (TwinsUK)	FFQ; USDA	HOMA-IR insulin concentration hsCRP concentration	highest versus lowest quintile of intake of anthocyanins: ↓ 0.1 ↓ 0.7 µU/mL ↓ 0.3 mg/L	Higher intake of anthocyanins is associated with significantly lower insulin resistance	99
Effects of cranberry juice on vascular function in subjects with coronary artery disease	chronic RCT cross-over placebo-controlled	44 men and women with coronary artery disease	US	Cranberry juice, daily intake of 480 mL providing 94 mg anthocyanins	cfPWV	↓ 0.5 m/s	Chronic cranberry juice consumption reduced carotid femoral pulse wave velocity—a clinically relevant measure of arterial stiffness.	84
Effects of daily blueberry consumption for 8 weeks on blood pressure and arterial stiffness in postmenopausal women with pre- and stage 1-hypertension.	chronic RCT parallel placebo-controlled	Treatment: 25 Placebo: 23 postmenopausal women with pre- and stage 1-hypertension	US	22 g freeze-dried blueberry powder/d providing 470 mg anthocyanins	SBP DBP baPWV NO level	↓ 7 mmHg ↓ 5 mmHg ↓ 97 cm/s ↑6.24 µmol/L	Daily blueberry consumption may reduce blood pressure and arterial stiffness (compared to baseline levels), which may be due, in part, to increased nitric oxide production.	85
Effects of purified anthocyanins on dyslipidaemia, oxidative status, and insulin sensitivity in patients with type 2 diabetes.	chronic RCT parallel placebo-controlled	Treatment: 29 Placebo: 29 men and women with type 2 diabetes	China	purified anthocyanins from bilberry and black currant; 160 mg twice daily	LDLC TG Apo B-48 Apo C-III HDLC Fasting glucose Adiponectin β-hydroxybutyrate	↓ 7.9% ↓ 23.0% ↓ 16.5% ↓ 11.0% ↑ 19.4% ↓ 8.5% ↑ 23.4% ↑ 42.4%	Beneficial metabolic effects in subjects with type 2 diabetes (compared to placebo group) by improving dyslipidaemia, enhancing antioxidant capacity, and preventing insulin resistance.	86

Table 1-3. Continued

Study aim	Study type	Study population	Country (cohort)	Study product	Outcome assessment	Effect estimate	Conclusion	Ref
Effect of daily dietary supplementation with bioactives from blueberries on whole-body insulin sensitivity	chronic RCT parallel placebo-controlled 6 weeks	Treatment: 15 Placebo: 17 men and women, obese, nondiabetic, and insulin-resistant	US	smoothie containing 22.5 g blueberry bioactives twice daily; 668 mg anthocyanins in total	Mean change in insulin sensitivity	↑ 1.3 mg/kg FFM/min	Favourable change in insulin sensitivity	87
Effects of pomegranate juice supplementation on pulse wave velocity and blood pressure	chronic RCT parallel placebo-controlled 4 weeks	Treatment: 24 Placebo: 24 healthy young and middle-aged men and women	UK	330 ml/day of pomegranate juice	SBP DBP MAP	↓ 3.14 mmHg ↓ 2.33 mmHg ↓ 2.60 mmHg	Pomegranate juice supplementation has benefits for blood pressure in the short term	88
Effects of berry consumption on haemostatic function, serum lipids, and blood pressure	chronic RCT parallel placebo-controlled 8 weeks	Treatment: 35 Placebo: 36 middle-aged men and women with cardiovascular risk factors	Finland	2 portions of berries daily: berry products alternated between black currants, lingonberries, bilberries, chokeberries, strawberries and raspberries. Mean daily anthocyanin intake: 515 mg	HDLC CADP-CT (platelet function)	↑ 5.2% compared to 0.6% placebo ↑ CT prolonged by 11.0% compared to -1.4% (placebo)	Favourable changes in platelet function, HDL cholesterol, and blood pressure after berries consumption.	89
To investigate whether dietary anthocyanins exert direct effects on endothelium-dependent vasodilation	acute RCT crossover placebo-controlled	12 hypercholesterolemic men and women	China	purified anthocyanin from bilberry and black currant; 320 mg anthocyanins daily	Flow-mediated dilation	↑ from 8.3% at baseline to 11.0% at 1 h and 10.1% at 2 h	Purified anthocyanins improve endothelial function. Maximal plasma concentration of anthocyanins 1 h after supplementation was associated with maximum FMD.	90
Long term effects of anthocyanin supplementation on endothelial function in hypercholesterolemic individuals	chronic RCT parallel placebo-controlled 12 weeks	Treatment: 75 placebo: 75 hypercholesterolemic men and women	China	purified anthocyanin from bilberry and black currant; 320 mg anthocyanins daily	Flow-mediated dilation plasma cGMP	↑ 28.4% vs 2.2% (placebo) ↑ 12.6% vs -1.2% (placebo)	Anthocyanin supplementation improves endothelium-dependent vasodilation in hypercholesterolemic individuals. Suggests activation of the NO-cGMP signalling pathway.	90

Table 1-3. Continued

Study aim	Study type	Study population	Country (cohort)	Study product	Outcome assessment	Effect estimate	Conclusion	Ref
Dose-dependent impact of blueberry flavonoid intake on endothelial function in healthy men	two acute RCTs crossover placebo-controlled	21 healthy men	UK	Blueberry drink containing 129, 258, 310, 517 or 724 mg total blueberry anthocyanins	Flow-mediated dilation Polyphenol absorption as AUC of plasma metabolites Neutrophil NADPH oxidase activity	For 310 mg blueberry anthocyanin intake: ↑ 2.4% at 1h compared to -0.3 (control) ↑ 16.1 compared to 4.3 µmol · h/l (control) ↓ up to 35% reductions in neutrophil NADPH oxidase activity	Blueberry intake acutely improves vascular function in healthy men in a time- and intake-dependent manner. Suggested link of circulating phenolic metabolites and impact on neutrophil NADPH oxidase activity.	91
Effect of strawberry antioxidants on meal-induced postprandial inflammatory and insulin responses in overweight human subjects.	acute RCT cross-over placebo-controlled	26 overweight men and women with BMI ≥ 25 kg/m^2	US	Strawberry beverage containing 10 g of freeze-dried strawberry powder providing 82 mg total anthocyanins. Consumed together with a high-carbohydrate, moderate-fat meal (HCFM)	Postprandial response at 6h: IL-6 IL-1β plasma insulin level	↓ 3.4 vs 4.5 pg/ml (control) ↓ 0.15 vs 0.27 pg/ml (control) ↓ 56 pmol/l compared to (control)	Data provides evidence for favourable effects of strawberry antioxidants on postprandial inflammation and insulin sensitivity.	95
Effects of berry-derived anthocyanin supplements on the serum lipid profile in dyslipidaemia patients.	chronic RCT parallel placebo-controlled 12 weeks	Treatment: 60 Placebo: 60	China	purified anthocyanin from bilberry and black currant; 320 mg anthocyanins daily	HDLC LDLC Cholesterol efflux capacity	↑ 13.7% vs 2.8% (placebo) ↓ 13.6% vs -0.6% (placebo) ↑ 20.0% vs 0.2% (placebo)	Anthocyanin supplementation in humans improves LDL- and HDL-cholesterol concentrations and enhances cellular cholesterol efflux to serum.	97
Examine cognitive performance in middle-aged adults following wild blueberry (WBB) consumption	acute RCT cross-over placebo-controlled	35 middle-aged adults	UK	Wild blueberry beverage consisted of 25 g freeze-dried whole wild blueberry powder providing 475 mg anthocyanins	Cognitive function	↓ errors ↓ response times ↑ memory-related performance	Acute cognitive benefits of blueberry intake.	106

Table 1-3. Continued

Study aim	Study type	Study population	Country (cohort)	Study product	Outcome assessment	Effect estimate	Conclusion	Ref
Effect of supplementing an obesogenic (high-fat) diet with whole blueberry powder to test whether blueberry powder can protect against adipose tissue (AT) inflammation, insulin resistance (IR) and other obesity complications.	In-vivo (mice) 8 weeks	24 mice	NA	High-fat (60% of energy) diet 4% (wt:wt) whole blueberry powder (31g anthocyanins/kg dry food)	insulin resistance (glucose AUC) blood glucose frequency of dead adipocytes	↓ 550 vs 788 mmol/L·90min (control) ↓ 7.3 vs 8.6 mmol/l (control) ↓ 0.33% vs 0.64% (control)	Blueberry supplementation provides favourable cytoprotective and anti-inflammatory metabolic benefits to alleviate obesity-associated pathology	96
Effect of anthocyanin metabolites on vascular inflammation in human endothelial cells	Cell culture study	NA	NA	NA	TNF-stimulated secretion of (metabolite concentration, incubation time) VCAM-1 (20 µM, 6h) VCAM-1 (44 µM, 24h) IL-6 (2 µM, 1h) IL-6 (20 µM, 6h) IL-6 (44 µM, 24h)	 ↓ −65% ↓ −66% ↓ −37% ↓ −31% ↓ −36%	Signatures of anthocyanin metabolites reduce VCAM-1 and IL-6 production.	94

↑: increase; ↓: decrease; Adp: adiponectin; Apo: apolipoprotein; AUC: area under the curve; baPWV: brachial-ankle pulse wave velocity; CADP-CT: collagen-adenosine diphosphate closure time; cfPWV: carotid-femoral pulse wave velocity; cGMP: cyclic guanosine monophosphate; DBP: diastolic blood pressure; FMD: flow-mediated dilation; HDLC: high-density lipoprotein cholesterol; HOMA-IR: Homeostatic Model Assessment for Insulin Resistance; hsCRP: high-sensitivity C-reactive protein; IL-6: interleukin-6; IL-1β: interleukin-1β; LDLC: low-density lipoprotein cholesterol; MAP: mean arterial pressure; NADPH: nicotinamide adenine dinucleotide phosphate; NO: nitric oxide; PWV: pulse wave velocity; SBP: systolic blood pressure; TG: triglyceride; TNF: tumour necrosis factor; VCAM-1: vascular cellular adhesion molecule-1

1.4 Dietary intake of anthocyanins

Anthocyanins are readily available in the habitual diet (see **Table 1-4**) and can reach nearly 500 mg/100g of fresh weight in the highest sources such as elderberries. Rich dietary sources of anthocyanins are mainly berries such as blueberries, elderberries, and black currants, but also dark coloured vegetables like red cabbage, black beans, eggplant, and red onions.[109]

However, sources such as elderberries, red cabbage, and eggplants are seldom consumed raw. Raw food processing such as steaming, boiling, cooking can lead to an anthocyanin degradation of 10% to 80%, dependent on duration and temperature of the treatment.[110]

Table 1-4. Anthocyanin content in common dietary sources. (Based on [71])

Food	mg/100g fresh weight
Elderberries	485.3
Red cabbage	210.0
Blueberries	163.3
Black currants	157.8
Grapes (Concord)	120.1
Cranberries	104.0
Blackberries	100.6
Eggplant	85.7
Red currants	75.0
Plums	56.1
Raspberries	48.6
Black beans	44.5
Lingonberries	40.2
Red wine	35.6
Cherries, sweet	32.0
Strawberries	27.0
Pecan nuts	18.0
Grape juice	16.1

Despite anthocyanins being highly abundant in some foods, the intake levels at the population level are comparably small; anthocyanins make up only a small fraction of total flavonoid intake and the average intake has been reported to range from 1.6% in the USA[111] to 10% in Mediterranean countries[112]. Mean intake for adults (see **Figure 1-11**) were estimated to be 16 mg/d in the UK[113]; 3 mg/d[111] to 11.6 mg/d in the USA[114]; 19 – 65 mg/d in Europe[115]; 27.2 mg/d in Australia[116]; 27.6 mg/d in China[117].

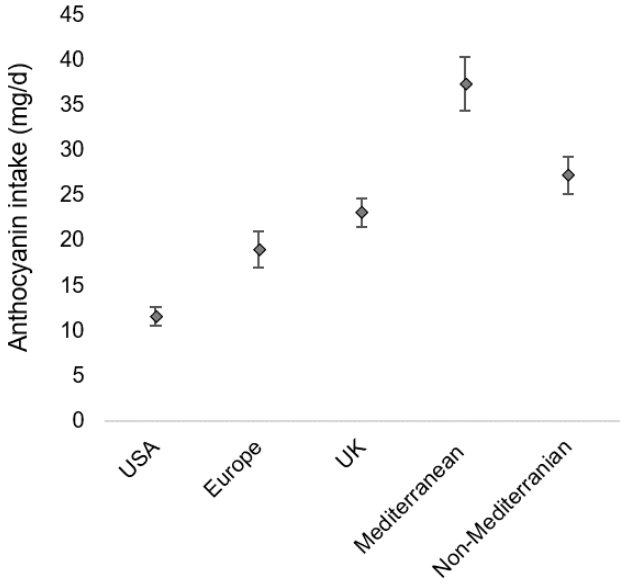


Figure 1-11. Mean anthocyanin intake across various countries and regions.

Estimated values for daily mean anthocyanin intake taken from various publications.[111–116] Dietary intake was assessed using 24h dietary recall[111–116] or 3-day food records[113]. Different flavonoid databases were used by the authors to estimate anthocyanin intake based on the dietary data (US Department of Agriculture Flavonoid Databases[111–116], Phenol Explorer[113,115,116], or AUSNUT 2011-2013[116]).

So far, no upper intake level has been published by food safety authorities nor an acceptable daily intake for anthocyanins in general. Based on toxicologic animal studies derived from grape-skin, blueberries and blackcurrant extracts, the European Food Safety Authority considered anthocyanins unlikely to be of safety concern.[118] Daily intake values of anthocyanins from currants, blueberries or elderberries at amounts of 25 mg/kg (mice), >3 g/d (guinea pigs and rats), >2.4% body weight (beagle dogs), and 9 g/kg (over 3 generations in rats, mice, and rabbits) have not resulted in any identified toxic effects.[74,119]

Usually, a mixture of polyphenol and flavonoid subclasses can be found in our diet. The respective proportion is largely dependent on the fruit or plant-product (**Figure 1-12**). For example, heavy coffee drinkers will likely consume more phenolic acids than flavonoids.[120]

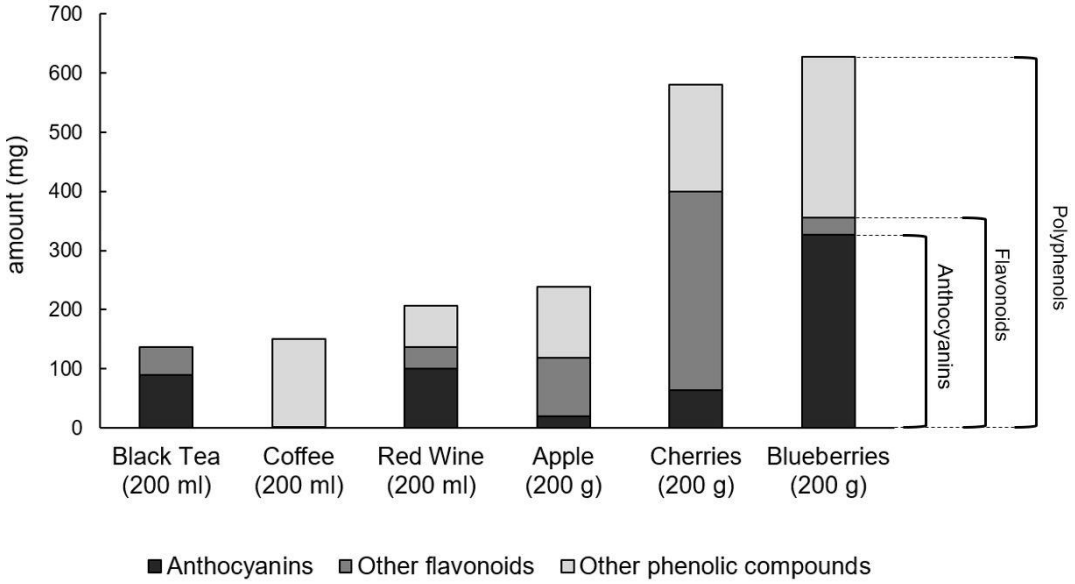


Figure 1-12. Total phenolic compounds, flavonoid, and anthocyanin content in selected examples

Examples of total phenolic compounds found in some regularly consumed fruits (in mg/200g) and beverages (in mg/200ml) and the contribution of flavonoids (dark grey + black) and anthocyanins (black) to the total phenolic compounds (light + dark grey + black). These values should only give a rough idea of phenolic and flavonoid content in common sources, as levels may vary widely, depending on growth conditions, ripening status, and soil.[120,296–300]

One of the first attempts from 1976 to measure and calculate an average daily consumption of total polyphenols, reported an estimate of 1g/day.[121] In a more recent study with 35,628 participants, typical dietary flavonoid intake was reported as 373 mg/day for non-Mediterranean countries and 370 mg/day for Mediterranean countries, in both cases the main contribution food sources were fruits.[112]

1.5 Metabolism of anthocyanins

After consumption, anthocyanins can be detected after a short amount of time (maximum concentration in plasma has been observed around 1.5 h after eating[122,123]), suggesting absorption in the upper gastrointestinal tract (GIT)[124] and possibly also absorption through the gastric mucosa.[125] The presence of intact anthocyanins in circulation suggests an active transport mechanism across the epithelium of the GIT, as they would be otherwise unlikely to pass the cell membrane due to their polarity and size. The transport mechanisms purported to be involved include glucose transporter 2 (GLUT2)[126] and sodium-dependent glucose transporter (SGLT1) which transport glucose and galactose across the intestinal barrier.[127]

However, while anthocyanins can be absorbed intact, they represent only about 2% of the total metabolites and only briefly appear in circulation ($t_{1/2}$ = 0.4h, not detected anymore after six hours).[123] In fact, the majority of anthocyanins is quickly metabolised, with one of the first metabolic reactions being deglycosylation in the GIT. This process is mediated either via mammalian enzymes in the small intestine (lactase phlorizin hydrolase in the brush border of the small intestine epithelial cells, or the cytosolic ß-glucosidase in intestinal epithelial cells) or by the microbiome.[124]. Following deglycosylation, anthocyanidins can chemically degrade in the small intestine to an aldehyde and a hydroxybenzoic acid due to their instability at physiological pH.[128] The expected products of this spontaneous degradation is given in **Table 1-5**.

Deglycosylated phenolic compounds are much less polar and can pass into the intestinal cells via passive diffusion. After absorption of the parent compounds or their metabolites, the phenolic compounds undergo phase I and II metabolism in the intestinal epithelium or liver to increase water solubility; similar to xenobiotics. Phase I metabolism, which involves the oxidation, reduction, and hydrolysis of molecules, appear to only play a minor role due to the rapid phase II conjugation following phase I.[129] Phase II enzymes convert compounds to more water soluble metabolites by conjugating them with polar groups such as sulphate, glucuronide, glycine, or methyl. These can then be excreted via urine or recirculated into the intestine via bile for faecal excretion. This re-entering of previously absorbed compounds into the GIT is called enterohepatic circulation. This exposes the compound to microbial metabolism once more and the possibility of reabsorption.[124]

Generally, the appearance of a compound in plasma within an hour after ingestion indicates absorption in the small intestine, while an appearance after 5 h is considered to be mainly from the large intestine[128]. Up to 85% of anthocyanins (dependent on the sugar moiety) pass from the small intestine unchanged into the colon[130], where these, as well as recirculated metabolites, are exposed to successive break down by the colonic gut microflora (described in more detail below).

Table 1-5. Anthocyanidins and their expected B-ring derived phenolic acids upon degradation

Anthocyanidin	Expected phenolic acid
Pelargonidin	4-hydroxybenzoic acid
Cyanidin	3,4-dihydroxybenzoic acid (protocatechuic acid)
Delphinidin	3,4,5-trihydroxybenzoic acid (gallic acid)
Peonidin	4-hydroxy-3-methoxybenzoic acid (vanillic acid)
Petunidin	3,4-dihydroxy-5-methoxybenzoic acid
Malvidin	4-hydroxy-3,5-dimethoxybenzoic acid (syringic acid)

Microbial metabolism of anthocyanins

The GIT includes all parts of the digestive system: mouth, oesophagus, stomach, small intestine and colon. Each section has its own specific conditions relating to absorption, the chemical environment, the presence of digestive enzymes and its specific microflora. In terms concentrations and diversity, the vast majority of the organisms in the GIT are located in the colon. It is estimated that 500 to over 1000 different species of bacteria, archaea and eukaryota reside within the colon of an individual. Their sheer number almost rivalling the human cell count (**Figure 1-13**).[131]

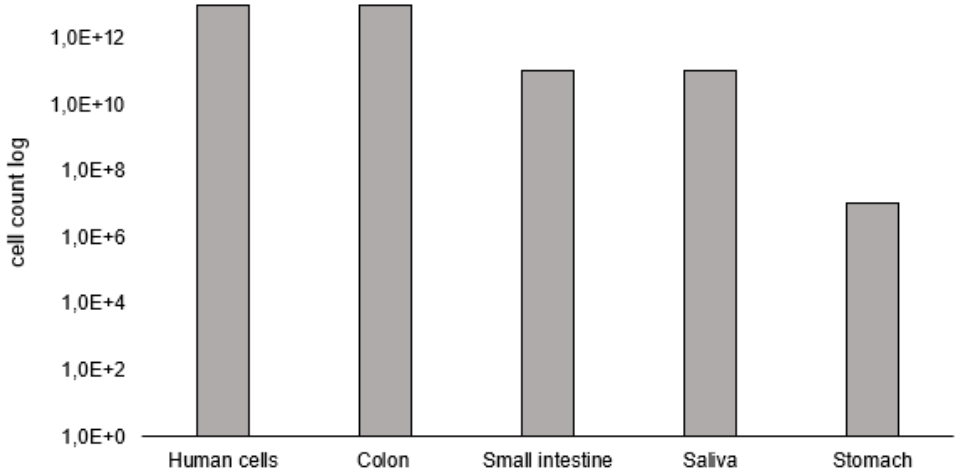


Figure 1-13. Comparison of estimated numbers of human and bacterial cells in the human body

The estimates for total human cells (first bar) and bacterial cells (second to fifth bar) in the human body are of similar magnitude. Bacterial cells are shown by distribution in the GIT. Based on Sender et al (2016)[131]

Known metabolic processes by the microbiome phenolic compounds include O- and C-deglycosylation, hydrolysis of esters and amides, deglucuronidation, and ring fission. The cleavage of the C-ring results in small microbial metabolites, such as cinnamic acids, phenylpropionic acids, and phenylacetic acids, which are then exposed to hydrogenation, α-oxidation and β-oxidation of the aliphatic elements, aromatic dehydroxylation, demethoxylation and demethylation. The gut microbiota can also break down the aromatic A-ring into CO_2 or into short chain fatty acids, which are absorbed through the colon, used for energy metabolism and exhaled as CO_2 (approximately up to 7% of consumed anthocyanins).[123,132]

For many of these transformations, the responsible organisms, corresponding enzymes, and the substrate specificity have not yet been characterised. In fact, this knowledge gap may remain for some time, considering the complexity and vast array of possible microbial candidates and how many are likely to be uncharacterised.

This immense complexity is compounded further by inter-individual variations in the microflora, which can change with age, diet and health status.[128] Presently, these factors remain poorly understood.

Phase I metabolism – modification

In the first phase of the human drug metabolism (also called xenobiotic metabolism) the absorbed anthocyanins undergo an 'activation' step to prepare them for the subsequent phase II conjugation. This is achieved by enzymatic introduction of reactive and polar groups into the substrate by oxidation, reduction, or hydrolysis.

During a feeding study with humans utilizing the ^{13}C isotope-labelled anthocyanin cyanidin-3-glucoside and subsequent identification of the metabolites, only small amounts of phase I metabolites were detected in serum and urine. It was therefore concluded that phase I metabolism was a minor pathway for anthocyanin metabolism.[123] Examples of phase I metabolites are shown in **Figure 1-14**.

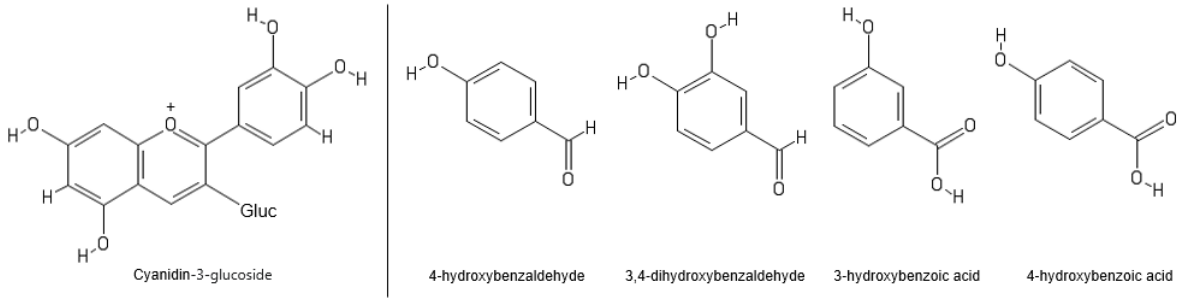


Figure 1-14. Examples of possible phase I metabolites derived from cyanidin-3-glucoside

Cyanidin-3-glucoside is first deglycosylated and is then metabolised or degrades to 3,4-dihydroxybenzoic acid. This then undergoes reduction reactions in phase I: a dehydroxylation results in the loss of a hydroxyl group at position 3' or 4'. The reduction of the carboxyl group yields a benyzaldehyde. Possible products are a benzoic acid or benzaldehyde with one or both hydroxyl groups.

Besides this, little is published specifically for phase I metabolites of polyphenols, flavonoids and anthocyanins. This could be because it is difficult to differentiate them from microbiome-derived metabolites prior to absorption. However, there have been efforts to predict phase I metabolites of flavonoids *in silico*,[133] or to produce them *in vitro* via recombinant enzymes.[134]

Phase II metabolism – conjugation

Phase II reactions are conjugation reactions which include acetylation, glucuronidation, sulphation, methylation, and glycine conjugation. The substrate must possess a reaction site for conjugation; these can already be present in the molecule or have been introduced during phase I metabolism. Reaction sites for the conjugation include carboxy (-COOH), amino (NH_2), thiol (-SH) and hydroxy (-OH) groups, with the latter being quite abundant in anthocyanins.

Enzymes that mediate phase II metabolism are found in many organs and tissues within the human body. The two main sites with the highest phase II metabolic activity are also the ones that come in contact with the substrate first: the epithelial cells of the intestinal mucosa and the liver.[135] Sulfation, glucuronidation, acetylation can all occur at the epithelial cells of the intestinal mucosa, which has been implicated as a major metabolic organ for many substances.[136] With regard to anthocyanins, phase II metabolites of anthocyanins or their degradation products (methyl, sulfate, glycine and glucuronic acid conjugates) could already be detected 0.5 h to 1 h after anthocyanin consumption.[123]

1.6 Bioavailability of anthocyanins

Bioavailability refers to the fraction of compound that is absorbed and present in the bloodstream. In a broader perspective it includes the intestinal metabolization and absorption, biotransformation, transportation and distribution to tissues, and excretion. Bioavailability of dietary flavonoids is usually characterised by maximum plasma or serum concentrations (C_{max}) or the fraction recovered via urinary excretion as a measure of the exposure.

Early assumptions on the bioavailability of anthocyanins were derived from animal and human feeding studies.[66,137–139] In comparison with other flavonoid classes, anthocyanins appeared to be one of the least bioavailable flavonoids.[66] Across these feeding studies, bioavailability of anthocyanins was observed to be poor, with less than 2% and often less than 0.1% of the consumed amount detected in urine. Anthocyanins were found to be rapidly absorbed and eliminated: the time maximum plasma concentration was reached (t_{max}) between 0.5 and 2 hours in most studies and could not be detected in plasma anymore 6-8 h post ingestion.

However, the reported poor bioavailability of anthocyanins is likely underestimated due to two main factors. First, analytical limitations[66,137] may lead to loss of anthocyanins in biological samples. Because anthocyanins are degraded easily in less than an hour in near neutral pH, the amount of anthocyanins present in biological samples, such as urine and serum, may be underestimated. Therefore, to preserve the most stable form, anthocyanins are typically extracted and analysed in acidic medium with low pH.[138]

In addition, the spontaneous degradation of anthocyanins into phenolic acids and aldehydes as well as the extensive metabolization of anthocyanins in the GIT to smaller molecules (more detailed in chapter ‘Microbial metabolism of anthocyanins’), likely leads to an underestimation of the bioavailability. Another source of error relates to accuracy of detection, as degradation products may easily escape detection, particularly if they are present in concentrations below detection limits of the quantifying method (such as liquid chromatography tandem mass spectrometry).

As a partial solution to these assessment limitations, labelling anthocyanins with carbon isotopes allows the tracking of metabolites and a more accurate assessment of bioavailability by measuring the recovered amount of the carbon isotope. Studies using such compounds are rare, however a ground-breaking human study which fed 500 mg of ^{13}C-labelled cyanidin-3-glucoside indeed showed that the bioavailability of anthocyanins is much higher than formerly perceived: a relative bioavailability of 12.38% of ^{13}C was recovered in urine (5.37%) and breath (6.91%).[140] Including the recovery from faeces (32.13%), a total of 43.9% of ^{13}C was recovered over 48h. Whilst this was a significant increase on previous estimates, it was notable that over half of the ^{13}C remained unaccounted for in this study. One explanation could be that the anthocyanins broke down into many diverse (and unmeasured) metabolites or existed at such low concentration that they could not be detected. It may also be that a large proportion remained in faeces after 48 h and a greater recovery may have been achieved with a longer sampling period.

Importantly, as already mentioned above, the parent compound cyanidin-3-glucoside represented only 2% of the total metabolites found in circulation and was detected only up to 6 h in serum. Conversely, ^{13}C-labelled metabolites were present in higher abundance throughout the 48h post-ingestion period of observation. This emphasised the role of phenolic metabolites as bioactive molecules and the likely mediators of potential clinical effects of anthocyanins in comparison to the parent compounds. In support of this, *in vitro* studies have shown anthocyanin metabolites to be more effective than their parent compounds in reducing the expression of inflammatory markers[141–143].

Factors which influence the absorption of anthocyanins

The chemical structure of anthocyanins impacts their absorption. For instance, the size and hydrophobicity can determine whether the molecule may enter the cell via diffusion or is transported across the membrane with a carrier. In support of this, a study investigating the uptake of blueberry anthocyanins in Caco-2 human intestinal cells showed differences in transport efficiency depending on aglycone structures and sugar moiety.[144] Delphinidin-glucoside was least efficiently absorbed, while malvidin-glucoside was most efficiently absorbed. This suggests that absorption efficiency increases with hydrophobicity – malvidin has two methoxy groups as B-ring substituents, while delphinidin has three hydroxyl groups and is therefore more polar.

It is also apparent that the type of sugar plays a role in the bioavailability of anthocyanins. In the same study as above, glucosides appeared to have higher transport efficiencies than galactosides. This was also demonstrated *in vivo* in rats where 22.4% of ingested cyanidin-3-glucoside were absorbed in the small intestine versus 13.6% of cyanidin-3-galactoside.[145] Moreover, acylation and glycosylation with more than one sugar molecule enhanced

anthocyanin stability and made them more resistant to metabolization and transformation by intestinal microflora.[146]

It is highly likely that other dietary components in the food matrix may also interact with anthocyanin absorption. For example, a human study investigated the effects on the bioavailability of strawberry anthocyanins over 6h when co-ingested with a meal consisting of a croissant with butter and jelly, sausage links, corn flakes, frosted flakes cereal and whole milk. [147] In this study, the authors observed that C_{max} dropped from 38.0 nmol/l when the strawberry drink was taken without the meal to 12.8 nmol/l when the drink was taken together with the meal. The meal also extended t_{max} from 1.7 h to 2.8 h. These findings suggest that the timing of the meal can have a substantial impact on the bioavailability of anthocyanins. Another feeding study comparing the bioavailability of strawberry anthocyanins when given with or without cream showed that C_{max} was not significantly different, but t_{max} was delayed by over one hour from 1.1 h to 2.4 h.[148] Interestingly, the urinary excretion over 24 h was not significantly affected by cream, suggesting that the cream did not impact on total absorption, but delayed gastric emptying and gastrointestinal transit time.

1.7 Inter-individual variability and phenotypes

Although a body of population-based studies have associated clinical benefits with anthocyanin intake, findings from human randomised controlled trials (RCTs) are equivocal. Likewise, it is widely accepted that inter-individual variability in response to flavonoids interventions exists, which may partly be responsible for the contradictory findings from RCTs.

Inter-individual variability relates to differences in absorption and metabolization of flavonoids between individuals and implies that individuals are exposed to different levels of bioactive compounds.

For example, in a one-year study, epicatechin levels in 24h pooled urine were measured at the beginning of the intervention and after one year of daily intake of 27 g flavonoid-enriched chocolate, providing 85 mg of epicatechin per day.[149] A wide inter-individual variability was observed for urinary excretion rates of epicatechin (flavan-3-ol), with individual excretion amounts ranging 9.6 to 327.0 µmol/24h (**Figure 1-15**).[150]

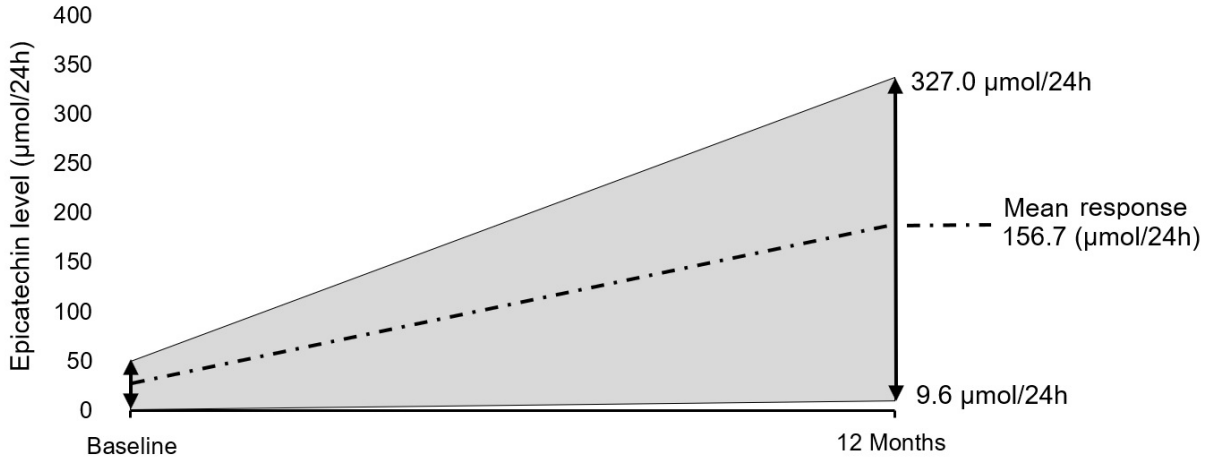


Figure 1-15. Figure illustration of the range of individual responsiveness

Figure shows the range of urinary excretion levels of epicatechin in 47 participants after one year daily intake of chocolate enriched with 85 mg epicatechin per day.

Adapted from Cassidy & Minihane (2016)[150], based on data published in Curtis (2012)[149]; figure copyright 2016 Oxford University Press used under the Creative Commons CC BY license (https://creativecommons.org/licenses/by/4.0/)

Differences between individuals in any step in the absorption, metabolization, distribution, and elimination of metabolites may contribute to variability in the bioavailability of flavonoids. To be more specific, such factors could include genetic variability (e.g., in the xenobiotic metabolism or transporters and carriers involved in the absorption from the gut), composition of the habitual diet, integrity of the intestinal barrier and constitution of the gut microbiome, as well as age and gender.[150]

Because the majority of anthocyanins pass into the colon intact, the metabolization by colonic microbiota plays a central role in the metabolism of flavonoids[128,151]. Thus, key determinants of inter-individual variability are likely to be the gut microbiota and differences in its composition between individuals. On this basis, several metabolic phenotypes (‘metabotype’) have been discovered which classify an individual into a ‘responder’ and ‘non-responder’ to flavonoid intake.

One of the earliest and most prominent examples of a metabolic phenotype is the ability to produce equol from the soy isoflavone daidzein, which is dependent on certain bacterial strains.[152] Equol has been shown to have vasoactive and immune modulatory functions,[152–154] however only 30% of the western population are equol producers. A further example is

the production of urolithins from ellagitannins (found in foods such as tea, pomegranate, and walnuts). [155] The production of urolithins is also dependent on certain strains of bacteria, and enables the classification of individuals into one of three metabotypes Uro-A, Uro-B, and Uro-0 (Uro-0 do not produce any urolithins). Interestingly, the prevalence of the metabotypes differs between healthy normoweight (20% Uro-B) and unhealthy individuals with metabolic syndrome (45% Uro-B), which may be associated with a gut dysbiosis in the unhealthy individuals. Importantly, the clustering by urolithin metabotype was recently shown to be associated with the observed improvements in cardiometabolic markers, such as oxidized LDL-cholesterol and small LDL-cholesterol.[156]

Unfortunately, not many polyphenols with unique catabolites are known. The microbial degradation of different and diverse flavonoids largely produces a similar set of catabolites of aromatic and phenolic acids, which include cinnamic acids, phenylpropionic acids, phenylacetic acids and benzoic acids [157]. This means that the different subclasses of flavonoids present in one dietary source converge at the metabolite level, which add difficulty to the identification of specific metabotypes in response to most flavonoid classes. In this case, multivariate analyses techniques which identify correlations between metabolites and identify clusters of similar participants based on their metabolic profile can help deduce the existence of metabotypes. These methods were applied recently to putatively define specific metabotypes following the intake of high amounts of flavan-3-ols from green tea and green coffee extracts (3469 mg (corresponding to 23 cups of green tea) daily for eight weeks).[158].

1.8 Research Gap

Following consumption, flavonoids are rapidly metabolised and there is strong support for the hypothesis that phenolic metabolites rather than their parent compounds are the bioactive molecules and likely mediators of potential clinical effects. Although in population-based studies the habitual intake of anthocyanins has been shown to be the most effective of the six main flavonoid subclasses in reducing the risk of CVD, findings from RCTs remain equivocal. It is widely accepted that inter-individual variability in the absorption and metabolization of flavonoids exists, in part driven by differences in the gut microbiota between individuals, which may partly be responsible for the contradictory findings from RCTs. In support of this rationale, several metabolic phenotypes (‘metabotype’) characterised by unique metabolism profiles following the intake of flavonoids have been identified, i.e. *equol* following isoflavone intake, and *urolithins* following ellagitannin intake. In both cases, these metabotypes have successfully classified individuals as ‘*responders*’ and ‘*non-responders*’ to flavonoid intake. Conversely, the main phenolic metabolites of anthocyanins are not unique, being commonly produced following the intake of different

flavonoids. Consequently, there has been no attempt to establish a metabolite profile to identify responders to anthocyanin intake to date. In addition, we lack an understanding how using such a metabolite profile to differentiate between responders and non-responders to anthocyanin intake may impact on assessing the clinical efficacy of anthocyanins in RCTs.

1.9 Thesis objectives

The first objective of this thesis was to investigate whether inter-individual variability in response to intake of anthocyanins from blueberries among participants of the CIRCLES study existed and could be characterised based on the urinary excretion of a single or a group of metabolites. A further aim was to determine the relationship between such a urinary metabolite profile and cardiometabolic markers, including vascular function.

The second objective was to assess whether the same urinary metabolite profile could be applied in an ongoing dietary intervention study to prospectively recruit individuals as high, medium, or low anthocyanin metabolisers.

2 Methods pertaining to the retrospective analysis of urinary anthocyanin metabolites and their association with cardiometabolic markers (Chapter 3)

2.1 The CIRCLES Study design, participants and intervention material

The aim of the CIRCLES study was to assess the effects of blueberry anthocyanins on insulin resistance and vascular function in a population with metabolic syndrome. The study design is briefly described in the following. Further details can be found in the paper published in 2019 by Curtis et al.[159]

The study was designed as a double-blind, six-month, placebo-controlled parallel dietary intervention conducted in men and women with overweight or obesity (BMI ≥ 25 kg/m^2), aged 50 -75 years, with at least 3 characteristics of metabolic syndrome (MetS) (central adiposity, elevated fasting glucose, elevated triglyceride levels, reduced high-density lipoprotein cholesterol, hypertension). Eligibility was based on a range of criteria, with exclusion criteria relating to smoking status, disease status, medication, and supplement intake. Eligibility to participate in the study was assessed through a combination of approaches: 1) the administration of a health and lifestyle questionnaire, collected after initial expression of interest to participate, and 2) a fasted clinical screening, conducted at the clinical research and trials facility based at the University of east Anglia and performed by study research nurses to assess general health parameters (assessed through clinical chemistry analysis following biological sampling).

Prior to the baseline assessment, participants were required to adhere to dietary restrictions for 21 days. Throughout the study, participants took a daily dose of one of three treatments (described below). Assessments were done at 0 months and 6 months in the morning after an overnight fast (≥ 10 h). Assessments included the collection of biological samples (24h pooled urine and blood sampling) and the assessment of cardiometabolic endpoints.

The primary outcome of the study was change in insulin resistance from baseline (0 months) to endpoint (6 months). Secondary outcomes were changes in blueberry metabolites, changes in vascular function (flow-mediated dilation (FMD), carotid-to-femoral pulse-wave velocity (PWV), augmentation index (AIx) and blood pressure), and changes in cardiometabolic biomarkers (fasting blood lipids, nitric oxide (NO) and related metabolites, fasting glucose, inflammatory markers).

Treatment groups, intervention material and compliance

A total of 138 participants were randomly assigned to one of three groups: full-dose blueberry (equivalent to 1 US cup or 150 g fresh blueberry per day), half-dose blueberry (equivalent to ½ US cup or 75 g fresh blueberry per day), or placebo. These doses were chosen to reflect dietarily achievable portions that could easily be incorporated into the habitual diet. Participants were randomised by four factors considered to likely affect insulin resistance and vascular function: sex, number of MetS criteria, age, and statin/anti-hypertensive medication use. Both participants and researchers conducting the study and analysing the biological samples remained blind to the treatment group throughout the study.

Participants were instructed to daily consume one of three intervention materials corresponding to their group allocation, which was given to participants in single-serve opaque sachets labelled only with the letters A, B, C to uphold the double-blind approach to masking the treatment allocation. 82% of participants were unable to judge their treatment group, as assessed by an exit questionnaire completed after six months.

The full-dose group received 26g of freeze-dried blueberry powder; the half-dose group received a mixture of 13 g freeze-dried blueberry powder and 13 g placebo powder; and the placebo group received 26 g of the placebo material (dextrose, maltodextrin, fructose, purple colouring, and blueberry aroma). All intervention foods were matched for calories, colour, sugar, and taste. The blueberries were provided by the United States Highbush Blueberry Council (USHBC; Folsom, California, US) and the placebo was produced by the National Food Laboratory (Plymouth, MN 55331, US).

The anthocyanin and phenolic content of the products were: 364, 182 and 0 mg anthocyanins and 879, 439, and 0 mg phenolics for the full-dose, half-dose, and placebo groups, respectively. Similar amounts of blueberries have been observed to improve insulin sensitivity, blood pressure and arterial stiffness[85,87], and the anthocyanin doses have been shown previously to affect endothelial function[90,91].

Compliance was measured by the number of returned wrappers and unused sachets. Overall compliance was high with 94.1% of sachets consumed across all three intervention groups.

Dietary and lifestyle restrictions

To minimise the interference of effects on the study outcomes from other dietary sources, participants were required to adhere to dietary restrictions for 21 days prior to the first assessment and throughout the study. Particularly, the intake of anthocyanin-rich foods as well as other commonly consumed rich in other flavonoids were limited. This allowed for a cleaner assessment of metabolites derived from the intervention material rather than other foods. Throughout the study, the intake of blueberries was completely prohibited, while the intake of

anthocyanin-rich foods (berries, red and purple fruits and vegetables) was limited to one portion a week. The intake of other commonly consumed foods rich in other flavonoids was also limited, such as dark chocolate (2 portions per week), red wine (1 small glass per week), tea and coffee (4 cups total per day). The intake of oily fish was also restricted due to potential effects on cardiometabolic measurements[160]. Immediately prior to assessment visits, further restrictions were implemented to improve the accuracy of metabolite and vascular assessments. This included the strict avoidance of strenuous exercise for 48 h before the study visit[161] and for 24h before the visit foods rich in anthocyanins, nitrates, nitrites[162], caffeine[163] and alcohol[164]. In addition, a standardised evening meal was given to control for any effect from the background diet on the repeated measurements at the two timepoints. The meal was free of anthocyanins and low in flavonoids, nitrite, and nitrate, and for 24 h before vascular assessments only low-nitrate bottled water (Buxton, Derbyshire, UK) was consumed.

The adherence to dietary restrictions was monitored using a validated food-frequency questionnaire[70,165] at baseline, interim (3 months), and six months.

Participants were also instructed to maintain habitual activity levels throughout the study which was monitored using the International Physical Activity Questionnaire (long last 7-day version[166]) at baseline, 13-week interim point, and 6-month endpoint.

2.2 Vascular assessments

Vascular assessments included measures of endothelial function, arterial stiffness and blood pressure.

Brachial artery flow-mediated dilation as a measure of endothelial function

Endothelial dysfunction describes a disorder of the vascular endothelium in which the endothelium-dependent vasodilation is decreased. It is recognized as an early key factor in the development of atherosclerosis and a marker for cardiovascular disease[167]. Endothelial function can be measured by flow-mediated dilation (FMD) of the brachial artery[168]. This non-invasive method uses ultrasound to measure the widening (dilation) of the artery as a reaction to acute reactive hyperaemia and a subsequent increased blood flow which causes shear stress. The increased blood flow is induced through a temporary occlusion (5 minutes) of the brachial artery in the upper arm through inflation of a sphygmomanometric cuff to 220 mmHg. Upon release of the cuff (after 5 minutes occlusion) a shear stress response is generated on the endothelium which increases NO production through the activation of the endothelial nitric oxide synthase (eNOS). NO induces relaxation of the smooth muscle cells of the vessel and leads to vasodilation. FMD is expressed as the percentage change at the maximum dilation from the baseline diameter, which is assessed for one minute under resting conditions. An

impaired endothelium is characterised by reduced NO availability and thus would dilate less than a healthy endothelium.

For the FMD measurement, the participant was lying in supine position with the arm rested in a comfortable position. A sphygmomanometric cuff was placed below the elbow, distal to the measurement site, which was at the upper arm. FMD was captured using ultrasound (Philips iE33; 11-3MHz linear transducer, Philips, Surrey, UK) with the image acquisition triggered by 3-lead ECG gating (Vascular Imager software; Medical Imaging Applications LLC, Coralville, USA). The protocol was as follows: 1 min baseline assessment of the vessel diameter, occlusion of the artery for 5 min by inflating the cuff to 220 mmHg, followed by a 5 min assessment of the vessel diameter post-occlusion after release of the cuff.

FMD images were independently assessed by two researchers (Dr Peter Curtis and Dr Lindsey Berends) and automated edge-detection software (Brachial Analyzer v5; Medical Imaging Applications LLC) was used to determine the vessel diameter. The pre-occlusion baseline diameter was averaged across all viable frames during the 1 min baseline assessment and the maximum diameter was averaged across at least 3 frames at the peak of post-occlusion diameter. The percentage change from baseline was calculated as

$$FMD\ (\%) = \frac{maximum\ diameter - baseline\ diameter}{baseline\ diameter} \times 100$$

The %FMD value assigned for each assessment was the mean of the independent analyses performed by the two researchers. An FMD measurements were excluded if both researchers judged the images as ‘failed’.

Measurement of systolic and diastolic blood pressure

With every heartbeat, the ejection of blood from the left ventricle of the heart into the aorta generates a pressure wave that travels down the arterial tree. The maximal pressure in the aorta and large arteries is termed systolic blood pressure (SBP). While the left ventricle refills, the pressure in the arteries drop. The lowest blood pressure is termed diastolic blood pressure (DBP) which occurs just before the next heart contraction. Blood pressure is influenced by a number of factors, including cardiac contractility, blood volume, the afterload (pressure in the arterial system) and vascular compliance. The latter factor is related to arterial stiffness – when compliance is reduced with increasing stiffness, resistance to blood flow is increased which in turn increases blood pressure. Hypertension, i.e. high blood pressure, is one of the characteristics of metabolic syndrome[10] and a well-established risk factor for cardiovascular disease and mortality[169,170].

Blood pressure was measured in the brachial artery using an automated sphygmomanometric cuff (Omron 705IT, Omron Healthcare Co., Kyoto, Japan) after 15 minutes of supine rest.

Blood pressure was measured in triplicate separated by three minutes each and was calculated as a mean of the last two measurements.

Pulse wave velocity and augmentation index as measures of arterial stiffness and wave reflection

Arterial stiffness involves structural and cellular changes of the vessel wall[171]. The structural changes are largely defined by the balance of two scaffolding proteins which differ in their elasticity, collagen (inelastic) and elastin (elastic). Increase in arterial collagen and degradation of elastin results in a stiffening of the arterial wall. In addition, cellular changes to the endothelium can promote arterial stiffness through an increased vascular tone. Further contributors to arterial stiffness are inflammation, oxidative stress and changes in the extracellular matrix of the wall[171]. Chronic hypertension can induce a thickening of the arterial wall mainly through hypertrophy of vascular muscle cells. Conversely, increased arterial stiffness is associated with an increase in blood pressure, resulting in a positive feedback loop and common progression of both arterial stiffness and hypertension[172]. While arterial stiffness is part of the normal aging process, it has been shown to be associated with increased risk for future cardiovascular events, including myocardial infarction, heart failure, or stroke[173].

Arterial stiffness affects the way blood pressure and the arterial diameter change with every heartbeat[174]. As the heart pumps blood from the left ventricle into the aorta, a pulse wave is generated that moves along the arterial tree. The speed at which this wave moves down a section of the aorta can be assessed and used as a measure for arterial stiffness. The stiffer the artery, the faster the transit time of the pulse. This measure is called pulse wave velocity and is often measured between the carotid and femoral artery.

The pulse wave is reflected at multiple points along its path, mostly at bifurcation points or changes in vascular resistance[175]. The final shape of the aortic pressure wave is the sum of the forward wave and the reflected backward wave. The backward wave mainly affects the systolic pressure, as diastolic and mean arterial pressure remain roughly the same (**Figure 2-1**). Increased amplitude of the reflected wave increases the systolic load at the left ventricle and was shown to cause a variety of cardiovascular complications, such as myocardial hypertrophy and fibrosis[174]. The augmentation pressure, i.e., the contribution of the reflected wave to the systolic pressure, can be measured and expressed as the percentage of augmentation pressure to total pulse pressure (augmentation index, AIx): $AIx = \frac{AP}{PP} \times 100$

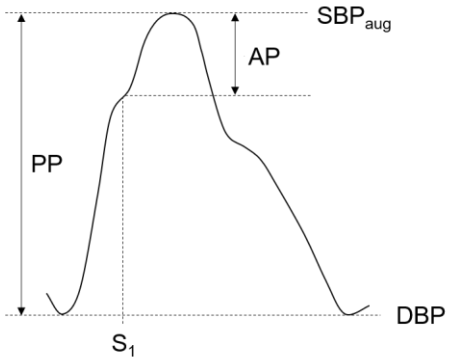


Figure 2-1. Aortic pulse pressure wave

The shape of the aortic pulse pressure wave is the sum of the forward wave generated by the pumping of the heart and the backward wave, which is the reflection of the forward wave at multiple points along the arterial tree. The reflected wave increases the systolic pressure of the forward wave (S_1) by the augmented pressure (AP). The pulse pressure (PP) is the difference between the total augmented systolic pressure (SBP_{aug}) and diastolic pressure (DBP). The augmentation index is the percentage of AP to PP.

In this study, arterial stiffness was assessed via the pulse wave velocity (PWV) between cuffs placed at the carotid and femoral artery. The distance between the cuffs was measured in duplicate and the calculation of the transit time was done automatically (Vicorder, Smart medical, UK).

Wave reflection (AIx) was automatically generated by the Vicorder software and standardised to 75 bpm. Up to six measurements were made with the aim to achieve ≤ 10% CV.

2.3 Assessment of biomarkers of CVD risk in fasted blood

After an overnight fast (≥10 h), venous blood was collected into heparin and serum tubes. Serum tubes were allowed to clot for 30 min before centrifuging, while heparin plasma tubes were centrifuged right away at 1300 g for 10 minutes. Samples were then aliquoted and stored at -80°C.

The cardiometabolic markers assessed were blood lipids, glucose, insulin, nitric oxide levels, and inflammatory markers.

Measurement of blood lipids, glucose, and insulin

Fasting glucose, total cholesterol, high-density lipoprotein cholesterol (HDLC) and triglycerides (TG) were assessed at the Norfolk & Norwich University Hospital using a clinical chemistry autoanalyzer (ARCHITECT c; Abbott Laboratories). Low-density lipoprotein cholesterol was calculated using the Friedewald equation [176]:

$$LDLC = {}^{(Total\ cholesterol - HDLC - TG)}/_{5}$$

Fasting serum insulin was measured in-house by the CIRCLES team using an enzyme-linked immunosorbent assay (Mercodia, Uppsala, Sweden) according to the manufacturer's instruction.

The intra-assay coefficients of variation were 1.51% (glucose), 0.86% (total cholesterol), 2.91% (HDL cholesterol), 1.23% (TGs), and 11.8% (insulin).

Dyslipidaemia, i.e., abnormal lipid levels, are one of the key attributes of metabolic syndrome. It is characterised by increased plasma levels of TG (hypertriglyceridemia), increased small density LDL and low HDLC. While LDL transports cholesterol to peripheral tissues such as muscles, HDL takes up free cholesterol from the periphery and transports them back to the liver for excretion.

Dyslipidaemia is likely the result of an interplay between excess nutrients, insulin resistance, and inflammation. Increased release of free fatty acids (FFA) from adipose tissue due to chronic low-grade inflammation, reduced clearance of FFA and *de novo* fatty acid synthesis in the liver due to insulin resistance may all contribute to the increased availability of FFA. This leads to the overproduction of TG-rich very-low-density lipoproteins (VLDL) and subsequently hypertriglyceridemia. In addition, the postprandial clearance of TG from dietary fats is impaired, further augmenting TG levels in circulation. Hypertriglyceridemia also stimulates the exchange of TG for cholesteryl esters between HDL and LDL particles causing a decrease in HDLC and an increase in small dense LDL particles, which have minimal amounts of TG left. Small dense LDL (sdLDL) are a subtype of LDL and sdLDL-cholesterol (sdLDLC) has been found to be a major contributor of the total LDLC associated with risk of coronary heart disease[35]. In contrast, HDLC has a protective effect on vascular function, therefore a decrease in HDLC is associated with greater CVD risk[36].

Assessment of insulin resistance

Insulin resistance (IR) is a state in which cells are less responsive to normal or elevated levels of insulin. In this study, IR was calculated using the homeostatic model assessment (HOMA)[177], which is calibrated to a normal IR of 1. It approximates IR with the following equation:

$$HOMA\text{-}IR = \frac{Glucose\ \left[\frac{mmol}{l}\right] \times Insulin\ [\frac{mU}{l}]}{22.5}$$

The model is based on the relationship between fasting glucose and insulin levels, which are regulated by the feedback loop between pancreatic β-cells and the liver. The liver plays a major role in balancing glucose levels, by storing glucose as glycogen molecules or releasing glucose via gluconeogenesis or glycogen breakdown. Insulin is secreted by pancreatic β-cells when blood sugar levels are high. It stimulates the uptake of glucose into liver, fat, and muscle cells

and suppresses gluconeogenesis in the liver. However, when cells are insulin resistant, the uptake of glucose into cells is reduced and plasma levels of glucose are increased. In the insulin resistant liver, reduced suppression of gluconeogenesis augments the already high glucose levels and contributes to hyperglycaemia. This in turn results in more insulin being synthesized by the pancreatic β-cells causing hyperinsulinemia. The increased production of insulin eventually causes endoplasmic reticulum (ER) stress and failure to alleviate this stress leads to death of the β-cells and progression to type 2 diabetes mellitus (T2DM). Impaired regulation of gluconeogenesis was found to be the main contributor to hyperglycaemia in individuals with type 2 diabetes[178].

Measurement of NO-related metabolites

To investigate underlying mechanisms of vascular function, components of the NO pathway were quantified. This was done by Prof. Martin Feelisch and his group at the University of Southampton.

Gas phase chemiluminescence and liquid chromatography were used to assess nitrite (NO_2^-), nitrate (NO_3^-), and total nitroso species (RXNO) levels in EDTA plasma[179]. Cyclic guanosine monophosphate (cGMP) was measured using an enzyme immunoassay (KGE003; RnD Systems). The intra-assay coefficients of variation were 1.6% (NO2−), 16.2% (NO3−), 15% (RXNO), and 9.9% (cGMP).

Endothelial dysfunction is characterised by an impairment of the endothelium-dependent vasodilation.[8] This is largely mediated via nitric oxide (NO), an oxidative signalling molecule produced in endothelial cells and a potent mediator of smooth muscle cell relaxation. The direct measurement of NO is technically difficult due to its reactivity and short half-life of approximately a few seconds.[180] Indirect measurements include nitrate, nitrite, and nitrosylated species. Nitrate and nitrite are products of the NO metabolism, but have also been shown to be recycled to form NO.[181] There are two main sources for nitrate and nitrite: the endogenous production of NO through nitric oxide synthases (NOS) and dietary sources. Particularly nitrate is abundant in many vegetables and requires the commensal bacteria in the gastrointestinal tract to convert it into nitrite. After ingestion of nitrate, both nitrate and nitrite levels increase in plasma. Nitrite has been reported to exhibit vasodilatory properties, and could therefore be seen as stable storage pool for NO-like bioactivity.[181]

Measurement of inflammatory biomarkers

Insulin resistance and obesity are accompanied by a chronic low-grade proinflammatory state. This includes increases in interleukin 6 (IL-6), tumour necrosis factor α (TNF-α), C-reactive protein (CRP), IL-8 and a decrease in adiponectin[10,182].

To evaluate the effect of the intervention on the inflammatory state, pro-inflammatory markers were assessed in heparin plasma using a sandwich immunoassay with an electrochemiluminescent detection method from Meso Scale Discovery (MSD). This assay allows the detection of up to ten analytes of interest simultaneously. Electrodes at the bottom of a 96-well plate are coated with antibodies which bind the analytes of interest in the sample. The addition of labelled detection antibodies, read buffer and the application of a voltage causes the labels to emit light. The intensity of the emitted light is measured and quantified against a standard curve.

The manufacturer provided immunoassay plates coated with antibodies on 1 to 10 spots per well, the detection antibody, calibrators, required diluents, blocker buffer, and read buffer. Calibrators and detection antibodies were prepared according to the manufacturer's protocol. In total, 40 inflammatory biomarkers were analysed. This included the human kits for adiponectin, RBP4, and Neuroinflammatory Panel 1 (IFNγ, IL-1b, IL-2, IL-4, IL-6, IL-8, IL-10, IL-12p70, IL-13, TNF-α, GM-CSF, IL-1a, IL-5, IL-7, IL-12/IL-23p40, IL-15, IL-16, IL-17A, TNF-b, VEGF, SAA, CRP, VCAM-a, ICAM-1**,** Eotaxin, MIP-1β, Eotaxin-3, TARC, IP-10, MIP-1α, MCP-1, MDC, MCP-4, VEGF-C, VEGF-D, Tie-2, Flt-1, and PlGF.

A general protocol provided by the manufacturer was followed and was similar for all analytes. Differences were in the type of diluents used and the sample dilution, which was adjusted to ensure sample readings were within range of the standard curve. Sample dilutions varied from 2-fold (most analytes) to 50000-fold (RBP4). In brief, first the plate was blocked with 150 µl blocking buffer (5% BSA in PBS) for 1 h at room temperature, shaking at 600 rpm. Then the plate was washed three times with 1X PBS + 0.05% Tween. The wash buffer was discarded and the plated tapped out on paper towels to remove any excess liquid. 50 µl of prepared calibrator, control or sample (heparin plasma) was added. The plate was shaken for 5 min at 600 rpm and then incubated overnight at 4°C for 16 h with mild shaking. The next morning, the plate was washed three times with 1X PBS + 0.05% Tween and tapped dry. The plate was incubated with 25 µl of prepared antibody detection solution for 2 h at room temperature while shaking at 600 rpm and then washed again three times with 1X PBS + 0.05% Tween. After the plate was tapped dry, 150 µl read buffer was added to the plate and analysed on the MESO QuickPlex SQ 120 imager.

Quality control criteria included intra- and interplate CV ≤ 20%; standard curve within 80 – 120% recovery; data completeness ≥ 85%. Due to very low concentrations, the measurement of some analytes was unreliable, resulting in high CVs. After quality control (QC), data for 28 analytes was available.

2.4 Quantification of blueberry anthocyanin metabolites

Blueberry anthocyanin metabolites were quantified in serum and urine. Blood samples were collected via venepuncture and allowed to clot for 30 min followed by centrifugation at 1300 g for 10 minutes. A 24h pooled urine collection was made prior to the assessment visit and included the first sample on the morning of the assessment day. Upon receipt, both serum and urine samples were acidified with 95% formic acid (reagent grade, Sigma Aldrich, Dorset UK) to prevent degradation of anthocyanins using 52.5 µL/mL for serum and 32µL/mL for urine. Aliquots of both serum and urine were stored at -80°C before analysis.

The samples were analysed for blueberry anthocyanin metabolites using liquid-chromatography tandem mass spectrometry (LCMS, the combination of HPLC (high performance liquid chromatography) and mass spectrometry) based on a published method [183] and adapted for this study by Dr Vera van der Velpen. I was substantially involved in the lab analysis of serum and urine samples and the subsequent data analysis and processing.

To summarise the method: 100 µl serum or 50 µl urine samples were purified through solid phase extraction prior to mass spectrometric analysis using a 96 well plate containing Strata-X- 33 µm polymeric reversed phase (Strata-X 33u Polymeric Reversed Phase, 60mg/1ml well, Phenomenex, Cheshire, UK). The samples were concentrated to 20 µl and spiked with 5 µl of 5 µM scopoletin as an internal standard (for adjustment to final sample volume). Matrix-matched standard curves were prepared from commercial serum or pooled urine ranging from 0 - 10 µM for serum (10 points) and 0 - 20 µM for urine (11 points). For hippuric acid and benzoic acid in urine an external standard curve was prepared which extended to 2000 µM.

Following the SPE, 1 µl of sample was injected onto a Kinetex polyfluorphenol column (2.6µM, 100 x 2.1mm; Phenomenex, Cheshire, UK) and separated with an Agilent 1200 HPLC at a column temperature of 37°C and a flow rate between 0.3 and 0.45 ml/min using 0.1% formic acid in water (mobile phase A) and 0.1% formic acid in acetonitrile (mobile phase B) over a total run time of 34 min. The elution gradient is shown in **Table 2-1**. The HPLC was coupled to an ESI-MS/MS (SCIEX 3200 series Q-trap MS/MS; SCIEX, Warrington, UK). Source parameters were as follows. curtain gas: 30 psi, ion source gas 1: 35 psi, ion source gas 2: 50 psi, ion spray voltage: -4000V/+4000V, temperature: 700°C. Compound parameters were independently optimised for each compound using syringe infusion from purchased standards (Sigma, Dorset, UK). The metabolites were identified based on retention time and three to four transitions (see **Table 2-2**).

Table 2-1. HPLC elution gradient for mobile phases A and B

Time (min)	Flow (µl/min)	A (%)	B (%)
0	450	99	1
7	300	92.5	7.5
14	300	92.4	7.6
17	450	90	10
18.5	450	88	12
20	450	87.5	12.5
24	450	70	30
25	450	10	90
28	600	10	90
29	450	99	1
34	450	99	1

Analyte peak areas and concentrations were determined using Analyst 1.5.1 (AB Sciex). Matrix-matched standard curves were prepared from commercial serum or pooled urine ranging from 0 - 10 µM for serum (10 points) and 0 - 20 µM for urine (11 points). For hippuric acid and benzoic acid in urine an external standard curve was prepared which extended to 2000 µM. For each compound six consecutive points were chosen for the quantification.

Further data processing was handled in Excel. This included adjustment of the concentrations for 1) a baseline peak in the 0 µM standard if present to obtain the absolute amount of analyte in the sample; 2) a final sample volume of 25 µl using the internal standard scopoletin; 3) the dilution factor (4 for serum and 2 for urine). To account for variations in urine volume, urine sample concentrations were further adjusted to the total volume of the 24h pooled sample, giving the total metabolite amount in µmol/24h.

After all adjustments, the dataset for final statistical analysis was generated in SAS (SAS v9.3/9.4, SAS Institute Inc., Cary, NC, USA).

LODs were calculated as the average of 3 times the signal-to-noise ratio from seven plates (see Table 2-2).

Table 2-2. HPLC-MS identification of phenolic compounds and their LOD in serum and urine

Compound	Mode	R_t (min)	ion transitions (m/z)	LOD[a] Serum	LOD[a] Urine
2,3-dihydroxybenzoic acid	-	8.2	153/109, 108, 91	389 ± 304	736 ± 174
2,4,6-trihydroxybenzaldehyde	-	15.6	153/151, 125, 107	95 ± 38	91 ± 32
2,4-dihydroxybenzaldehyde	+	11.6	139/121	N/A	3395 ± 850
2,5-dihydroxybenzoic acid	-	7.5	153/109, 108, 91	N/A	180 ± 50
2,6-dihydroxybenzoic acid	-	19.4	153/135, 109, 91, 65	116 ± 41	N/A
2,6-dimethoxybenzoic acid	-	10.9	181/137, 107, 92	501 ± 156	597 ± 184
2-hydroxy-4-methoxybenzaldehyde	+	23	153/92, 77, 63	965 ± 432	N/A
2-hydroxy-4-methoxybenzoic acid	-	23.4	167/123, 108, 80	17 ± 2	20 ± 8
2-hydroxy-6-methoxybenzoic acid	-	19.6	167/123, 108, 80	63 ± 19	59 ± 11
2-hydroxybenzoic acid	-	16	137/93, 75, 65	64 ± 28	86 ± 9
2-hydroxycinnamic acid	-	19.3	163/119, 117, 93	46 ± 32	48 ± 17
3,4,5-trihydroxybenzoic acid	-	2	169/125, 107, 79	48 ± 13	72 ± 10

Table 2-3. Continued

Compound	Mode	R_t (min)	ion transitions (m/z)	LOD[a] Serum	LOD[a] Urine
3,4,5-trimethoxybenzaldehyde	+	23.8	197/169, 154, 138, 123	50 ± 14	55 ± 13
3,4,5-trimethoxyphenylpropionic acid	-	23.4	239/180, 164, 149, 121	349 ± 87	594 ± 378
3,4-dihydroxy-5-methoxybenzoic acid	+	7.5	185/155, 146, 107, 78	490 ± 248	1208 ± 147
3,4-dihydroxybenzaldehyde	-	5.4	137/108, 92, 81	55 ± 25	48 ± 20
3,4-dihydroxybenzoic acid	-	3.8	153/108, 91, 81	52 ± 9	57 ± 14
3,4-dihydroxycinnamic acid	-	10.6	179/135, 106, 89	262 ± 74	236 ± 117
3,4-dihydroxyphenylacetic acid	-	4.2	167/123, 108, 95	405 ± 219	127 ± 73
3,4-dihydroxyphenylpropionic acid	-	7.1	181/137, 135, 121, 109	348 ± 189	506 ± 331
3,4-dimethoxybenzoic acid	-	17.8	181/137, 122, 107, 79	633 ± 302	1976 ± 794
3,4-dimethoxybenzyl alcohol	+	12.2	169/109	665 ± 348	N/A
3,4-dimethoxyphenylacetic acid	-	17.7	195/151, 136, 121, 93	51 ± 15	51 ± 25
3,5-dihydroxybenzaldehyde	-	5.8	137/109, 95, 93, 91	182 ± 60	206 ± 38
3,5-dihydroxybenzoic acid	-	3.9	153/109, 108, 91	157 ± 32	60 ± 6
3,5-dihydroxybenzyl alcohol	-	1.8	139/121, 109, 97, 95	98 ± 21	169 ± 63
3,5-dihydroxyphenylpropionic acid	-	6.6	181/137, 122, 95	171 ± 83	94 ± 19
3,5-dimethoxybenzaldehyde	+	25.9	167/139, 123, 105, 79	336 ± 226	N/A
3-hydroxy-4-methoxybenzaldehyde	-	11.8	151/136, 108, 92	1781 ± 1867	N/A
3-hydroxy-4-methoxybenzoic acid	+	6.7	169/151, 125, 93	1117 ± 482	141 ± 43
3-hydroxy-4-methoxycinnamic acid	-	19.4	193/178, 149, 134	170 ± 62	198 ± 53
3-hydroxy-4-methoxyphenylacetic acid	-	11	181/166, 137, 122, 94	719 ± 209	840 ± 105
3-hydroxybenzaldehyde	-	8.4	121/93, 92, 65	133 ± 34	163 ± 34
3-hydroxybenzoic acid	-	7.6	137/93, 75, 65	152 ± 31	151 ± 43
3-hydroxybenzoic acid-4-glucuronide	+	3	331/155, 153, 115, 113, 111, 109	12 ± 3	20 ± 7
3-hydroxybenzoic acid-4-sulfate & 4-hydroxybenzoic acid-3-sulfate[b]	+	7	235/191, 189, 155, 153, 110, 109	7 ± 5	13 ± 7
3-hydroxyhippuric acid	-	4.4	194/150, 148, 93	125 ± 100	13 ± 11
3-hydroxyphenylpropionic acid	-	12.1	165/121, 119, 106	99 ± 29	95 ± 20
3-methoxybenzoic acid	+	19.5	153/93, 75, 65	N/A	431 ± 412
3-methoxybenzoic acid-4-glucuronide	+	5.1	345/169, 167, 154, 152, 115, 113	21 ± 14	24 ± 7
3-methoxybenzoic acid-4-sulfate & 4-methoxybenzoic acid-3-sulfate[b]	+	10	249/169, 167, 154, 152, 110, 108	14 ± 9	12 ± 5
3-methoxyphenylpropionic acid	-	24.5	179/135, 120, 105	N/A	1958 ± 741
3-methylhippuric acid & 4-methylhippuric acid[b]	-	12.5	192/148, 146, 91	109 ± 87	140 ± 80
4-hydroxy-2-methoxybenzaldehyde	+	15.6	153/92	N/A	318 ± 92
4-hydroxy-3,5-dimethoxybenzoic acid	+	12.3	199/155, 140, 125	80 ± 45	68 ± 14
4-hydroxy-3,5-dimethoxycinnamic acid	-	21.1	223/208, 193, 164	15 ± 4	14 ± 2
4-hydroxy-3,5-dimethoxyphenylacetic acid	-	12.5	211/167, 152, 109	N/A	595 ± 188
4-hydroxy-3-methoxybenzoic acid	-	9.5	167/152, 123, 108	153 ± 87	135 ± 57
4-hydroxy-3-methoxycinnamic acid	-	18.8	193/178, 149, 134	49 ± 12	35 ± 11
4-hydroxy-3-methoxyphenylacetic acid	-	9.8	181/166, 137, 122, 79	1020 ± 322	1226 ± 464
4-hydroxy-3-methoxyphenylpropionic acid	-	14	195/136, 121, 119	42 ± 7	74 ± 80
4-hydroxybenzaldehyde	-	7.9	121/108, 92, 65	76 ± 63	49 ± 35
4-hydroxybenzoic acid	-	6	137/93, 75, 65	78 ± 33	72 ± 17
4-hydroxybenzoic acid-3-glucuronide	-	4.1	329/153, 113, 109	11 ± 8	21 ± 10
4-hydroxybenzyl alcohol	-	4.2	123/121, 105, 77	N/A	135 ± 60
4-hydroxycinnamic acid	-	16.1	163/119, 117, 93	64 ± 48	58 ± 16
4-hydroxyhippuric acid	-	3.7	194/150, 100, 93	11 ± 4	4 ± 2
4-hydroxyphenylacetic acid	-	6.4	151/123, 107, 93	1007 ± 305	N/A
4-hydroxyphenylpropionic acid	-	10.5	165/121, 119, 106, 59	428 ± 203	1447 ± 471
4-methoxybenzoic acid	+	19.5	153/135, 109, 92, 77	250 ± 46	N/A
4-methoxybenzoic acid-3-glucuronide	-	6.5	343/167, 152, 113	27 ± 4	31 ± 6
4-methoxyphenylpropionic acid	+	23.6	181/121	1213 ± 963	1196 ± 304
Benzoic acid[c]	-	12	121/93, 77	N/A	14 ± 1
Benzoic acid-4-glucuronide	+	2.5	315/177, 175, 115, 113, 95, 93	11 ± 4	25 ± 9

Table 2-3. Continued

Compound	Mode	R_t (min)	ion transitions (m/z)	LOD[a] Serum	LOD[a] Urine
Benzoic acid-4-sulfate	+	8.4	219/175, 173, 139, 137	25 ± 6	11 ± 9
Benzoylglutamic acid	-	9.7	250/206, 162, 121	23 ± 7	24 ± 8
Chlorogenic acid	-	11.7	353/191, 161, 127	47 ± 24	34 ± 13
Cyanidin-3-glucoside	+	12.7	449/287, 241, 157	20 ± 15	33 ± 9
Hippuric acid[c]	-	6	178/134, 132, 77	85 ± 108	8 ± 3
Methyl 3,4,5-trihydroxybenzoate	-	7.4	183/124, 106, 78	26 ± 15	26 ± 13
Methyl 3,4,5-trimethoxybenzoate	+	26.5	227/195, 168, 91	309 ± 236	N/A
Methyl 3,4-dihydroxybenzoate	-	12.5	167/108, 107, 91	27 ± 18	28 ± 10
Methyl 3,4-dimethoxybenzoate	+	25.6	197/165, 153, 138	54 ± 25	47 ± 9
Methyl 3,5-dihydroxybenzoate	-	12.2	167/152, 125, 123, 108	N/A	160 ± 71
Methyl 3,5-dimethoxybenzoate	+	26.8	197/165, 153, 138, 122	59 ± 35	59 ± 25
Methyl 3-hydroxybenzoate	-	19	151/136, 92	N/A	1292 ± 426
Methyl 3-methoxybenzoate & methyl 4-methoxybenzoate[b]	+	26.8	167/145	237 ± 86	403 ± 63
Methyl 4-hydroxybenzoate	-	18	151/136, 92, 91	18 ± 5	23 ± 8
Rosmarinic acid	-	25	359/197, 179, 161	8 ± 3	7 ± 2
trans-3-hydroxycinnamic acid	-	16,9	163/119, 93, 91	79 ± 56	75 ± 24

R_t = retention time in minutes.
[a] Limit of Detection (LOD) values are LOD ± SD. Calculated as 3 times the signal-to-noise ratio and averaged across 7 plates. N/A = no concentration data was available for this compound
[b] Isomers could not be separated sufficiently with HPLC and are presented as the cumulative of both isomers.
[c] Hippuric acid and benzoic acid were quantified using an external standard curve run in triplicate. SD reflects the triplicate measurement.

2.5 Statistical analyses

All statistical testing was performed in Stata 16.0 (Stata Corp, College Station, Texas, USA).

In part 1 of the results (Section 3.1) an exploratory factor analysis (EFA) was used to reduce the data dimension to two factors (see Section 2.6 for an introduction to EFA). For each factor, two types of composite measures were calculated: 1) factor scores, in which factor loadings are applied as weights to all variables, were computed using the Stata factor postestimation command and 2) average scores (in µmol/24h) were calculated as the average change of selected variables most represented by the extracted factors. Associations between composite scores from EFA and cardiovascular risk factors were assessed using Pearson correlation. High and low group comparisons were tested using Welch's t-test. Participants were clustered by plotting the ranks of the average scores against each other. The clustering was initially done manually and followed up by a cluster analysis using k-means clustering algorithm. One-way ANOVA was performed to test differences between the clusters, homogeneity of variance was tested using the Brown-Forsythe test.

In part 2 of the results (Section 3.2) the following additional analyses were performed:

- The metabolite panels were calculated as the average $\log_2$ fold change from baseline for selected metabolites: $log_2FC = log_2(endpoint) - log_2(baseline)$.

- Equality of variance between groups was tested based on the median using the Brown-Forsythe test before using one-way ANOVA to compare the metabolite panels across treatment groups.
- Univariable ordinary least-squares linear regression was used to assess the relationship between the metabolites/panels and cardiovascular risk factors; validity of the model was assessed by inspecting the residuals plot and assessing normality of the residuals using normal probability plots.
- In case of outliers, which were confirmed to not be due to measurement errors, a robust regression was performed using an MM-estimator[184,185].
- A non-parametric ROC analysis was performed using Stata's "roctab" command.

Across all treatment groups, participants were excluded on the basis of incomplete urine collections (n= 22), lacking metabolite data (n = 2), lower quality FMD data (n = 16), and change in blood pressure medication (n = 1).

Note to examiners:

The development of a working panel of target metabolites which characterized blueberry high metabolisers and low metabolisers was an exploratory and iterative process. The exploratory factor analysis described in Section 3.1 and the common metabolite analysis presented in Section 3.2 within this chapter were originally performed with all available data (n = 35) in the full dose group as the underpinning assessment to then apply in a prospective ongoing study (described in Chapter 5). Due to the limitations confirmed in Section 3.1.2 (see Challenge 1) regarding the use of factor analysis in small study sizes, and the need to have a confirmed panel to employ consistently in the AMP study (described in Section 3.2.3), it was considered justified to proceed prior to a more stringent application of data quality assessments in the latter stage of the thesis.

Recruitment for the study and use of the metabolite panel as a screening tool began at the end of my third year. A delay to the start of the study was introduced due to the move of the clinical research facility to the Quadram Institute, and was the reason I got an extension to my studies.

After completion of the AMP study recruitment, the process was reexamined and stepwise refined:

1) excluding those with incomplete urine collections

2) excluding those with FMD image sequences scored as poorer quality by the ultra-sonographers.

What follows in Chapter 3 (all analyses) represents the final, refined analysis (after the implementation of steps 1 and 2 above), as this was considered the most robust for purposes

of extensive thesis analysis (i.e., n = 25 participants in the full dose group were used for the exploratory factor analysis and subsequent analyses). The panel which resulted from the original analysis was reexamined in Section 3.2.3 (see Panel 5) to assess validity of its use when a more stringent data quality assessment was applied.

2.6 Exploratory factor analysis

Research question

In metabolomics, multivariate methods are applied to identify a few biologically relevant features among a large background of metabolites which uniquely define the system[186]. In other words, the goal is to find a pattern within the data. The first step in such an analysis is commonly a dimension reduction technique, which summarizes variables into a smaller, more manageable set of variables based on their correlations. There are two types of dimension reduction – the grouping of variables and the grouping of observations. The former is commonly done using principal components analysis or factor analysis. The latter can be achieved using a cluster analysis algorithm such as hierarchical clustering or k-means clustering. These two methods can also be combined. First, a 'metabolic fingerprint' of an underlying cause is identified, and then individuals are clustered based on their similarities in regard to the identified metabolic fingerprint.

The research objective in Chapter 2.7 was to determine if a metabolic phenotype existed which underpinned the formation of metabolite patterns in the CIRCLES study population. This analysis was approached in two steps: first, an exploratory factor analysis was performed to reduce the dimension of the metabolite data and identify clusters of urinary metabolites which may suggest a metabolic phenotype. Second, after identification of high and low metabolisers, the metabolic profiles of each group were examined more in-depth to create a panel of metabolites indicative of a high metaboliser of blueberry compounds. Hereafter is a description of the methods used in the first step.

Sample type and variables

Of the 72 metabolites quantified in urine, 46 metabolites were included in this analysis. Because of the data pre-treatment methods used, measurement errors particularly at low concentrations would have been inflated. To ensure high data quality, 26 metabolites were excluded if the compound was not detected or there was uncertainty in measurement accuracy. Uncertainty included low concentrations close to the LOD or unclear peak detection due to noisy chromatograms.

24h pooled urine samples (at baseline and 6 months) were chosen to determine metaboliser signatures in preference to single, fasted serum samples, as this enabled an assessment of the accumulated metabolite elimination over 24h. In contrast, metabolite concentrations in fasted serum samples would have shown only the metabolites in circulation at the time of blood sampling; thus providing an incomplete picture of likely metabolites in response to blueberry intake. In previous studies, urinary metabolites have correlated well with maximum serum concentrations [64], anthocyanin metabolite levels have been shown to be elevated over 24 h[140] and urine has been confirmed as the likely main pathway of phenolic acid elimination[48]. For these reasons, it was considered a more accurate reflection of the bioavailability of blueberry metabolites in this study than fasting serum.

The EFA was restricted to the full dose treatment group only. This was done as it was of particular interest to identify metabolic differences in people who received the same amount of blueberry compounds.

Brief introduction to EFA

This section is intended to give a brief overview of EFA to facilitate understanding of the results. More details of some concepts concerning EFA, such as sample size considerations, are included in the results chapter where pertinent.

Exploratory factor analysis (EFA), or also called common factor analysis, is a multivariate analysis technique which is used to identify an unobserved (latent) variable represented in the common variance of the observed variables. In simpler terms, it aims to identify patterns among the correlations between the variables and tries to find the best way to express those patterns by summarising them in a smaller number of variables called factors. It is similar to principal component analysis (PCA), which is one of the most widely used methods for dimension reduction in metabolomics[186]. Although EFA is less widespread, it has been used in nutritional research to identify dietary patterns associated with metabolic syndrome[187–189] and has also found application in the interpretation of variations in mass spectrometry data[190].

The fundamental difference between PCA and EFA is that PCA tries to identify components which are composites of the observed variables, whereas EFA assumes there is an underlying factor which affects the observed variables. EFA was therefore better suited than PCA in the context of the research question set out in this study. The rationale behind choosing EFA for this analysis was the hypothesis that a metabolic phenotype (= the unobserved factor) affected the metabolism of blueberry compounds and thereby the formation of phenolic metabolites (= the observed variables). The two different principles are illustrated in **Figure 2-2**.

Next to the conceptual difference, a further important difference between PCA and EFA is the way the variance in the data is viewed. EFA assumes that the total variance can be partitioned

into two parts: a common variance, which is the shared variance among the variables, and a unique variance, which is any variance which is not shared among the variables. The unique variance includes a specific variance, which is specific to a certain variable, and an error variance, which may arise from measurement errors. The presence of unique variance and measurement errors is a reasonable assumption in real data (i.e., not simulated data). When creating factors, EFA only considers the common variance. In contrast, PCA assumes that the total variance is equal to the common variance. While it preserves as much of the variance as possible, it is reasonable to assume that each extracted component includes some unique variance and measurement error.

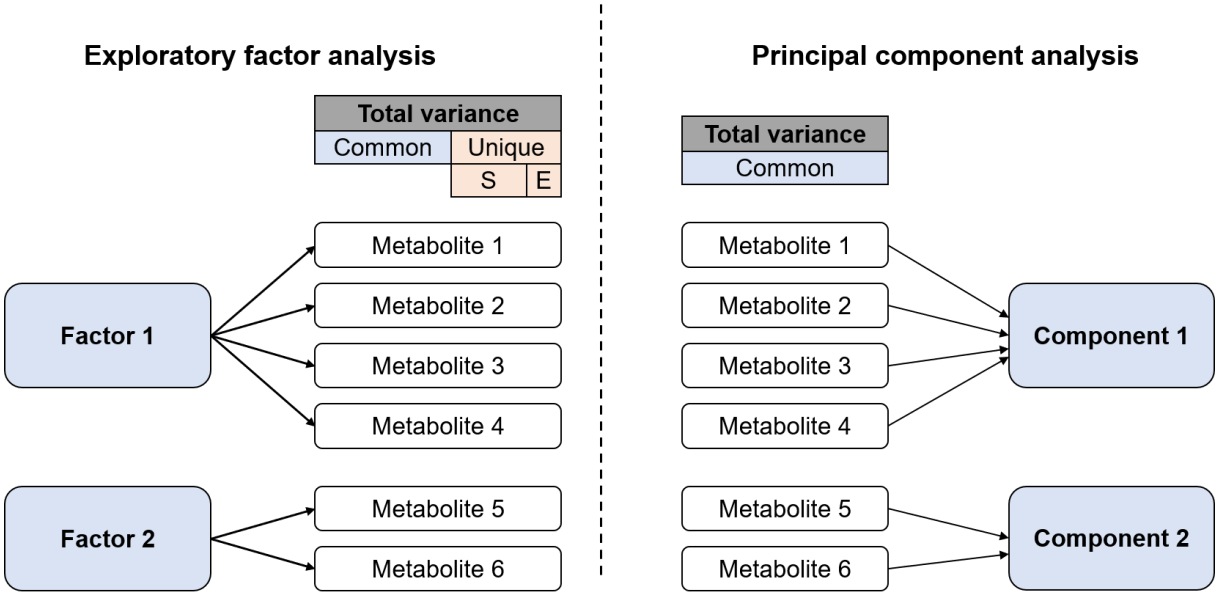


Figure 2-2. Concept of EFA and PCA

Exploratory factor analysis (left) assumes that an underlying factor gives rise to the observed variables. It differentiates between common variance among the variables and a unique variance, which includes a variance specific to a variable, and an error variance, which could arise from a measurement error. A factor represents only the common variance among variables, not the total variance. In contrast, principal component analysis (right) focusses on the observed variables and tries to find optimised weighted linear combinations of the variables. A component represents the total variance in the data and will therefore likely include specific variance and measurement error. S: specific variance; E: error variance

Decision process in EFA

A key component in EFA is the exploratory nature. It goes through multiple stages and finding the best solution requires returning to and repeating previous stages. The decision process followed in this study is outlined in **Figure 2-3**.

Stage 1. Data screening and pre-treatment

Data pre-treatment is crucial to the outcome of factor analysis and improves the biological interpretation and reliability of the results[191,192]. This stage involves considerations regarding

the available sample size (in this case n = 25), the cleaning of data (26 of 72 quantified metabolites were excluded to ensure high data quality), identification and handling of potential outliers, treatment of missing data, scaling of data (affects the range of each variable), and variable selection (affects the number of variables for the model). These considerations are explained in detail in the results section.

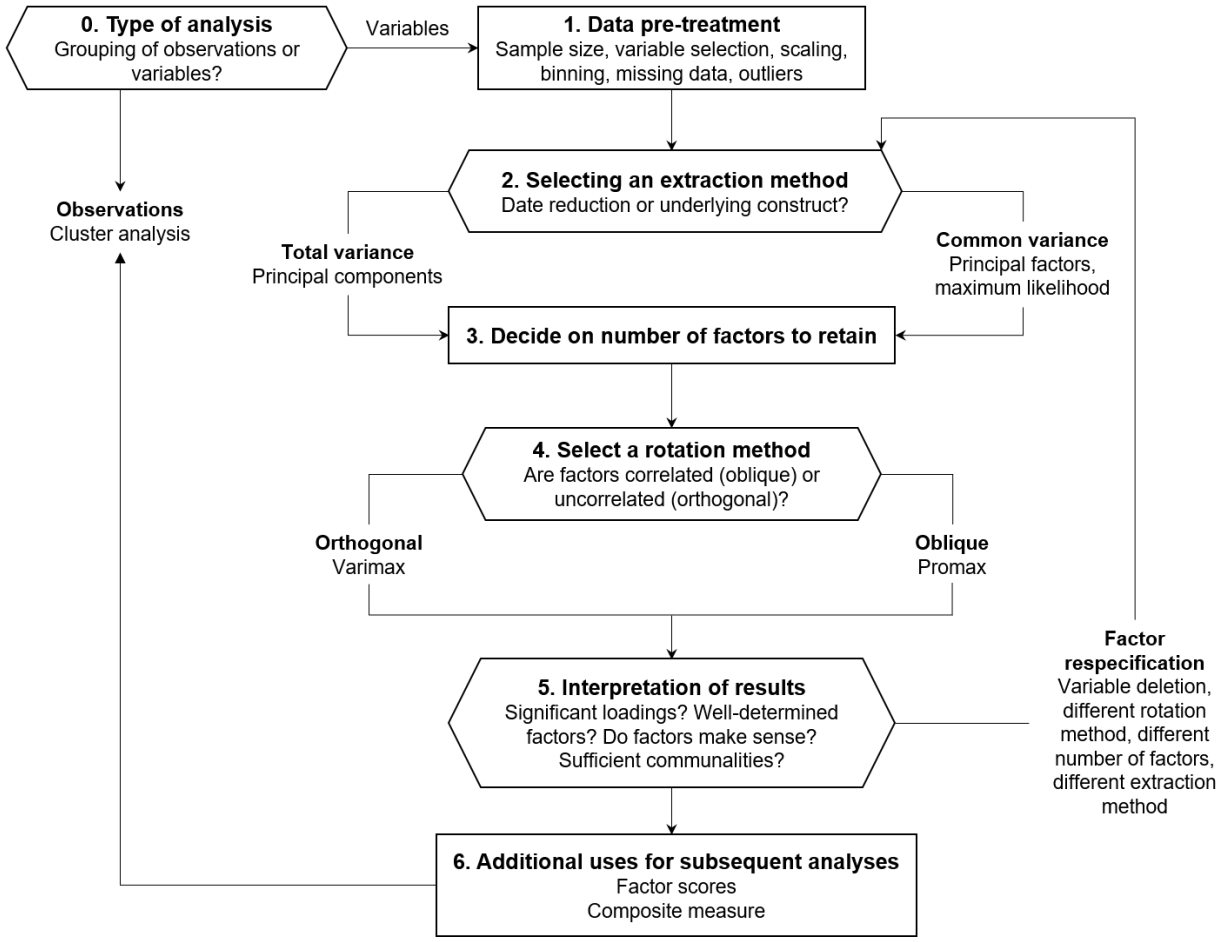


Figure 2-3. Decision process in factor analysis

The figure summarises the decision process used in this thesis to perform the EFA. It was informed by the guidelines described in Chapter 3 of Hair et al, 2010, *Multivariate Data Analysis*, 7th Edition, Pearson Education, Upper Saddle River, NJ

The data was screened for factorability, i.e., the suitability of the data for factor analysis. This is based on the correlation between the variables as well as the sample size. If multiple variables are correlated with an underlying factor, we expect the variables to be correlated with each other. Thus, a degree of correlation is desired. Different measures were used to ascertain factorability in this study.

- **Correlation matrix:**
  As a rule of thumb, no variable should have many correlations under 0.3 [193]
- **Bartlett's test of sphericity:**
  If all variables are perfectly uncorrelated, then the correlation is 0 and the matrix would resemble an identity matrix, i.e., 1 on the diagonal and 0 everywhere else. Bartlett's test tests the correlation matrix against an identity matrix and a significant p-value would indicate that the correlations significantly deviate from the identity matrix.
- **Kaiser-Meyer-Olkin measure of sampling adequacy (KMO):**
  KMO compares correlations to partial correlations and takes a value between 0 and 1, with values closer to 1 indicating better suitability of the data for factor analysis. Values close to 0 indicate that the partial correlations are greater than the correlations, which means that the correlation between the variables is due to only a few variables sharing variance rather than being more widespread.
- **Anti-image correlation matrix:**
  This is the negative of the partial correlation matrix. Many large values indicate that the correlation between some variables is not related to other variables and that the data is not suitable for factor analysis.

An important consideration is the sample size required for exploratory factor analysis. This discussed in more detail in results section 3.1.2. Many different recommendations exist for a minimum sample size in absolute numbers or ratios of observations to variables. While it is widely accepted that larger sample sizes lead to a more stable and reproducible factor pattern, MacCallum, Widaman, and Zhang[194] have demonstrated that the common rules of thumb for a minimum sample size are not generalisable and that the impact of the sample size rests on the communalities (the proportion of shared variance explained by the factors) and factor overdetermination (at least three or four variables strongly correlate with the factor and simple structure). Furthermore, the impact of sample size reduces with increased factor loadings (the strength of the relationship between a variable with a factor)[192].

Although the sample size in this study was small with n = 25, putting thresholds in place can help reduce the impact of a small sample size to still achieve a meaningful result. The thresholds used were 0.5 for communalities and factor loadings and 0.7 for KMO. As will be discussed in the results chapter, a factor with less than three significant loadings was not considered sufficiently overdetermined.

Stage 2. Selecting an extraction method

An extraction method is the technique by which the correlation matrix is examined, and factors are extracted. The factors are linear combinations of the variables and are extracted in a progressive fashion until a number of factors has been extracted that equals the number of original variables. The first factor accounts for the most variance in the data, the second the second highest variance and so on. Each extracted factor has an *eigenvalue*, which is the amount of variance explained by that factor. Each method also estimates *factor loadings* and the *communalities* of the model.

The factor loading represents the unique relationship of the original variable with the extracted factor. Higher loadings therefore represent a bigger contribution of a variable to a factor. The sum of squared factor loadings across all original variables equals the amount of variance explained by the factor.

Communality is the proportion of variance for each variable that is explained by the extracted factors. It is estimated by summing the squared factor loadings across all extracted factors per original variable and ranges from 0 to 1.

The maximum number of factors that can be extracted is the number of original variables. However, the first few factors will account for a substantial portion of the variance. Therefore, the next step is to decide on how many factors to extract.

Stage 3. Deciding on the number of factors to extract

The number of factors retained in the model has implications on factor pattern, the reliability of the structure and interpretability[195]. There a several criteria available to guide the factor retention decision[196]. Two of these, the minimum average partial test (MAP) and scree plot were used in this study and are presented in the results chapter.

In addition, the decision on the number of factors to extract should also be guided on a theoretical basis when possible. In this case, we might expect at least two factors which differentiate at least two metabolic profiles.

Stage 4. Selecting a rotation method

Factor rotation helps achieve a simpler and theoretically more meaningful factor pattern[193]. Each variable is modelled as a linear combination of the extracted factors and the relationship of each variable with each factor is quantified by the factor loading. Ideally, each variable should load only on one factor and each factor should have multiple high-loading variables. This is often not the case with unrotated factor loadings. Because the first factor explains the largest amount of variance, most variables will load highly on the first factor.

To understand the rotation, imagine a factor diagram which plots the unrotated factor loadings of two factors **Figure 2-4**. The variables form two clusters on the diagram, however this is not easily visible from the factor loadings. When factors are rotated, this means that the axes are rotated around the origin so that the factor loading pattern changes, but the relative position of the variables remains the same. This can clarify the factor structure such that each of the clusters is now represented by one factor, i.e., each variable loads highly only on one factor.

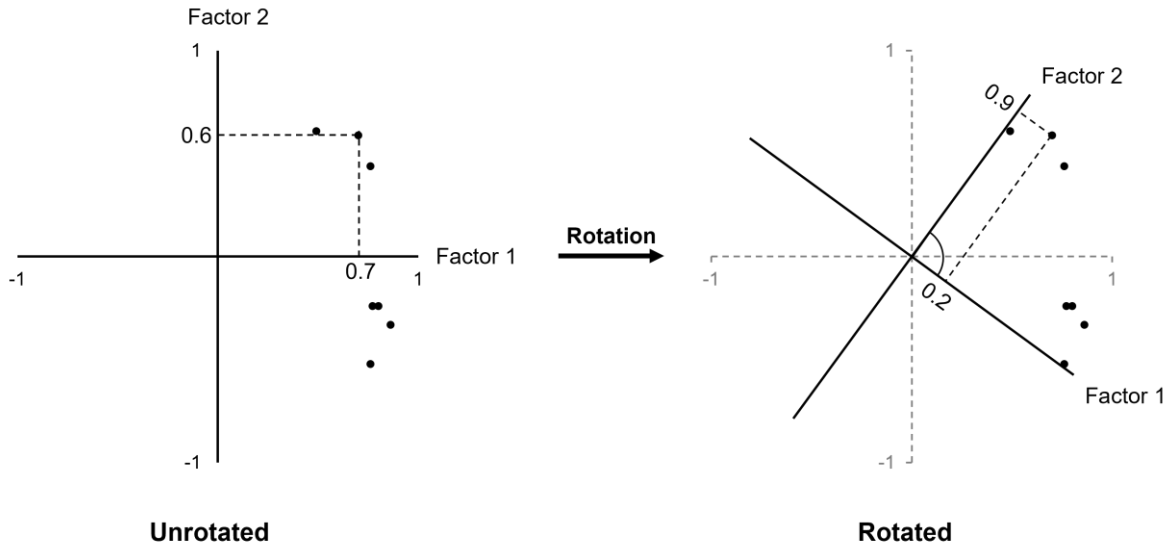


Figure 2-4. Explanation of factor rotation

In essence, factor rotation can be understood as the rotation of the axes around the origin of a coordinate system while the relative position of the variables remains the same. When the angle between the axes remains 90°, as shown in the figure, then the rotation is orthogonal and resulting factors are uncorrelated. When the angle is not constrained to 90°, i.e. the angle may be larger or smaller than 90°, then the rotation is oblique and results in correlated factors.

There are two main categories of rotation: the orthogonal factor rotation and the oblique factor rotation. In the orthogonal rotation, the rotated axes are kept at a 90° angle which results in uncorrelated factors. **Varimax** is an orthogonal rotation method and aims to maximise the variance within a factor so that high loadings are increased and low loadings are reduced. In contrast, **promax** is the oblique rotation method. In oblique rotation, the rotated axes are not constrained to a 90° angle which means that the factors can be correlated after rotation.

The structure simplification can be seen from the example in **Table 2-4**. For this example, two factors were extracted from seven variables. Any loading below 0.5 is omitted for clarity. The unrotated factor pattern is not easy to interpret, as most variables load onto factor 1 with a factor loading of over 0.7. In addition, variables 1 and 3 cross-load onto factor 2, i.e., they significantly load on a second factor. The rotated factor solution however, clearly separates variables 1, 2, 6, and 7 from variables 3, 4, and 5. The former are described by factor 1 and the latter by factor 2. There are no cross-loadings. Essentially, the rotation redistributes the variance explained by each factor, thus the communalities (far right column in Table 2-4) are

the same for both factor patterns. The communalities in this example are all high with a communality of at least 0.6, indicating that at least 60% of the variance of each variable was explained by this factor solution.

Table 2-4. Comparison of rotated and unrotated factor loadings

	Unrotated		Rotated		
Variables	Factor 1	Factor 2	Factor 1	Factor 2	Communality
Var 1	0.76	-0.52	0.92		0.84
Var 2	0.77		0.76		0.66
Var 3	0.70	0.59		0.90	0.85
Var 4		0.61		0.78	0.62
Var 5	0.76			0.81	0.78
Var 6	0.86		0.89		0.85
Var 7	0.80		0.78		0.70

Stage 5. Interpreting the results and respecification

After rotating the factor matrix, the factor matrix should be interpreted and the model respecified if required. Interpretation includes identifying the variables which load significantly on each factor, evaluating the communalities, and labelling the factors. Variables which load highly on a factor are considered to contribute more to the factor. The labelling of a factor involves trying to find an appropriate term descriptive of those variables. This is derived from the nature of the variables rather than a statistical computation.

Respecification of the model is required when a variable does not load significantly on any factor, there is cross-loading, communalities are low, or the resulting factors do not make any sense. In this case, the model may be respecified by extracting a different number of factors or using a different rotation method.

Stage 6. Additional uses for subsequent analyses

After an acceptable solution has been found, appropriate variables representing the factors can be generated and used for subsequent statistical testing. This effectively reduces the dimension of the data to a smaller set of new variables. There are two types of scores, refined and non-refined[197]. A refined score is a linear combination of all variables using the factor loadings as weights. A non-refined score is a composite measure such as the total sum or average of variables which loaded highly on the respective factors. In contrast to the factor scores, each variable was weighted the same.

The new variables can be used for further statistical testing. In this study, they were used for association studies with cardiovascular risk factors and the identification of groups of individuals which were classed as high and low metabolisers. The metabolic profile of these groups was further investigated in Part 2 of the results chapter (Section 3.2).

2.7 My involvement in the CIRCLES study and acknowledgements

The CIRCLES study was designed and conducted by the CIRCLES research team (Chief Investigator Prof Aedin Cassidy, see [159] for other team members and contributions). Dr Vera van der Velpen optimised the LCMS method, supported by Dr Colin Kay. The flow-mediated dilation (FMD) images were captured and analysed by Dr Peter Curtis and Dr Lindsey Berends. Prof Martin Feelisch and his group assessed NO-related metabolites and cGMP levels.

My role in the study included the quantification of metabolites using LCMS in urine and serum, and the assessment (including method optimisation) of inflammatory biomarkers in plasma. For a period of time towards the end of the CIRCLES study and during my first year, I also had some patient facing responsibilities and received and processed biological samples (urine and blood). The data analysis presented in Chapter 3 was designed and performed by me.

3 Retrospective analysis of urinary anthocyanin metabolites to identify a metabolic signature of high metabolisers and its association with cardiometabolic markers

The objective of the study in this chapter was to investigate if differential responses to intake of blueberry anthocyanins among participants of the CIRCLES study existed and if these mediated differential clinical responses, such as vascular function. The CIRCLES study was a six-month blueberry intervention study conducted in people with metabolic syndrome (MetS). MetS is a term which encompasses the compounded metabolic risk, attributable to a cluster of conditions that together significantly increase the risk of developing cardiovascular disease and type 2 diabetes mellitus (T2DM)[10]. MetS is characterised by visceral obesity, dyslipidaemia, hyperglycaemia, and hypertension. While the exact underlying cause of MetS is uncertain, insulin resistance is widely believed to be central to the pathophysiology. In addition, each of the components of MetS can cause endothelial dysfunction[46]. There is evidence from population-based studies and human intervention trials that dietary flavonoids, and in particular the subclass of anthocyanins, may improve cardiometabolic health and alleviate several components of the metabolic profile associated with MetS[70,87,89,99,198]. Thus, for those with MetS a daily intake of anthocyanins may improve the MetS metabolic profile as a whole and complement existing therapies to address MetS.

Anthocyanins are found in red-, blue-, and purple-coloured fruits and vegetables and are especially abundant in commonly consumed berry fruits including blueberries.[199] In population-based studies, the habitual intake of anthocyanins has been shown to be the most effective of the six main flavonoid subclasses in reducing the risk of cardiovascular diseases (CVD), such as myocardial infarction, ischemic stroke, and hypertension, and is inversely associated with insulin resistance, inflammation, and body weight[59,68,69,99,108,200,201]. Beneficial effects on intermediate markers of cardiovascular risk, including endothelial function, arterial stiffness, blood pressure, insulin resistance and lipid profile, have also been observed in randomised controlled trials (RCTs)[84–91].

However, findings from RCTs are equivocal and there is a well-established body of research which highlights the inter-individual variability in response to flavonoid interventions and the importance of addressing this issue in future research[151,156,202–208]. The inter-individual variability in the absorption and metabolisation of flavonoids implies that individuals are exposed to different levels of bioactive compounds. For example, inter-individual variability is demonstrated by a study with 18 participants which assessed the flavanone metabolites

hesperitin and naringenin (aglycones and phase II conjugated metabolites) in urine after consumption of 400 ml fresh orange juice[209]. In this study variable mean excretion rates were observed for high excretors (15% of intake, n = 4), medium excretors (9% of intake, n = 7), and low excretors (3% of intake, n = 7). In another example, a human tracer study which fed 500 mg ^{13}C-labelled Cyanidin-3-glucoside recovered on average a total of 44% ^{13}C in urine, breath, and faeces over 48h. However, per individual, the recovery ranged from 15% to 99%[140]. The authors note that the observed variability was likely due to unrecovered label remaining in faeces, demonstrating a variability in gastric and intestinal transit time, which may also affect bioavailability over time due to prolonged time in the gut. The between-person variability in bioavailability is thought to be one of the factors leading to contradictory results obtained from clinical trials.

Importantly, the tracer study outlined above observed a high abundance of ^{13}C-labelled phenolic metabolites, some of which persisted over 48 h[123]. In contrast, the parent anthocyanin cyanidin-3-glucoside was only observed for a short period of time and made up only 2% of the total compounds found in circulation. This result emphasised the potential of the anthocyanin metabolites as the bioactive mediators of clinical effects rather than the parent compounds. In support of this, phenolic metabolites have been found to be effective in *in vitro* studies. For example, in a cell model a greater anti-inflammatory effect was observed for anthocyanin metabolites in comparison to their parent anthocyanins.[141–143]

There are a multitude of factors which likely contribute to inter-individual variation, including genetic variability in the absorption of phenolic compounds from the gut, further metabolism once absorbed (including the phase I and II of the xenobiotic metabolism), the elimination through urine excretion or recirculation to the gut via bile. Next to the genetic component, particularly the gut microbiota are thought to play a major role in the metabolism of flavonoids.[128,151]. Up to 85% of anthocyanins reach the colon intact where they are exposed to microbial degradation. Thus, the composition and bioactivity of the gut microbiota may significantly contribute to inter-individual variation. On this basis, several phenotypes have been established which differentiate a responder from a non-responder to flavonoid intake. A prominent example of such a phenotype is the ability to produce the microbial metabolite equol from soy isoflavones[152] which depends on the presence of certain strains of bacteria. Equol production has been linked to vasoactive functions[153,210], however, only about 30% of the western population are equol producers. This highlights the importance of being able to differentiate between metabolic phenotypes (‘metabotypes’) when assessing the clinical efficacy of flavonoids. Unfortunately, only few unique, metabotype-defining catabolites for flavonoids such as equol are known (another example are urolithins from ellagitannins[156]). Predominantly, microbiota degrade many different flavonoids into a set of similar catabolites [157] which include hydroxycinnamic, phenylpropionic, phenylacetic and benzoic acids. Because

of this, metabotyping on the basis of produced metabolites in response to dietary interventions is not straightforward for the majority of flavonoids and likely includes a group of metabolites rather than a specific metabolite.

The main aim of the current study was to identify a metabolic signature of responders to blueberry anthocyanins. The second aim was to investigate the relationship between the metabolic signature and vascular function. These aims were driven by the underlying hypothesis that a single or a group of metabolites may mediate the potential cardiovascular benefits a person receives from blueberry intake and that this group of metabolites is preferentially produced by individuals of a certain metabolic phenotype. As has been shown for other dietary flavonoids (e.g. isoflavones[152,211] and ellagitannins[156]), underpinning metabolic phenotypes may explain some of the variability in cardiometabolic response to intake.

The data analysis was approached in two steps: first, an exploratory factor analysis was performed to identify clusters of urinary metabolites which may suggest a metabolic phenotype. Second, the metabolic profile of high metabolisers was examined more in-depth to create a panel of metabolites indicative of a high metaboliser of blueberry compounds.

The data presented hereafter, is the first known assessment (to the author) of a metabolic phenotype from an expansive range of anthocyanin-derived metabolites on cardiometabolic endpoints.

3.1 Part 1: Exploratory factor analysis of urinary anthocyanin metabolites following blueberry intake for six months

3.1.1 Summary of the method and data used (see methods Section 2.6 for details) and statistical approach

The factor analysis and subsequent testing in this section was restricted to n = 25 participants (after quality assurance-based exclusions) from the full dose group of the CIRCLES study (see Section 2.1). To recap, the full-dose group consumed freeze-dried blueberries daily for six months providing 364 mg anthocyanins and 879 mg phenolics. The 24 h urinary analysis presented hereafter was limited to n = 46 blueberry metabolites quantified using liquid-chromatography mass spectrometry (see Section 2.4 for methodology). The data from a range of cardiometabolic assessments were included in this section (further described in Sections 2.2 and 2.3, i.e., FMD, PWV,SBP, DBP, fasting levels of glucose, insulin, HOMA-IR, HDLC, LDLC, TG and the inflammatory markers IL-6, TNF-α, CRP, IL-8, adiponectin. As described in the statistics section (Section 2.5), exploratory factor analysis (EFA) was used to identify clusters of metabolites; associations were assessed using Pearson correlation, group

comparisons were done using Welch's t-test and cluster analysis was performed using k-means clustering.

3.1.2 Screening and preparation of the metabolite data

After data pre-processing, which included the analysis of the LCMS output in Analyst 1.5 and further data handling in Excel and SAS, the data was pre-treated prior to the multivariate analysis. This included data cleaning, identification of potential outliers, scaling, and variable selection. Data pre-treatment affects the outcome of the analysis and is crucial for improving the biological interpretation of the results[191,192]. The challenges encountered with this dataset and how these were resolved is summarised in **Table 3-1** and described in more detail below.

Table 3-1. Steps involved in the preparation of the metabolite data for EFA

Challenge	Solution
1. No. of variables > sample size (n = 25)	Reduce no. of variables (group by chemical structure) and set thresholds: - factor loadings (> 0.5) - communalities (> 0.5) - KMO (> 0.7)
2. Variation at baseline	Adjust for baseline (calculate change)
3. Differences in scale/large range	Median and IQR normalization
4. Potential outliers	Median and IQR normalization
5. Missing values	Imputation of median
6. Factorability	Grouping by chemical structure attenuated the extreme multicollinearity of some variables

Challenge 1: Sample size and number of variables

This analysis included n = 25 observations from the original n = 37 full dose participants that completed the study. Participants were excluded from this factor analysis based on having poorer quality FMD data, incomplete urine collections or lacking metabolite data.

Although there is no generally agreed recommendation for an adequate EFA sample size, it is widely accepted that more stable, reproducible and generalisable factor patterns are observed with larger sample sizes[192]. Equally, the ratio of observations to variables assessed is important; with greater stability when the observations comfortably outnumber the variables assessed. Guidelines in textbooks for a minimum sample size vary largely, for example Hair et al[193] recommended a sample size of at least 50 and a ratio of 5 observations per variable, but added that 100 samples and a ratio of 10:1 would be preferable. Comrey and Lee[212] offered a rough scale ranging from n = 50 (very poor) to 100 (poor), 200 (fair), 300 (good), 500 (very good) and 1000 (excellent). In contrast, Gorsuch[213] considered a ratio of 10:1 above what was

needed and speaks of a factor analysis where a sample size of less than twice the number of variables gave the same factors as a larger sample size. MacCallum, Widaman, Zhang and Hong[194] demonstrated empirically that the common rules of thumb for a minimum sample size are not generalisable nor particularly useful in factor analysis and showed that the required sample size is related to the size of the communalities (i.e. the shared variance between variables) and number of variables per extracted factor. Furthermore, strong factor loadings have been shown to be good indicators of a stable factor pattern and the impact of sample size decreases with increasing factor loadings[192]. In sum, the impact of sample size decreases with high communalities, well-determined factors and strong factor loadings.

The data in the present dataset therefore had around half of what could be perceived as the minimal expected observations, whilst the number of metabolites of interest (n = 46) gave a ratio of about 1:2 outside the often-recommended boundaries. To account for these limitations, the following steps were taken to maximize the chance to obtain biologically meaningful results:

- The 46 quantified metabolites were grouped into 15 groups based on chemical structure (see **Table 3-2**). Each group was created by adding up the absolute change of individual metabolites between baseline and endpoint to a total sum.
- Benzoic acid and hydroxybenzoic acids were separated into two separate groups. These groupings were considered to better reflect the kinetics of metabolite formation where hydroxybenzoic acids can appear in the circulation as one of the first degradation products, whereas benzoic acid is usually considered as a later stage metabolite of microbial origin which is eliminated from the system as hippuric acid upon absorption[123].
- Thresholds were set at 0.5 for factor loadings and communalities and at 0.7 for the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy.

The groups shown in Table 3-2 were used as variables in the factor analysis.

Table 3-2. The 15 metabolite groups used as variables in EFA

Group number	Group	Metabolite
1	Methoxybenzoic acid	2-hydroxy-4-methoxybenzoic acid 3-hydroxy-4-methoxybenzoic acid 4-hydroxy-3,5-dimethoxybenzoic acid 4-hydroxy-3-methoxybenzoic acid 3,4-dihydroxy-5-methoxybenzoic acid
2	Methyl benzoates	methyl-3,5-dimethoxybenzoic acid methyl-3,5-dihydroxybenzoic acid methyl-4-hydroxybenzoic acid
3	Hydroxycinnamic acids	3,4-dihydroxycinnamic acid 4-hydroxycinnamic acid trans-3-hydroxycinnamic acid

Table 3-2. Continued

Group number	Group	Metabolite
4	Methoxycinnamic acids	4-hydroxy-3,5-dimethoxycinnamic acid 4-hydroxy-3-methoxycinnamic acid
5	Cinnamic acid esters	Chlorogenic acid Rosmarinic acid
6	Hydroxyphenylpropionic acids	3,4-dihydroxyphenylpropionic acid 3,5-dihydroxyphenylpropionic acid 3-hydroxyphenylpropionic acid 4-hydroxyphenylpropionic acid
7	Methoxyphenylpropionic acids	4-methoxyphenylpropionic acid 4-hydroxy-3-methoxyphenylpropionic acid
8	Hydroxyphenylacetic acids	3,4-dihydroxyphenylacetic acid
9	Methoxyphenylacetic acids	3-hydroxy-4-methoxyphenylacetic acid 4-hydroxy-3-methoxyphenylacetic acid
10	Sulfate conjugates	3&4-methoxybenzoic acid-3&4-sulf 3&4-hydroxybenzoic acid-3&4-sulf Benzoic acid-4-sulfate
11	Hippuric acids	3&4-methylhippuric acid 3-hydroxyhippuric acid 4-hydroxyhippuric acid Hippuric acid
12	Glucuronide conjugates	3-methoxybenzoic acid-4-GlcA 3-hydroxybenzoic acid-4-GlcA 4-methoxybenzoic acid-3-GlcA 4-hydroxybenzoic acid-3-GlcA Benzoic acid-4-GlcA
13	Glutamate conjugates	Benzoylglutamic acid
14	Benzoic acid	Benzoic acid
15	Hydroxybenzoic acids	2,3-dihydroxybenzoic acid 2,5-dihydroxybenzoic acid 2-hydroxybenzoic acid 3,4,5-trihydroxybenzoic acid 3,4-dihydroxybenzoic acid 3,5-dihydroxybenzoic acid 3-hydroxybenzoic acid 4-hydroxybenzoic acid

Challenge 2: Baseline variation

Although dietary restrictions were designed to limit the intake of anthocyanins and common sources of flavonoids as a means of standardising metabolite backgrounds, there was still some variation in metabolite values at baseline. It is highly likely that other food sources contributed to the measured metabolites, as many of the metabolites listed in Table 1-2 are attainable from other sources including many commonly consumed fruits and vegetables, coffee and tea[64]. To adjust for the variation in background diet, a calculation was made to

deduct the baseline from the endpoint value; this 'change from baseline' value was used for the factor analysis.

Challenge 3: Handling of differences in range

The metabolites varied greatly in their abundance and in the magnitude of absolute change from baseline to endpoint. For example, the change in hippuric acids ranged from -2797.6 to 7458.7 µmol/24h while the change in hydroxybenzoic acids ranged from -21.4 to 24.8 µmol/24h. The data was centred by subtracting the median and scaled by dividing by the interquartile range (IQR) to accommodate the offset between high and low abundant metabolites and to adjust for variations in absolute change. Median and IQR were chosen instead of the mean and standard deviation to minimize the effect of potential outliers.

Challenge 4: Dealing with outliers

Values which could be considered as outliers were checked for experimental or measurement errors. In the absence of such errors, these values were intentionally not excluded from the analysis due to exploratory aim of the study to find differences in how people metabolise blueberries.

Challenge 5: Handling of missing values

In circumstances of missing data, data was imputed using the median values of the full-dose treatment group. In the dataset used in this chapter, 'change' (0 to 6 month) data was partially missing for three metabolites: 3-hydroxy-4-methoxybenzoic acid (n = 4 participants (14.81 %)), benzoic acid-4-sulfate (n = 8 participants (29.63%)), and 3,4-dihydroxyphenylacetic acid (n = 1 participants (3.7%)).

When assessing the completeness of the metabolite data, it was apparent that the missing data occurred at random and was due to incomplete quantification using LCMS. Because a substantial proportion of data was missing these metabolites, a complete-case analysis was not a viable option. Instead, for the factor analysis presented in this section, the missing data was imputed using the median values of the full-dose treatment group. While imputing the median is known to reduce the variance of the data[214], the effect should be largely attenuated because the metabolites were part of a bigger group of metabolites. This was tested by excluding benzoic acid-4-sulfate and 3,4-dihydroxyphenylacetic acid from the analysis (while preserving the sample size). Because the exclusion did not change the overall results, the median imputation was favoured for its simplicity over other generally more accepted methods such as multiple imputation.

Challenge 6: Testing the factorability of the data

Factor analysis assumes that an underlying structure exists in the set of observed variables. It tries to identify a pattern among the correlation between the variables and describe them with a smaller set of factors. To this extent, a degree of correlation between the variables is required. Because this is a key requirement for a successful factor analysis, the interrelationship between variables and thereby the factorability of the data was tested using four different methods.

First, the correlation matrix was inspected for a substantial number of correlations above 0.3 (as recommended by Hair et al[193]). For each of the n = 15 metabolite groups (see Table 3-2), a correlation with multiple other groups above 0.3 was observed (**Table 3-3**). In most cases, correlation coefficients were between 0.3 and 0.7; with some highly correlated (>0.8), such as the hydroxy- and methoxycinnamic acids (groups 3 and 4) and hydroxy- and methoxybenzoic acids (groups 1 and 15). The high multicollinearity between variables is not generally a problem in factor analysis unless variables are perfectly correlated and therefore redundant.

Table 3-3. Pairwise correlations of the metabolite groups

Group	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1.00														
2	0.37	1.00													
3	0.86	0.47	1.00												
4	0.81	0.25	0.83	1.00											
5	0.50	0.09	0.62	0.58	1.00										
6	0.46	0.36	0.49	0.21	0.38	1.00									
7	0.45	0.38	0.30	0.19	0.09	0.55	1.00								
8	0.16	-0.15	0.18	0.09	0.18	0.41	0.38	1.00							
9	0.54	0.31	0.50	0.33	0.31	0.68	0.64	0.32	1.00						
10	0.81	0.18	0.66	0.65	0.48	0.61	0.58	0.39	0.77	1.00					
11	0.47	0.30	0.49	0.27	0.61	0.78	0.45	0.27	0.44	0.49	1.00				
12	0.92	0.30	0.85	0.82	0.51	0.40	0.45	0.32	0.54	0.83	0.39	1.00			
13	0.80	0.21	0.75	0.75	0.74	0.36	0.21	0.05	0.37	0.66	0.55	0.74	1.00		
14	0.46	0.73	0.55	0.18	0.15	0.57	0.52	0.16	0.62	0.43	0.43	0.47	0.21	1.00	
15	0.87	0.42	0.86	0.79	0.46	0.48	0.35	0.21	0.58	0.79	0.34	0.90	0.68	0.49	1.00

See table Table 3-2 for detail on metabolite groupings.

To further confirm the suitability of the data for factor analysis, three statistical measures were assessed. The Kaiser-Meyer-Olkin measure of sampling adequacy, which estimates the proportion of common variance among the variables, was 0.80 (above the commonly recommended value of 0.5) and Bartlett's test of sphericity, which tests if the variables are uncorrelated, was significant ($\chi^2 = 844.0$, $p < 0.001$), i.e., the variables significantly diverged from being uncorrelated. Finally, the anti-image describes the portion of variance that cannot be explained by the other variables. Because the variables should be correlated, the anti-image

should be small. None of the anti-image correlation coefficients were high enough to indicate any problems.

Taken together, all four tests suggested reasonable factorability.

After pre-treating the data and confirming its suitability for factor analysis, a series of factor analyses was performed to find the most optimal factor model for this dataset.

3.1.3 Results of the initial factor analysis

Choosing an extraction method

The initial factor analysis was performed with 15 variables (metabolite groups). As mentioned previously, factor analysis extracts a number of factors based on the interrelatedness of the variables. The greater the correlation between variables, the greater the shared variance. Factor analysis proceeds by extracting the first factor which accounts for the most variance in the data. The second factor is then extracted from the variance that is still unexplained. Subsequent factors are progressively extracted from the remaining variance accounting for less and less of variance until all variance is explained. Although up to *n* factors can be extracted from a dataset with *n* variables, the goal of factor analysis is to explain a great amount of the variance with only a few factors and thus to represent the entire dataset with a much smaller set of variables.

The extraction method used in this analysis was *principal factors*. Stata offers several methods of extraction: principal components, principal factors, and maximum likelihood. Principal components (the extraction method used in principal components analysis) assumes that the common variance among the variables (communality) is equal to the total variance (i.e., there is no error in measurement) and is useful when dimension reduction of the data is the main goal. In contrast, principal factors (used here) assumes that the total variance is comprised of a common variance and a unique variance that is not common. Factors resulting from this method are based only on the common variance, which makes this method better suited to identify an underlying latent variable and thus was more appropriate for this analysis. The maximum likelihood factor method also partitions the total variance into common and unique, however assumes multivariate normality which was not the case here.

Using the principal factor method for extraction, nearly all of the variance was explained by the first five extracted factors (98.1%). This meant that the 15 variables could almost entirely be represented by five factors.

Tests to determine the number of factors to extract

The next step was to determine the number of factors to retain for further analysis. To avoid loss or dilution of information and uninterpretable models, it is critical to ensure that neither too few or too many factors are retained[195]. Therefore, two test criteria were considered to help with the decision process, including the recommended minimum average partial (MAP) test and the scree test as an adjunct criteria[196].

Minimum average partial (MAP) test: The MAP test sequentially partials out the common variance for each successive number of factors[215]. This test has been shown to have good accuracy and is considered superior to the other two tests.[192,196] The MAP criterion will decrease until the point at which only unique variance remains and then start to increase. With this criterion, only factors consisting of common variance are retained. In this case, the MAP test indicated the retention of two factors (**Figure 3-1A**).

Scree test findings: The scree test is one of the most commonly used criteria to determine the number of factors. It plots the eigenvalues of the factors[216]. The point at which the curve breaks or levels out, indicates the number of factors to retain. A disadvantage of this test is the subjective judgement on the number of factors when there are multiple breaks in the curve or the bend is unclear. In this case, Costello and Osborne[217] suggest to test multiple factor solutions retaining the predicted number of factors given by the scree test as well as one above and below. Velicer et al[196] also point out the advantage of the visual representation, as it has the potential to avoid overextraction because minor factors will appear as insignificant. The authors therefore value its use as an adjunct method to the MAP test.

As demonstrated in **Figure 3-1B**, the scree plot did not definitively confirm the point at which factors should be included, as it displayed two breaks. At these breakpoints, the scree test suggested the retention of either two or five factors. The plot also clearly shows that factors six and onwards were trivial factors because they accounted for insignificant amounts of variance.

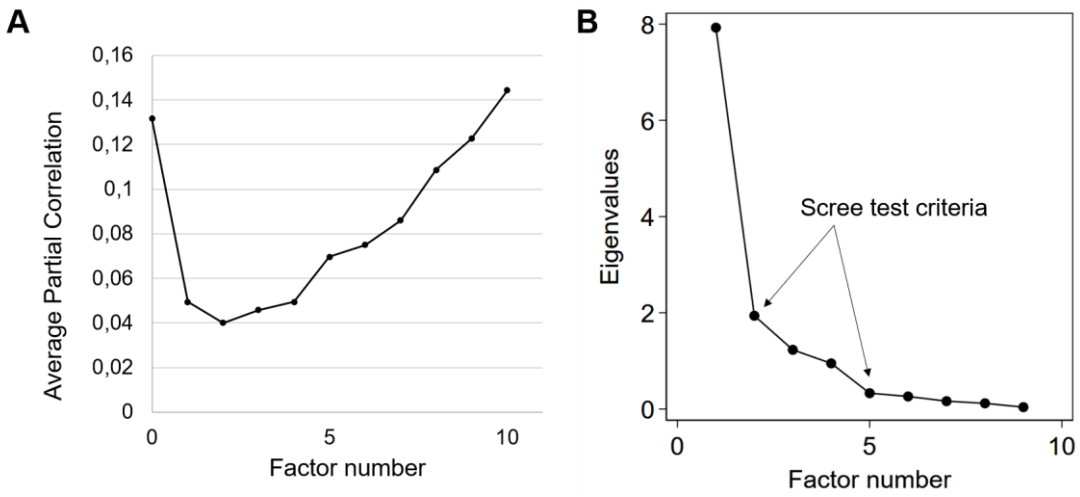


Figure 3-1 MAP and scree test plots.

The minimum average partial (MAP) correlation and the scree test were used to determine the number of factors to retain.
(A) The MAP correlation suggested two factors for this data (the number of factors where the average partial correlation was lowest).
(B) The scree test identified two bends in the curve, suggesting to retain between two and five factors. It also showed that factors six and beyond accounted for insignificant amounts of variance.

Comparison of factor solutions with two, three, four, and five extracted factors.

The MAP and scree tests indicated the retention of two or five factors. To ensure that the best structure was identified, a number of factor solutions was tested and the outcomes of retaining two, three, four or five factors were explored.

An important process in determining the factor solution which gives the best structure involves the ability to make sense of a factor, i.e., to give it a name and determine how the high loading variables of a factor fit together. In the absence of a sensible interpretation, respecifying the model and for example extracting a different number of factors can help achieve sensible results[192].

Each factor solution was examined using two types of factor rotation methods, an orthogonal rotation (varimax) and an oblique rotation (promax). An explanation of factor rotation is given in the methods section (see Section 2.6, stage 4). In brief, factor rotation helps achieve a simpler and theoretically more meaningful factor pattern[193]. A simple factor pattern is characterised by each variable only loading on one factor and each factor should have multiple high-loading variables. After an orthogonal rotation, the factors are uncorrelated, while an oblique rotation allows the factors to be correlated.

Due to the connection through metabolic pathways, we can expect some degree of correlation between the factors. Therefore, using an oblique rotation method was likely more appropriate

than supposing factors were uncorrelated, as assumed in the varimax rotation method. Both were tested to find out which one resulted in a simpler and interpretable factor structure.

Five-factor solution: The five-factor solution resulted in an unclear factor pattern with several variables cross-loaded onto multiple factors for both types of rotation. In addition, two factors were not well determined with only one and two significant loadings, respectively. With small sample sizes, as in the present dataset, the likelihood of improper solutions increases if factors are not well determined[194]. The determination of a factor relates to the number of variables which load highly on a factor; thus, two factors in this solution were not well determined with only one and two significant loadings. In addition to the possibility of an instable factor pattern, due to the cross-loading the pattern structure was unclear. Consequently, this model was discarded because likely too many factors had been extracted, resulting in cross-loading and insufficient significant loadings which made interpretation of the underlying variable not possible.

Four-factor solution: The four-factor solution resulted in a similar outcome as the five-factor solution, with cross-loading of variables, one factor with insufficient loadings and an unclear factor pattern with both types of rotation. This model was also discarded for the same reasons as the five-factor model.

Three-factor solution: The factor pattern was clearer when extracting three factors, however still had one factor with only two significant loadings. The variables in question were the methyl benzoates and benzoic acid. Methyl benzoates have been identified previously in urine following anthocyanin intake[123], and benzoic acid is a known microbial metabolite of phenols[128]. Apart from both groups being phenolic metabolites and containing benzoic acid, it was difficult to identify an underlying latent variable which might affect these two metabolite groups, but not other metabolites such as hydroxy- and methoxybenzoic acids as well as benzoic acid sulfates. It is possible that a third factor was present, but that the grouping of metabolites was obscuring some information. This resulted in an unclearly defined factor and structure.

Two-factor solution: The two-factor solution had two clearly defined factors with overall good communalities of ≥ 0.7 for most variables (i.e., > 50% of the variable was explained by this factor solution). It appeared that both benzoic acid and the methyl benzoates now loaded onto the second factor. However, the methyl benzoates had a low communality of 0.3, meaning in comparison to other metabolites it was underrepresented in this model. Extracting only two factors likely caused factors 2 and 3 to merge and potentially distort the loading of other variables on the second factor.

Because the methyl benzoates were underrepresented in the two-factor model, but a three-factor solution was not optimal due to the undetermined third factor, a further model was tested with *extracting two factors, but excluding the methyl benzoates*. This model resulted in the clearest loading pattern and was chosen for further analysis accepting that the methyl benzoates were not represented in this model. This factor solution is described in the following section.

3.1.4 Factor analysis with two retained factors

The following table shows the results for the two-factor solution with 14 variables for both varimax and promax rotations (**Table 3-4**). Cumulatively, the two factors explained a high amount of the variance (82.9%). Factor loadings were frequently > 0.7 and communalities for most variables were > 0.5. Both factors were well determined with 6 and 8 variables loading on the respective factors. These parameters gave reassurance this factor model was a valid solution despite the small sample size. As mentioned above, high communalities, well-determined factors, and strong factor loadings reduce the impact of a small sample size and increase the likelihood of finding a stable solution[194].

Both rotation methods resulted in similar factor patterns, however the promax rotation provided a clearer structure. This confirmed that the factors were correlated due to shared metabolic pathways. The correlation between the factors was r = 0.50. In addition, the promax rotation also clarified the primary loadings of the sulfate conjugates and hydroxyphenylacetic acids at a factor loading threshold of 0.5.

One variable (hydroxyphenylacetic acids) displayed low communality: only 23% of the variance was explained by this factor solution. The metabolite in question was 3,4-dihydroxyphenylacetic acid, also known as DOPAC, a direct metabolite of dopamine. The endogenous source is a likely reason why the factor model, which was comprised of phenolic metabolites mostly derived from the blueberry intervention, explained such a low amount of the variance. Nevertheless, DOPAC significantly increased in participants who received the blueberry intervention in comparison to the placebo group, suggesting that at least part of the measured metabolite was in response to blueberry intake[159]. Furthermore, DOPAC loaded clearly onto factor 2 and only very weakly on factor 1. For these reasons, DOPAC was included in the model and considered as a part of factor 2 despite the low communality.

Table 3-4 Factor pattern matrix for urinary anthocyanin metabolites.

Variable	Unrotated factor loadings		Rotated factor loadings (varimax)		Rotated factor loadings (promax)		Communality
	1	2	1	2	1	2	
Methoxycinnamic acids	0.766	-0.516	**0.924**	-0.007	**1.018**	-0.232	0.854
Methoxybenzoic acid	0.915	-0.218	**0.883**	0.324	**0.856**	0.152	0.885
Hydroxycinnamic acids	0.893	-0.244	**0.879**	0.289	**0.863**	0.114	0.856
Glucuronide conjugates	0.912	-0.212	**0.877**	0.327	**0.848**	0.157	0.877
Glutamate conjugate	0.787	-0.378	**0.864**	0.120	**0.907**	-0.074	0.761
Hydroxybenzoic acids	0.882	-0.194	**0.842**	0.325	**0.810**	0.163	0.815
Cinnamic acid ester	0.630	-0.239	**0.657**	0.149	**0.669**	0.009	0.454
Sulfate conjugates	0.889	0.103	**0.684**	**0.577**	**0.547**	0.485	0.800
Hippuric acids	0.639	0.325	0.353	**0.624**	0.167	**0.618**	0.514
Methoxyphenylacetic acid	0.708	0.439	0.348	**0.757**	0.114	**0.770**	0.693
Hydroxyphenylpropionic acids	0.673	0.557	0.253	**0.836**	-0.018	**0.882**	0.762
Benzoic acid	0.576	0.428	0.244	**0.675**	0.029	**0.703**	0.515
Methoxyphenylpropionic acids	0.539	0.517	0.164	**0.729**	-0.078	**0.783**	0.559
Hydroxyphenylacetic acid	0.325	0.354	0.076	0.475	-0.085	**0.517**	0.231

Factor correlation (promax rotation) = 0.50

The factor analysis revealed a grouping of metabolites by metabolism stage

The factor analysis revealed two clusters of metabolites (**Figure 3-2**) as follows:

- Factor 1: cinnamic acids, hydroxy- and methoxybenzoic acids (not benzoic acid), glucuronide conjugates, glutamate conjugates, and sulfate conjugates.
- Factor 2: phenylpropionic acids, phenylacetic acids, hippuric acids and benzoic acid.

On the basis of known metabolic pathways of anthocyanins and other flavonoids, these two metabolite groups were named ‘early-stage metabolites’ (factor 1) and ‘late-stage metabolites’ (factor 2). The stages are used here to refer to the likely absorption location and time. ‘Early-stage’ refers to metabolites likely absorbed in the small intestine. This is generally indicated by the appearance of a compound in plasma within an hour after ingestion, while an appearance after 5 h is considered to be mainly from the large intestine[128]. A great majority of anthocyanins reaches the colon[130], where they are exposed to degradation by the gut microflora. Thus ‘late-stage’ metabolites are likely to consist of phenolic acids originated from the microbial metabolism and to be more abundant than early-stage metabolites. Following absorption, phenolic compounds are treated as xenobiotics and undergo phase I and II biotransformation either in intestinal cells or the liver[124]. Phase I involves oxidation, reduction, and hydrolysis,

whereas phase II reactions are conjugation reactions which include glucuronidation, sulphation, methylation, and glycine conjugation. The phenolic compounds are then excreted via urine (main pathway for small conjugates) or bile (large, extensively conjugated metabolites) [64].

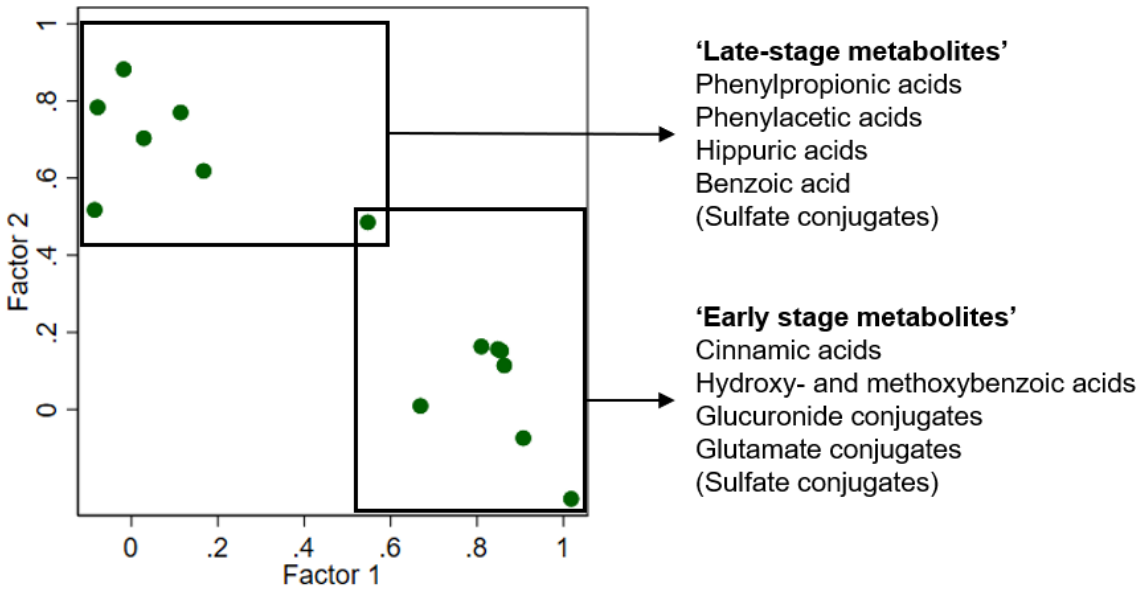


Figure 3-2. Loading plot using promax rotation

Loading plot of the two-factor solution shows two variable clusters of 'late-stage' and 'early-stage' metabolites. One variable (sulfate conjugates) cross-loaded onto both factors.
Early and late are used here to indicate the likely absorption location and time, with 'early' referring to the small intestine and 'late' to the colon. 'Early-stage' metabolites include compounds which have been observed to appear in plasma and urine within the first few hours after ingestion of berries or anthocyanins. 'Late-stage' metabolites include metabolites which are typically products of microbial transformations in the colon.

Early-stage metabolites: Hydroxycinnamic acids (such as ferulic and caffeic acid in free, esterified or glycosidic form) and smaller amounts of hydroxybenzoic acids (such as protocatechuic acid) occur naturally in blueberries[64]. They have been observed to appear in plasma and urine within the first few hours after ingestion of berries or anthocyanins[122,123,218,219]. In addition, hydroxybenzoic acids, such as gallic acid, syringic acid, protocatechuic acid, and vanillic acid, are formed as the initial metabolites from spontaneous chemical degradation in the small intestine[123,128]. Thus, protocatechuic acid and vanillic acid have been observed to peak early in urine between 1 – 4 hours after ingestion. Once absorbed, phenolic compounds undergo phase II metabolism and a corresponding peak for sulfated and glucuronidated protocatechuic and vanillic acid is observed in the first few hours. Ferulic acid could also arise from the 3-*O*-methylation of caffeic acid following hydrolysis of a small proportion of chlorogenic acid (ester of caffeic acid and quinic acid) through mammalian esterases[157].

Late-stage metabolites: The colon is the main site of metabolisation and absorption of anthocyanins[157]. The primary microbial catabolites of anthocyanins are phenylpropionic, phenylacetic and benzoic acids[128]. Microbial transformations in the gut include hydrolysis, ring fission, hydrogenation, demethylation, dehydroxylation, and decarboxylation. Phenylpropionic acids contain a sidechain of three carbons which are progressively shortened to benzoic acids, which are then glycinated upon absorption to form hippuric acids and excreted via urine. Hippuric acid could thus also be seen as a microbial metabolite, although the glycination is not a microbial transformation. The conversion to benzoic acid and then hippuric acid has been speculated to be the definitive elimination pathway of polyphenols[123]

The categorisation into early and late needs to be qualified in the sense that multiple phenolic acids appear in a biphasic pattern after ingestion of anthocyanins. For example, after the initial peak of vanillic acid, a second peak has been repeatedly observed at 24h after intake in different studies[122,123]. A similar pattern was observed for ferulic acid. The second peak is thought to be a result of metabolisation at different sites of the gut as well as enterohepatic recycling, i.e. phenolic compounds which have been absorbed are excreted via bile back into the gastrointestinal tract where they are exposed to the microbial metabolism once more and can be absorbed again[124]. Moreover, the majority of compounds such as chlorogenic acid are hydrolysed in the colon rather than in the small intestine[220]. The free caffeic acid is then rapidly absorbed and metabolised, as evidenced by the fact that it is frequently not observed in plasma or urine following anthocyanin intake[221]. The 3-O-methylation of caffeic acid, this time absorbed from the colon, could explain the second increase of ferulic acid. Furthermore, a chain of microbial reactions in the colon converts hydroxycinnamic acids, which fall under factor 1 (early-stage) to hydroxyphenylpropionic acids (factor 2 metabolites) through microbial hydrogenation, which are in turn progressively converted to hydroxybenzoic acids (factor 1 metabolites) via α- or β-oxidation[128]. It is clear therefore, that the factors are connected through metabolic pathways and their correlation is not surprising.

While some microbial degradation also occurs in the small intestine, the main site of microbial transformation is in the colon[157]. However, endogenous production and other dietary sources add difficulty in ascertaining the exact origin of the small phenolic acids. For example, in the present study, hippuric acid and homovanillic acid were two metabolites, which were already present at baseline, and loaded on factor 2 (presumed to be late-stage/colonic metabolites of the consumed blueberry anthocyanins). Hippuric acid can be derived from many aromatic compounds, including aromatic amino acids from protein catabolism, and quinic acid, which occurs in many fruits and vegetables[222]. Its daily urinary excretion is estimated to amount to 1 – 2 mM per day. Homovanillic acid is a metabolite of the neurotransmitter dopamine and is eliminated in urine at approximately 4 mg per day[223,224]. Nonetheless, a significant increase in urinary excretion above baseline after consumption of anthocyanin-rich products has been

observed across multiple studies for these two metabolites[122,159,219,225], including one study in which anthocyanins were labelled using ^{13}C carbon[123]. This indicated that the observed quantities of these two compounds are partially derived from anthocyanins and due to the late peak (> 5 h) in plasma are more likely to be metabolites of the colonic microbiota[123,226]. It

Based on the early and late explanation above, it could be hypothesised that the factors represent; factor 1 = 'dietary phenolic acids, initial degradation products, and proximal microbial metabolites' and factor 2 = 'colonic microbial metabolites'.

3.1.5 Computation of composite measures representing factors 1 and 2

To explore the association of the factors with clinical responses to the intervention such as vascular function, a composite measure was computed to represent each factor.

The factor scores were computed using the built-in Stata command (regression method), using the two-factor model described above (see Section 3.1.4). Calculating factor scores reduces the original set of variables into a smaller set of new variables through linear combination, by taking all variables into account and involves using the correlations of each variable with the respective factors as weights for the calculation of each score. The calculation method of the scores yields approximately standardised variables with a mean ≈ 0 and standard deviation ≈ 1.

The distributions of factor scores 1 and 2 is shown in **Figure 3-3A**. Nearly all factor 1 scores ranged from -1 to +1 with one outlier at 4.26, i.e., with one exception all values were within approximately ±1 SD from the mean. Because this outlier was inherent to the factor analysis on which these scores were based, this participant was included in all following tests unless also an outlier for the dependent variable (as defined by a value greater than 3 SD from the mean). In contrast to factor 1 scores, factor 2 scores were slightly more spread across a range from -2 to +3. The point at which the scores started to markedly decrease or increase and deviate from the average was at approximately the 15th and 85th percentiles. The datapoints in the upper and lower 15% are signified as black dots for both factor scores in the figure. The Pearson correlation coefficient between factor scores 1 and 2 was 0.53, which corresponded closely with the oblique rotation and a correlation between the original factors of 0.5 (**Figure 3-3B**).

To gain a more relatable overview from the factor relationships, a second type of score using the raw values (in µmol/24 h) was calculated to represent the average change of metabolites characterised by factors 1 and 2, respectively. The average score for each factor was calculated by averaging the metabolite variables with high loadings as determined in Section 3.1.4. Of note, the sulfate group was omitted from both average scores due to cross-loading on both factors. In total, the factor 1 average score was comprised of 7 metabolite groups (26

metabolites) and the factor 2 average score of 6 groups (14 metabolites). Notably, in contrast to the factor scores, the average score does not consider the relative importance of each variable but weights each variable the same.

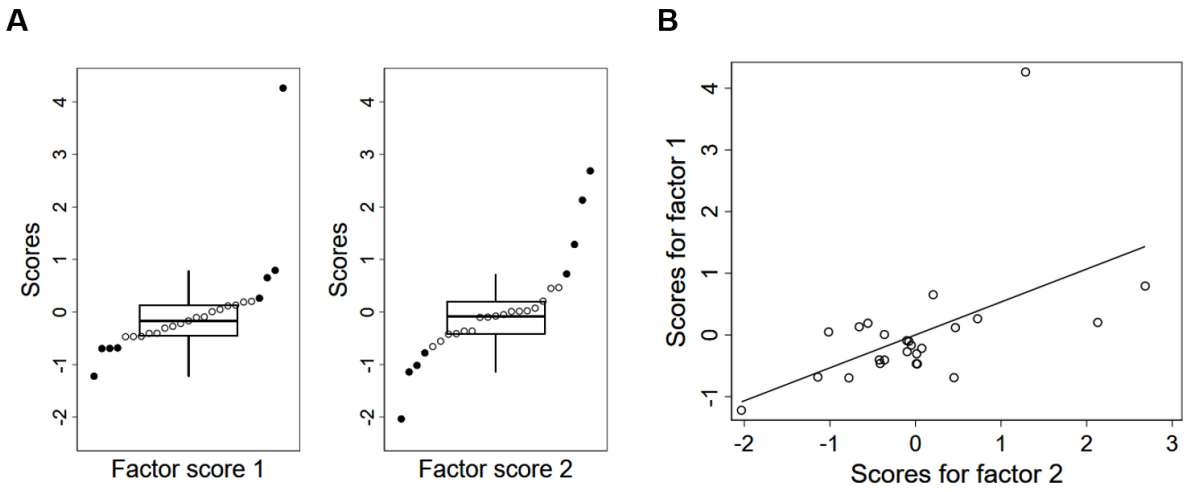


Figure 3-3. Distribution and correlation of factor scores

Scores were computed for each of the two factors. A factor score is a composite measure and its computation reduces the original set of variables into a smaller set through linear combination. It takes all variables into account and uses the correlation of each variable with the respective factor as weights for the calculation of each score.
(A) Both factor scores were approximately standardised variables with mean ≈ 0 and SD ≈ 1. The black datapoints denote high and low subsets, determined using the upper and lower 15 percentiles as cut-offs.
(B) The two scores were moderately correlated (r = 0.53).

Assessing the average scores for unidimensionality and reliability

Each of the average scores was assessed for both unidimensionality and reliability. Unidimensionality is the assumption that the variables together represent a single concept. Variables should therefore only load highly on a single factor, as was the case apart from the sulfate group, which appeared to cross-load on both factors. This core assumption supported the need to exclude the sulfate group from the factor 1 average score, as otherwise the average score would include a variable which shares similarities with factor 2.

A common way to measure reliability of a score is Cronbach's alpha[227]. This estimates the internal consistency, i.e., the average degree of interrelatedness of the variables comprising a score. A high alpha value means that each variable of the score measures something similar to at least some of the other variables. A high alpha is generally an indicator for better reliability of the composite measure, except when very high (in this case alpha may indicate redundancy within the score). The alpha values for the average scores were 0.71 (factor 1) and 0.34 (factor 2). The alpha value for average score 1 decreased to 0.51 when including the sulfate group, further reinforcing that this group should be excluded from the score. On the other hand, it did not improve or worsen the alpha value when included in average score 2, indicating that it

should not be included in this score instead. Generally, an (arbitrary) threshold of at least 0.7 is widely considered as desirable for the alpha value, i.e., the alpha value for average score 2 would be regarded as too low. However, rather than represent a fixed characteristic of a score, Cronbach's alpha applies to a particular sample[228]. Because of the small sample size used for the factor analysis in the present case, it was difficult to determine an appropriate threshold for the acceptance of the alpha values.

Two alternative measures were considered: (1) the correlation of each variable to the corresponding average score and (2) the correlation between the average score and relevant factor score. It is recommended that each variable should be at least moderately correlated (r > 0.3) with the relevant score[193]. Here, the correlations between the score and individual variables were regarded sufficient. Correlation coefficients for factor 1 variables ranged from 0.34 to 0.94 and from 0.59 to 0.95 for factor 2 variables.

There was a high degree of consistency between the average and factor scores with correlations of 0.98 and 0.83 for factors 1 and 2, respectively (**Table 3-5** and **Figure 3-4**). Between themselves, the correlation was 0.49 for average scores 1 and 2 which is very close to the correlation between the original factors of 0.5. Interestingly, including the sulfate group in average score 1 decreased the association with factor score 1 to 0.89 and increased the association with factor score 2 from 0.57 (close to what was expected) to 0.739, again reinforcing the notion to exclude the sulfates from average score 1. Figure 3-4 also clarified that a scoring around 0 on a factor did not mean 'no change' in metabolites, but rather 'average

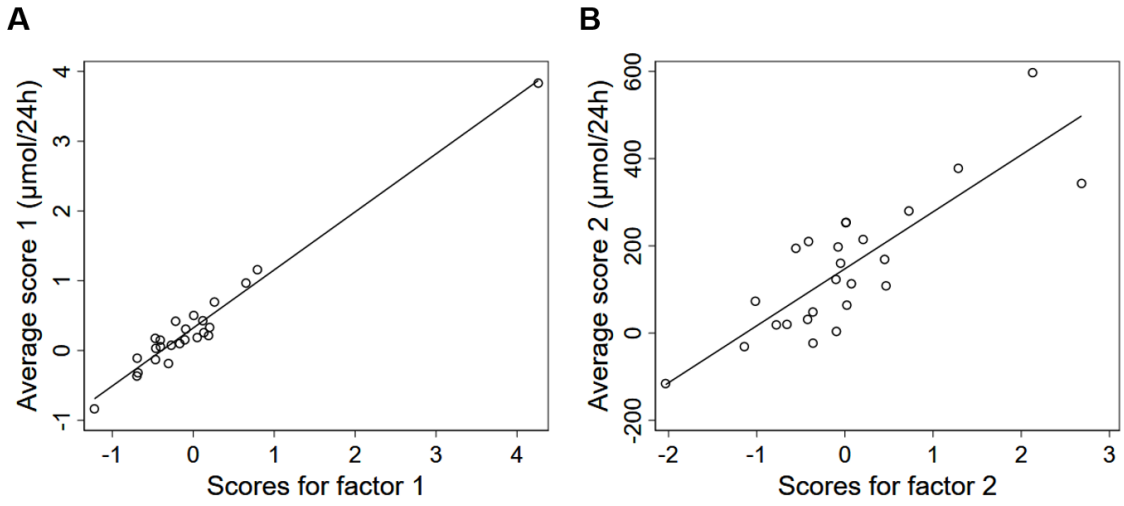


Figure 3-4. Correlation between average and factor scores

The figure shows that there was a strong correlation between the average and factor scores for both factor 1 (A) and factor 2 (B).
The average score for each factor equalled the average amount of change of metabolite variables with high loadings as determined in Section 3.1.4 and carried the unit µmol/24h. The sulfate group was omitted from both average scores due to cross-loading on both factors. Note that the average score differed from the factor score in that each variable was weighted the same.

change in metabolites'. This meant for the *factor scores* that the mean change in metabolites within 1 SD from 0 equalled 0.22 µmol/24h (factor 1) and 128.39 µmol/24h (factor 2).

Table 3-5. Correlation between average scores and factor scores

	Factor score 1	Factor score 2
Average score 1	0.983*	0.569*
Average score 2	0.517*	0.827*

* Significant at 0.01 level
** Significant at 0.05 level

In conclusion, the average scores were considered to be valid representations of the factor scores due to the close relationship between both types of scores. In addition, the assumption of unidimensionality, estimate of internal consistency and correlation between the average and factor scores for factor 1 all supported the exclusion of the sulfate conjugates from the average score for factor 1.

Factor 2 metabolites increased about two times as much as factor 1 metabolites

Figure 3-5 illustrates the average change in factor 1 and 2 metabolites. Particularly for factor 2 metabolites, a large difference in slopes of the curves was apparent, indicating a large variability between the participants in response to the intervention. Seven of 25 participants decreased in factor 1 metabolites, while only three decreased in factor 2 metabolites. Although factor 1 encompassed nearly twice the number of metabolites than factor 2, it was striking that the average amount of factor 1 metabolites at baseline and endpoint was about 220 to 280-fold smaller than factor 2 metabolites (**Table 3-6**). A big proportion of this difference was driven by hippuric acid, which made up 58% and 72% of factor 2 metabolites at baseline and endpoint, respectively. Without hippuric acid, the amount of factor 2 metabolites was still 98 and 84 times higher than factor 1 metabolites. Furthermore, on average, factor 2 metabolites (including hippuric acid) increased about two times as much as factor 1 metabolites between baseline and endpoint (76% versus 36%).

Both factors included metabolites which were shown to significantly change following blueberry intake[159], however it was clear that amount-wise the main urinary metabolites of blueberries in this study were represented by factor 2. The factor scores therefore not only represented the type of metabolite (early-stage, potentially dietary phenolic acids absorbed in the small intestine versus late-stage, likely colonic microbial metabolites), but also reflected their level of bioavailability.

Figure 3-5 also revealed the presence of a potential outlier for each factor. In both cases it was only the six-month value which significantly departed from the group average. There was no indication of an experimental or measurement error in either case. In addition, the metabolite

levels were within previously observed boundaries. For factor 1 the large value was mainly driven by vanillic acid and its glucuronide derivatives for which the total amount observed was 36.7 µmol/24h. This amount was comparable to a study which gave participants a bilberry and lingonberry puree containing nearly twice as much anthocyanins as given in this study and which reported 77.6 µmol/24h of vanillic acid and derivatives[122]. Similarly, for factor 2 the large value was mainly influenced by a single metabolite, hippuric acid. The measured amount was 10.5 mmol/24h for this participant. A study investigating hippuric acid as a biomarker for fruit and vegetable intake reported amounts up to 11.7 mmol/24h[229]. Moreover, because the potential outlier for factor 2 also had a high baseline, this participant was not an outlier when looking at change from baseline to endpoint. In conclusion, it appeared that the extreme values were due to individual fluctuation and that there was no practical reason, such as a quantification error or unrealistic amounts, to exclude these participants. Indeed, these individuals are particularly interesting for this exploratory study as its aim is to differences in the metabolisation of blueberries.

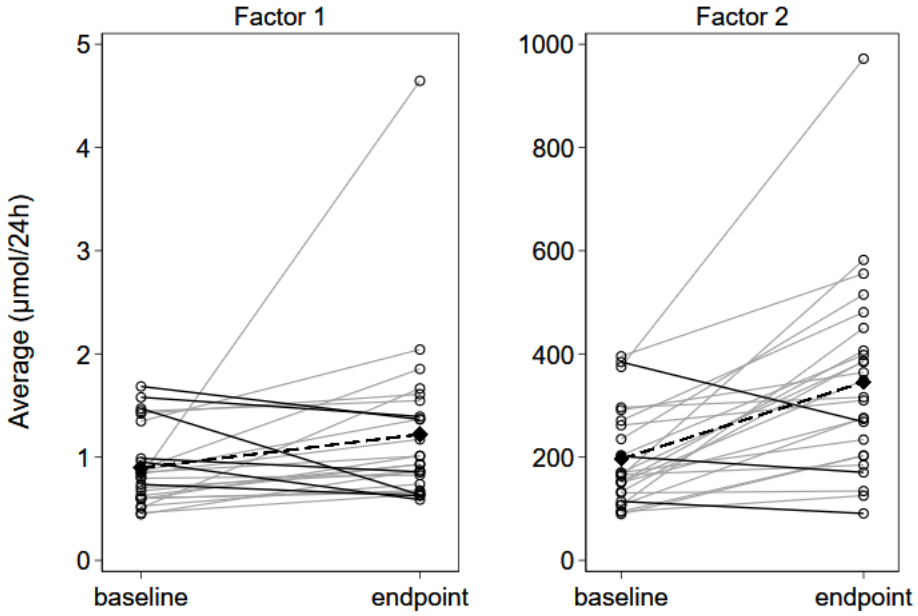


Figure 3-5. Average change of factor 1 and factor 2 metabolites

The average amount of metabolites at baseline and endpoint is shown for each participant and connected with lines. The gray lines indicate a positive change, while the black lines show negative change. The black diamonds and dashed line show the means. Mean changes from baseline to endpoint were 0.3 and 147.3 for factors 1 and 2, respectively.

Table 3-6. Mean values for factor 1 and 2 metabolites at baseline and endpoint

	Factor 1	Factor 2	Factor 2 (without hippuric acid)
Baseline	0.90 ± 0.37	196.35 ± 93.47	87.82 ± 61.19
Endpoint	1.22 ± 0.83	345.34 ± 187.76	101.96 ± 54.78
Change (% change)	0.32 ± 0.84 (36%)	147.35 ± 154.17 (75%)	14.14 ± 56.30 (16%)

Mean values ± SD in µmol/24h

3.1.6 Investigating the relationship between change in FMD and factors 1 and 2

To assess if one of the metabolite clusters associated with change in vascular function as indicated by FMD, ΔFMD was included in the factor analysis (**Figure 3-6**). FMD was singled out over other cardiovascular risk factors as the gold standard method to measure endothelial function, which is an important risk factor for cardiovascular disease and a central aspect in metabolic syndrome.

The analysis resulted in the same two factors as above with the same general loading pattern for the metabolite groups. Interestingly, ΔFMD clearly loaded onto factor 2 (factor loading = 0.64, communality = 0.32) and weakly loaded on factor 1 in the negative direction (factor loading = -0.25). This suggested that the change in FMD was aligned with the accumulation of late-stage metabolites in urine.

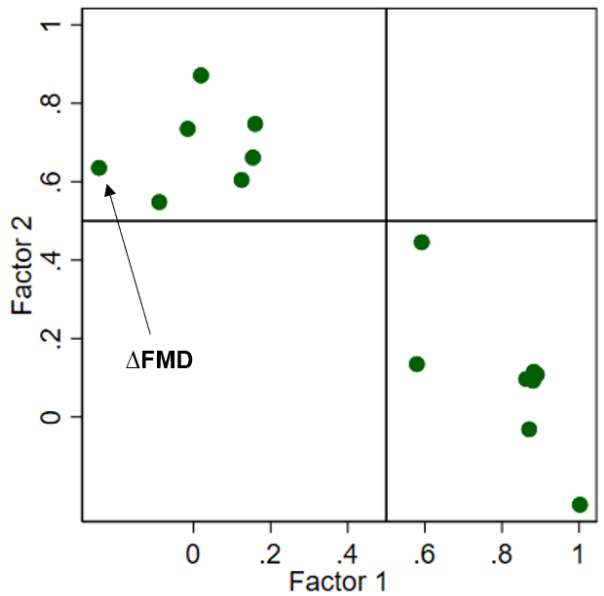


Figure 3-6. Factor loadings for factor analysis including ΔFMD

Including ΔFMD in the factor analysis resulted in the same two factors with the same general loading pattern. The loading of ΔFMD onto factor 2 suggested that the change in FMD was aligned with the accumulation of late-stage metabolites in urine.

Factor score 2 was moderately associated with ΔFMD

The relationship between each of the factor scores and ΔFMD was further explored. First, by determining Pearson's correlation coefficient (**Figure 3-7**) and then by comparing the FMD response between groups of individuals who scored highest and lowest on the factors (**Figure 3-8**).

There was a moderate positive correlation between ΔFMD and factor 2 (r = 0.41, p = 0.042) suggesting that an improvement in FMD could be associated with an increase in urinary late-stage metabolites. Considering the limited number of observations in this analysis, this was a significant finding. Moreover, the relationship was substantiated by a moderate correlation of the individual metabolite groups which loaded on factor 2 with ΔFMD (**Table 3-7**).

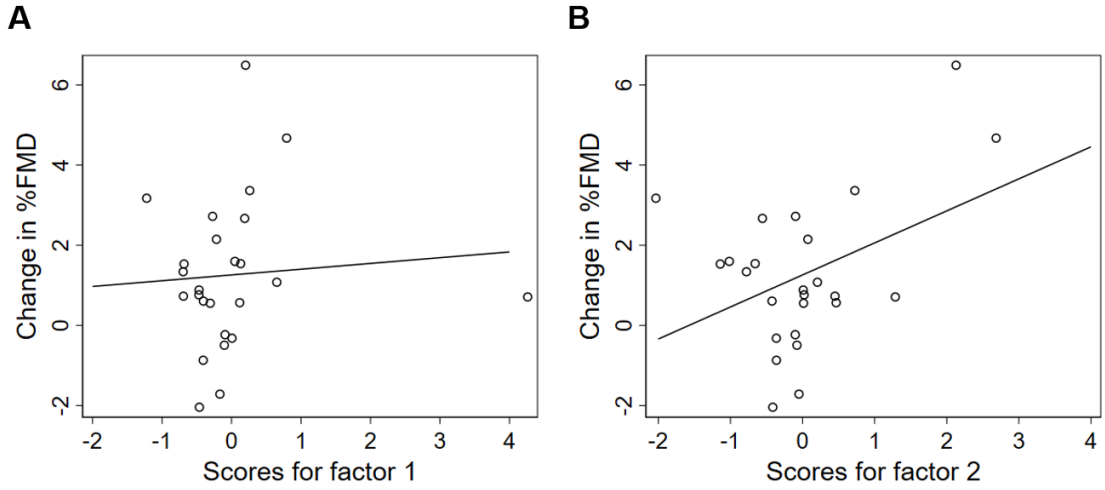


Figure 3-7. Correlation between ΔFMD and factor scores 1 (A) and 2 (B)

(A) No correlation was observed between ΔFMD and factor score 1 (r = 0.08, p = 0.722). When the outlier was removed, the correlation coefficient increased to r = 0.29 (p = 0.169), however there was no evidence of a relationship when pairwise correlations between ΔFMD and the individual metabolite groups comprising factor score 1 were inspected (Table 3-7).
(B) There was a moderate positive association with factor score 2 (r = 0.41, p = 0.042).

No correlation was observed between factor score 1 and ΔFMD (r = 0.08, p = 0.722). As Pearson's correlation is sensitive to outliers, for completeness the correlation was also determined excluding the participant with the very high factor 1 score. The correlation between ΔFMD and factor score 1 was then moderately positive (r = 0.29, p = 0.169), though to a lesser extent than factor score 2. However, of the eight metabolite groups loading primarily on factor 1, only one was weakly correlated with ΔFMD, which in turn was due to a moderate correlation of a single metabolite (chlorogenic acid) with ΔFMD (r = 0.28, p = 0.017). Because of the lack in evidence of an association between the majority of factor 1 metabolite groups and ΔFMD, it is possible that a potential correlation between factor 1 and ΔFMD was enhanced by the oblique rotation used for the factor analysis, i.e. although the individual variables primarily

constituting factor 1 may not have been associated with ΔFMD, the correlation between factors 1 and 2 and the inclusion of all variables in the computation of factor scores could have resulted in the factors partially representing the same underlying metabolic profile.

Table 3-7. Pairwise correlations between ΔFMD and metabolite groups

		ΔFMD
Factor 1	Methoxybenzoic acid	0.05 (0.80)
	Hydroxycinnamic acids	0.08 (0.70)
	Methoxycinnamic acids	-0.01 (0.95)
	Cinnamic acid esters	0.36 (0.07)
	Glucuronide conjugates	0.07 (0.75)
	Glutamate conjugates	0.11 (0.60)
	Hydroxybenzoic acids	0.00 (0.99)
Factor 2	Hydroxyphenylpropionic acids	0.42 (0.04)
	Methoxyphenylpropionic acids	0.33 (0.11)
	Hydroxyphenylacetic acids	0.36 (0.07)
	Methoxyphenylacetic acids	0.40 (0.05)
	Hippuric acids	0.39 (0.05)
	Benzoic acid	0.16 (0.45)

Values are Pearson correlation coefficients with p values in parentheses.

Interestingly, all individuals whose %FMD decreased following the intervention (ΔFMD < 0, n = 6) scored at around 0 on both factor 1 and factor 2, i.e., close to the average, while those who scored at the lower or upper end of the factors all experienced at least a small positive change in FMD. In this factor model, 32% of the variance in ΔFMD was explained with these two factors. The largest portion was therefore not captured by this model and hence it was not entirely surprising that the influence of other variables (such as age, BMI, blood pressure, lipid profile) could have caused a random scattering around zero. There may also be additional factors which were not identified in this dataset or other metabolites which were not measured in this study. To get a better idea of any potential effect the factors had on ΔFMD, those with scores at the either end of the range were compared.

The group with the highest factor score 2 showed the greatest improvement in %FMD

To assess the FMD response between individuals who scored highest and lowest on the factors, the score was used to group individuals by using the upper and lower 15 percentiles where those with particularly high or low scores separated visibly from the average. This resulted in two groups for each factor, high and low, each with n = 4 (**Figure 3-8** and see Figure 3-4 for the distribution of both scores). The low and high groups corresponded to a mean

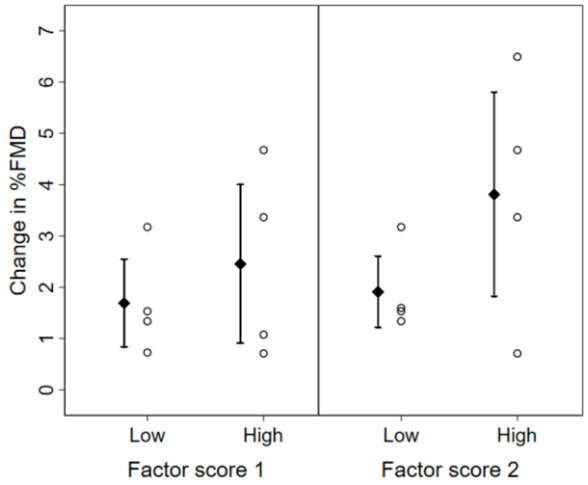


Figure 3-8. FMD for high and low factor score groups

Figure shows the change in %FMD for individuals who scored highest and lowest on the factor scores. These were identified by using the upper and lower 15 percentiles as cut-offs and resulted in n = 4 per group. Those who scored high on factor 2 tended to have a greater increase in %FMD in comparison to those in the low group and both high and low groups of factor score 1.
Hollow circles = individual data points. Mean (black diamonds) and 80% CI (error bars) displayed for each group.

average change in metabolites of -0.4 and 1.7 µmol/24h for factor 1 and -41.6 and 399.54 µmol/24h for factor 2.

From a visual evaluation of the means of ΔFMD, those who scored high on factor 2 tended to increase in %FMD more than the low group and both groups of factor score 1. Scoring high on factor 1 also slightly increased %FMD to a lesser degree in comparison to the low group. Mean differences between the groups and results of an independent *t*-test are listed in **Table 3-8**. Although evidence of a difference between the groups was not conclusive, the results provided some additional support that a positive association between factor score 2 and ΔFMD existed, i.e. that greater ΔFMD was aligned with higher levels of factor 2 metabolites. The mean difference in ΔFMD between the high and low groups of factor score 2 was 1.9% (80% CI [0.04, 3.75], p = 0.191). An increase in %FMD of this size would be of significant clinical relevance and has been associated with a reduction in risk for cardiovascular disease of nearly 16%.[230] However, the study was not powered for this analysis and with only four participants per high and low group, the test had low power. Therefore, results should be interpreted with caution and confirmed in further studies.

Table 3-8. Mean differences in ∆FMD between low and high groups of factor scores 1 and 2

	Mean ∆FMD (low)	Mean ∆FMD (high)	Mean difference (high – low) [80% CI]	t statistic	p value
Factor 1 score	1.69	2.46	0.76 [-0.79, 2.32]	0.707	0.513
Factor 2 score	1.91	3.81	1.90 [0.04, 3.75]	1.475	0.219

Values show ∆FMD in %. n = 4 for each group.

Identification of two clusters representing high and low metabolisers

It appeared that ∆FMD was slightly increased in those with a high factor 1 score, however in contrast to factor 2 metabolites, there was no evidence of an association of the individual metabolite groups comprising factor 1 with ∆FMD. The potential association may have been due to the correlation between the factors which resulted from the oblique rotation during the factor analysis, meaning that the two factors were at least partially estimators of the same thing. Indeed, the low and high groups of both factor scores overlapped considerably, with each factor sharing three of four participants, i.e., those in the high group scored highly on both factors and those in the low group scored low on both factors. This could imply that an improvement in vascular function may be associated with a general ability to absorb and metabolise blueberry compounds. This further suggests that those in the high group were high metabolisers, while those in the low group were low metabolisers.

Plotting the ranks of average change for each factor against each other (**Figure 3-9**) revealed two clear groups in the lower left and upper right corner, clearly demonstrating the overlap seen for the factor scores. The high group had the biggest positive changes in both factor 1 and 2 metabolites (1.4 and 362.5 µmol/24h, while the low group changed in the negative direction (-0.5 and -42.7 µmol/24h).

The remainder of participants scattered around the means of either factor (corresponds to mid-range ranks). Factor 1 metabolites were clustered closely around zero (-0.2 to 0.50 µmol/24h), while factor 2 metabolites were more spread (-23.3 to 254.0 µmol/24h), but within 1 SD from the mean. Although the graph hinted at the formation of two middle groups, there were no visually distinct clusters. Perhaps more observations could have clarified the presence of further clusters, however with the present dataset, it appeared that these participants were average for both factors. Therefore, the interesting groups resulting from this analysis were the high/high and low/low groups at the extremities of both factors 1 and 2 and the further analysis laid out in Section 3.2 was limited to these two groups.

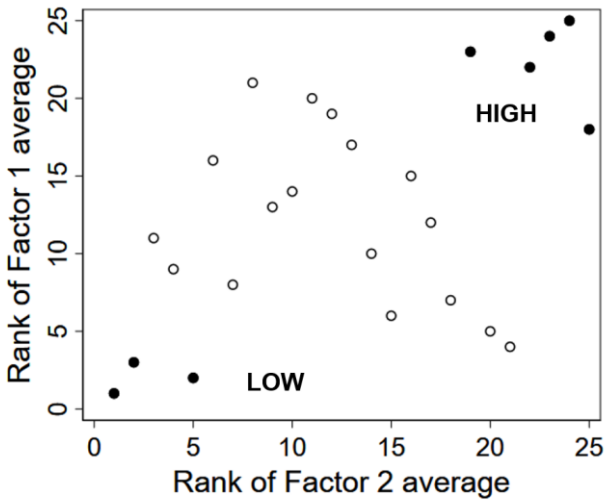


Figure 3-9. Ranks of factor averages

Figure shows the ranks of two factor averages plotted against each other. This highlights two groups: a high/high group, which is high in both factors 1 and 2 (black dots top right corner, n = 5), and a low/low group, which is low in both factors (black dots bottom left corner, n = 3).

Confirmation of high and low metaboliser clusters using cluster analysis

At a later date, after the analysis described in Section 3.2 was performed, the groups of high/high and low/low metabolisers identified via visual inspection in Figure 3-9 were confirmed using k-means cluster analysis. The optimal number of *k* clusters was determined using a scree plot which plotted the within sum of squares for cluster solutions with *k* ranging from 1 to 20[231]. In the scree plot, the point at which the curve breaks indicates the optimal number of clusters. Because the results depend on the initial cluster centres, this was repeated 50 times with random starting points. A cluster solution with $k = 4$ clusters was indicated in 82% of the cases (**Figure 3-10A**) and used as the optimal cluster solution.

Figure 3-10B shows the four clusters identified by k-means cluster analysis. Their corresponding factor average scores are shown in **Figure 3-10C**. The high/high cluster (cluster 4) was the same as determined in Figure 3-9, with 1.3 and 362.5 µmol/24h for factors 1 and 2. The low/low cluster (cluster 1) included three additional participants who had slightly higher factor 1 and factor 2 values, however the group mean in cluster 1 for both factors was still negative (-0.2 and -19.4 µmol/24h for factors 1 and 2). In addition, the cluster analysis partitioned the middle group into two groups, which were previously viewed as a single group. Relative to the sample, cluster 2 was characterised by medium increase in factor 1 and a low increase in factor 2 metabolites (0.3 and 93.0 µmol/24h). Cluster 3 was characterised by no change in factor 1 metabolites and a medium increase in factor 2 metabolites (0.0 and 207.3 µmol/24h).

Figure 3-10D and **Table 3-9** show the ΔFMD responses for each cluster. A one-way ANOVA was performed to test differences between the clusters. Cluster 4 had the highest response as expected and there was support for an effect of factor 2 metabolites on FMD when comparing cluster 4 to clusters with the lowest amounts of factor 2 metabolites, although this was statistically inconclusive (cluster 1 group difference = 1.85%, p = 0.235; and cluster 2 group difference = 2.17%, p = 0.098). These findings are similar to the results found when comparing the high and low groups of factor score 2 (see Table 3-8).

The contribution of factor 1 metabolites in vascular function was less clear. Cluster 2 increased slightly in both factor 1 and 2 metabolites in comparison to cluster 1. The slight increase in metabolites had no effect on the FMD response, however this may simply indicate that a certain threshold of metabolites must be reached to observe a noticeable effect. Unexpectedly, cluster 3 had the lowest FMD response, despite a much higher increase in factor 2 metabolites than clusters 1 and 2. This apparently contradicted the role of factor 2 metabolites in mediating vascular function. However, it may be that factor 1 metabolites could function as an indicator of other unobserved underlying causes. Like cluster 1, cluster 3 lacked a response in factor 1 metabolites. However, in cluster 1, the lack of both factor 1 and factor 2 metabolites seems more suggestive of a general lack of metabolism and/or absorption. In the case of cluster 3, the lack of factor 1 metabolites, but not factor 2, may point to other aspects of metabolism such as differences in the composition of small intestine microbiota. This could not only affect the metabolism and absorption of phenolic compounds, but also other dietary components which may have an impact on vascular function. The small intestine plays a big role in the digestion of macronutrients, such as lipids, and the proximal microbiota have been found to regulate lipid digestion and absorption in an animal study[232]. Differences in bacterial strains may therefore affect the absorption of lipids, which may affect vascular function. Acute postprandial effects of increased serum lipids have been shown to have a deleterious effect on endothelial function[233] and in a cross-sectional study, elevated levels of triglycerides were shown to be inversely associated with FMD[234]. In the present study, cluster 3 did not display an increased level of triglycerides in comparison to the other clusters (data not shown) and thus this theory could not be confirmed. Nevertheless, although cluster 3 was not further investigated in this thesis, other factors relating to the small intestine and vascular function may be an interesting subject for future research.

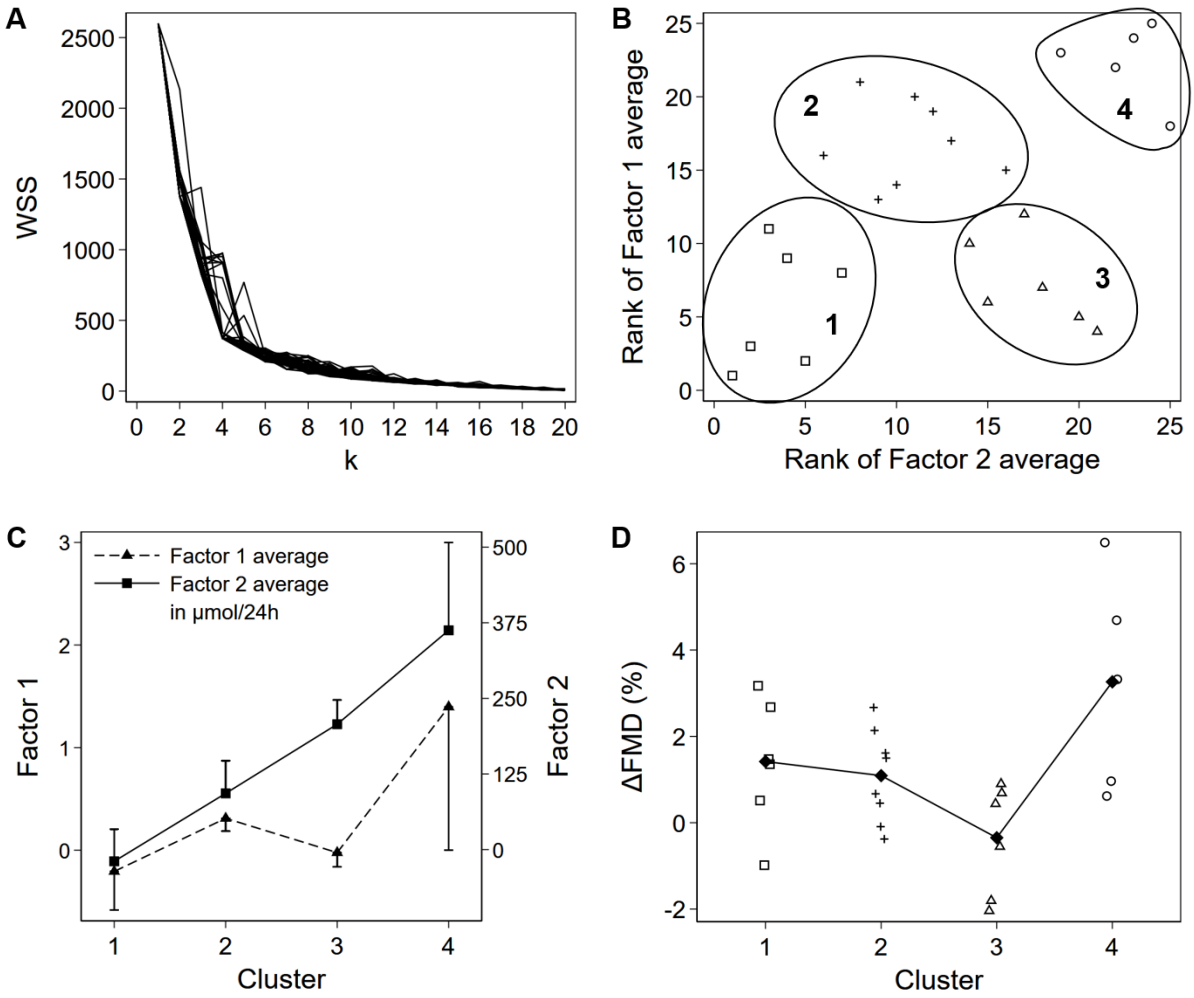


Figure 3-10. Identification of metaboliser clusters using k-means clustering of factor average ranks

(A) Scree plot of within sums of squares (WSS) to determine the optimal number of clusters from solutions with *k* ranging from 1 to 20. Scree plot was repeated 50 times using random initiation points for clustering. *k* = 4 was indicated in 82% of cases.
(B) Scatterplot of ranks of factor 1 and 2 averages. Clusters identified using k-means clustering shown by different symbols and numbers (1 = □, 2 = +, 3 = △, 4 = ○,). Group 1 corresponded to the low/low group, and group 4 corresponded to the high/high group identified in Figure 3-9.
(C) Scatterplot shows means of factor 1 and 2 average scores by cluster (µmol/24h) + or – SD (+ for factor 2; - for factor 1). SD shown in one direction only for clarity.
(D) Scatterplot shows the ΔFMD (%) response in each of the clusters identified in (B). Table 3-9 below this figure shows the Tukey adjusted pairwise comparisons of means.

Table 3-9. ΔFMD (%) pairwise cluster comparisons

Cluster comparison	Mean difference [95% CI]	t value	p value
1 vs 2	-0.33 [-2.67, 2.02]	-0.39	0.980
1 vs 3	-1.76 [-4.27, 0.74]	-1.96	0.233
1 vs 4	1.85 [-0.78, 4.47]	1.96	0.235
2 vs 3	-1.44 [-3.78, 0.9]	-1.71	0.342
2 vs 4	2.17 [-0.3, 4.64]	2.45	0.098
3 vs 4	3.61 [0.98, 6.24]	3.83	0.005

Results from one-way ANOVA. Tukey adjusted pairwise comparisons.

Clusters 1 and 4 are not significantly different for other cardiovascular risk factors

To test the potential effect of being a high or low metaboliser, only clusters 1 and 4 were compared for multiple cardiovascular risk factors (**Table 3-10**). Many markers moved in the expected direction, however none achieved statistical significance at a level of 0.05. There was however an indication for a reduction in inflammation. Interleukin-6 (IL-6), tumour necrosis factor α (TNF-α), and C-reactive protein (CRP) all reduced in cluster 4 in comparison to cluster 1. The support was strongest for TNF-α (mean difference -0.42 pg/ml, p = 0.08). The results further hinted at a reduction in insulin resistance (-2.55, p = 0.31), as assessed by HOMA-IR, due to a reduction in insulin, but not glucose. Although statistically inconclusive, insulin resistance and pro-inflammatory markers both play a role in endothelial dysfunction, thus these findings would be in alignment with an increase in FMD.

Curiously, both systolic (SBP) and diastolic blood pressure (DBP) increased in cluster 4 versus cluster 1, although again not statistically significant at a level of 0.05 (p = 0.191 and 0.296). Across treatments, no effect was observed on blood pressure in this study[159]. Although population-based studies have reported an inverse association of anthocyanin intake with blood pressure[70,235], this finding is in agreement with recent meta-analyses of clinical trials investigating the effect of anthocyanins on blood pressure[92,236,237]. The large mean difference in SBP observed between clusters 1 and 4 may be due to the fact that two participants in cluster 1 received blood pressure medication and reduced in SBP by 10 and 11 mmHg, whereas all other participants in the two clusters were unmedicated. Excluding the two medicated participants from the analysed greatly attenuated the effect observed on DBP, but not SBP (systolic: 7.75 mmHg, p = 0.311, diastolic: -0.62 mmHg, p = 0.870). This was mostly due to one participant in cluster 4 who had a very high SBP of 170 at the six-month endpoint (increase of 27.5 mmHg from baseline). In sum, these results suggested that metaboliser status likely did not affect DBP. However, due to the uncertainty of the group estimates and small group size, no conclusion on the effect of high metaboliser status on SBP could be made in either direction (attenuation, no effect, or increase).

In this assessment, no effect was observed by cluster for PWV, AIx or levels of glucose, HDLC, TG, IL-8, and adiponectin.

Table 3-10. High vs low group comparisons of cardiovascular risk factors[a]

Change in	n (clus1; clus4	Mean (clus1)	Mean (clus4)	Mean difference [80% CI]	t statistic	p value
PWV (m/s)	(2; 4)	0.59	-0.21	-0.80 [-2.06, 0.46]	-1.030	0.376
AIx@75 bpm (%)	(6; 5)	-2.88	-4.27	-1.37 [-6.71, 3.95]	-0.358	0.729
SBP (mmHg)	(6; 5)	-4.75	5.9.	10.65 [0.26, 21.04]	1.530	0.191
DBP (mmHg)	(6; 5)	-0.25	3.30	3.55 [-1.02, 8.12]	1.119	0.296
Glucose (mmol/l)	(6; 5)	0.22	0.10	-0.12 [-0.66, 0.42]	-0.311	0.764
Insulin (mU/l)	(6; 5)	3.53	0.98	-2.55 [-5.50, 0.41]	-1.242	0.254
HOMA-IR	(6; 5)	1.07	0.27	-0.80 [-1.85, 0.24]	-1.110	0.310
Chol (mmol/l)	(6; 5)	-0.05	0.24	0.29 [-0.09, 0.67]	1.101	0.310
HDLC (mmol/l)	(6; 5)	0.04	0.06	0.02 [-0.07, 0.10]	0.266	0.799
LDLC (mmol/l)	(6; 5)	-0.13	0.08	0.21 [-0.10, 0.53]	0.979	0.366
TG (mmol/l)	(6; 5)	0.14	0.17	0.03 [-0.29, 0.35]	0.148	0.887
IL-6 (pg/ml)	(5; 5)	0.05	-0.26	-0.31 [-0.68, 0.06]	-1.223	0.259
IL-8 (pg/ml)	(6; 5)	0.02	-0.05	-0.07 [-1.08, 0.95]	-0.097	0.926
TNF-α (pg/ml)	(5; 5)	0.24	-0.17	-0.42 [-0.71, -0.12]	-2.010	0.081
CRP (µg/ml)	(6; 5)	0.35	-0.86	-1.21 [-2.33, -0.09]	-1.562	0.175
Adp (µg/ml)	(6; 5)	0.20	6.54	6.35 [-1.89, 14.58]	0.183	0.862

[a] Results from Welch's t-test. Clus1 = cluster 1 and clus4 = cluster 4 from Figure 3-10.
PWV: pulse wave velocity; SBP: systolic blood pressure; DBP: diastolic blood presser; HOMA-IR: homeostasis model assessment estimated insulin resistance; Chol: cholesterol; HDLC: high-density lipoprotein cholesterol; LDLC: low-density lipoprotein cholesterol; TG: triglycerides; IL-6: interleukin-6; TNF-α: tumour necrosis factor α; CRP: c-reactive protein; Adp: adiponectin

3.1.7 Summary of results of the exploratory factor analysis

In summary, the exploratory factor analysis extracted two factors from 46 phenolic metabolites measured in 24 h pooled urine from n = 25 participants after a six-month dietary intervention with daily blueberry intake. Although the data could be perceived to have half of the minimal expected number of observations, careful preparation of the dataset was performed to account for this limitation as well as other challenges including baseline variation in metabolites, great differences in abundance and magnitude of change from baseline, outliers, and missing values. The extracted factors were labelled 'early-stage metabolites' (factor 1) and 'late-stage-metabolites (factor 2) as determinants of absorption location and time with early-stage metabolites predominantly absorbed in the small intestine and late-stage metabolites predominantly absorbed in the colon. Based on this definition, the factors could also represent dietary phenolic acids and initial degradation products (factor 1) and colonic microbial metabolites (factor 2).

The factors were distinct by the type of metabolites, but correlated to a moderately high degree (correlation coefficient = 0.5). Two types of scores were generated to represent each factor: (1) a standardised and weighted variable comprised of all metabolite groups and (2) an average score calculated using to absolute change of selected metabolite groups which loaded on the relevant factor. The average amount of change in metabolites differed greatly between the factors. It was clear that factor 2 metabolites, particularly hippuric acid, increased the most

from baseline, about 2-fold more than factor 1 metabolites. However, as the great majority of anthocyanins reach the colon, it is perhaps unsurprising that the colonic microbial metabolites (factor 2) were the main metabolites to increase in response to the intervention.

Including FMD in the factor model suggested that an improvement in FMD was associated more with factor 2 metabolites which seemed to imply a greater importance of the colonic metabolites metabolism. Plotting the ranks of each factor score against one another revealed four clusters of participants. A comparison of FMD response in the clusters which were high or low in both factors 1 and 2 reinforced the previous finding of a potential association between abundant (mainly factor 2) metabolites and FMD, however the association was not statistically conclusive and due to the small sample size of n = 4 per group, the evidence was not strong enough to conclude the presence of a real effect. Interestingly, the participant cluster which was high in factor 2 and low in factor 1 metabolites had the lowest FMD response of all four clusters. An increase in factor 2 metabolites, but a lack of a response in factor 1 metabolites could suggest that other, unobserved aspects of metabolism impacted on vascular function, such as the microbiome of the small intestine.

There was also an indication of an inverse association with inflammatory markers (TNF-α, IL-6, and CRP) and insulin resitance, with the strongest indication for TNF-α. TNF-α plays a pivotal role in endothelial dysfunction. Mechanisms inlcude an increase in oxidative stress and a reduction NO bioavailability, a major mediator of endothelium-dependent vasodilation[238]. A decrease in TNF-α would therefore be in alignment with an increase in FMD.

The next section investigates the two groups (high and low) as determined from the visual inspection of Figure 3-9 to identify common metabolite profile which may play a role in the mediation of vascular function as measured by FMD.

3.2 Part 2: Identifying an anthocyanin metabolite signature of high metabolisers

The factor analysis results indicated differential metaboliser profiles within the blueberry treatment group and within this, two subgroups were identified which were diametrically opposed. The high group (shown in Figure 3-9) represented a group of five individuals (20% of the analysed group) who had a large metabolite response to blueberry intake, i.e., had the largest rank change in urinary metabolites in both factor 1 and factor 2 from baseline to endpoint. Conversely, those in the low group (n=3; 12% of the analysed group) exhibited minimal responses to chronic six-month blueberry intake, i.e., with low or negative change in urinary metabolites relative to baseline levels.

Having confirmed a seeming disparity between the two groups of interest (high versus low), a series of assessments were made to identify a panel of key metabolites, which may characterise the main differences between high and low blueberry metaboliser phenotypes. Both groups were compared to the placebo group, as the control group which did not receive any blueberries.

To establish whether the panel of metabolites had heath consequences, associations between the metabolite panel and parameters related to vascular function were assessed.

The construction of a blueberry anthocyanin metabolite panel was the first of its kind and thus was exploratory and developmental in nature. It was considered that, if confirmed, that a panel of metabolites which predicted anthocyanin responsiveness would be useful in future studies a) as a screening tool, and b) to identify participants / patients that would benefit from anthocyanin based dietary strategies for health benefits. From a study design perspective, recruiting by metaboliser type would allow studies to more readily understand and account for the inter-individuality in metabolic responses which are conjectured to mask dietary health effects in randomised controlled studies.

3.2.1 A comparison of the metabolite patterns shows the high group strongly differed from the low group and placebo

In the first step to finding the key metabolites, the metabolite profiles of the high (n = 5) and low group (n = 3) were compared. First, the absolute change from baseline (µmol/24h) of each metabolite was ranked from smallest to greatest change for each individual. Then, the rank product of those metabolites was calculated to create a pooled rank across the participants in each group. The combined rank expressed which of the metabolites ranked highest per group. **Figure 3-11** illustrates that there are clear commonalities between ranks among the participants in the high group (identified as H1- H5 on the x axis). Likewise, **Figure 3-12** shows the absolute change for each metabolite for each participant in the high and the low groups (H1-H5; L1 – L3 respectively on the x axis). For comparative purposes, the mean changes in each metabolite *per* group (across the high and low groups; drawn from the blueberry intervention group) are shown alongside the changes in these metabolites in the placebo group, which received no blueberries for six months. In both figures the metabolites are shown in the order of the combined rank in the high group (highest to lowest).

Figure 3-12 clearly demonstrates the stark difference in absolute metabolite changes between the high group and both low and placebo groups.

High metaboliser group

In the high group, metabolites generally increased after the six-month intervention. Hippuric acid made up nearly 80% of total change (4098.8 ± 1974 µmol/24h), followed by benzoic acid (784.4 ± 83.3 µmol/24h (15% of total)) and 3-hydroxyhippuric acid (113.9 ± 65.9 (2.2% of total)). Over half of the metabolites (57%) increased by less than 1 µmol/24h in comparison to baseline. On average only one metabolite decreased, while for all participants individually, metabolites decreased in about 10% of the cases.

Low metaboliser group

A large percentage of metabolites (29 total, 63%) decreased on average in the low group (see Figure 3-12, L1 – L3). 14 metabolites (30%) increased by less than 1 µmol/24h, of which at least one participant had a negative change in most cases (see inset table in Figure 3-12, middle column). Only three metabolites increased more than 1 µmol/24h. Whilst hippuric acid was the metabolite with the greatest pre to post increase, the absolute level in the low metaboliser group was 10-fold less than in the high group (363.1 ± 66.0 µmol/24h). The other two metabolites were 3-hydroxyhippuric acid (8.0 ± 64.0 µmol/24h) and 4-hydroxy-3-methoxyphenylpropionic acid (1.6 ± 2.5 µmol/24h). Although on average 3-hydroxyhippuric acid was the metabolite with the second highest change in this group, it ranked much lower at rank 16 because the positive average was due to only one of the three participants in the group while the other two strongly decreased in this metabolite. Interestingly, benzoic acid strongly decreased in all three participants (-956.4 ± 957.1 µmol/24h).

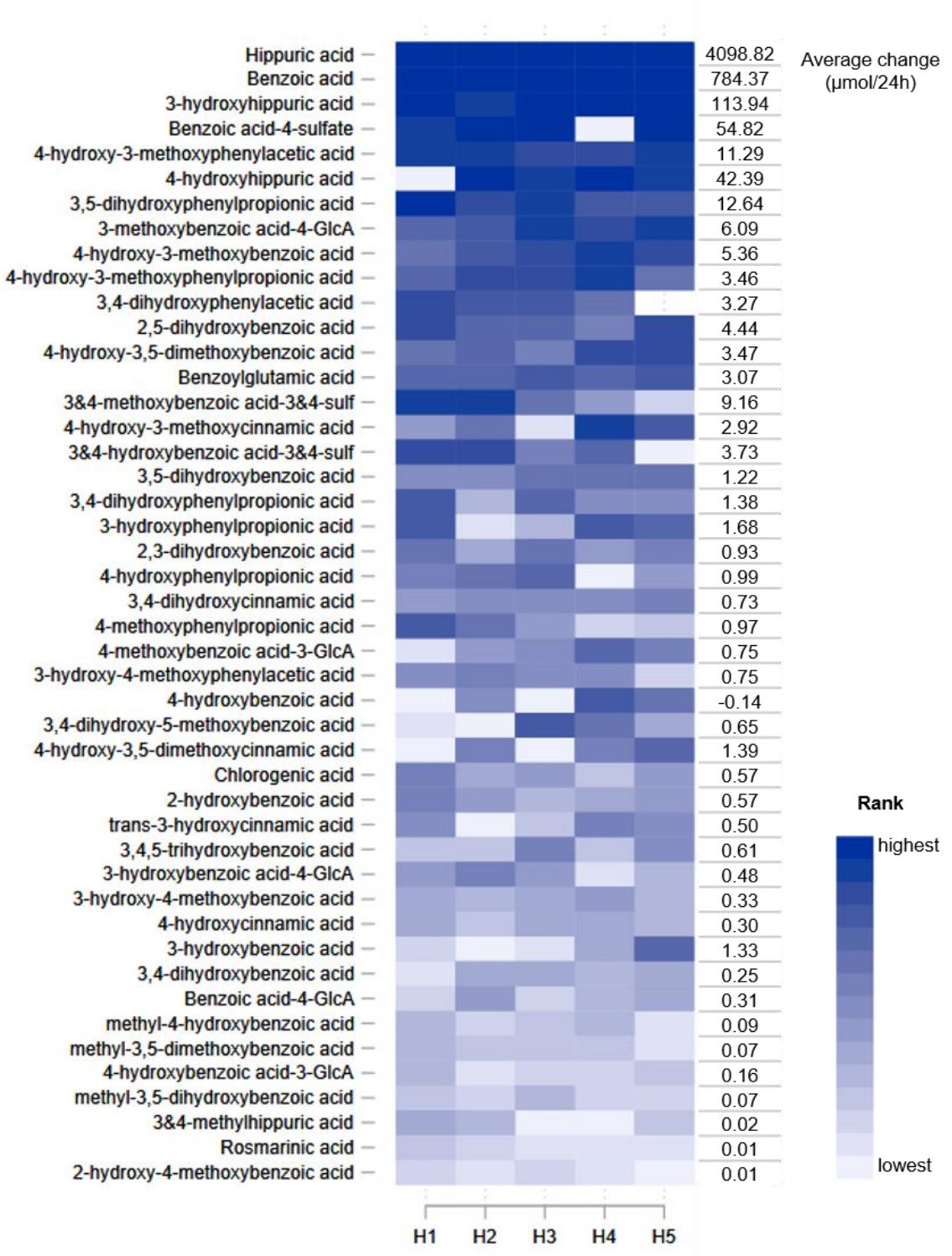
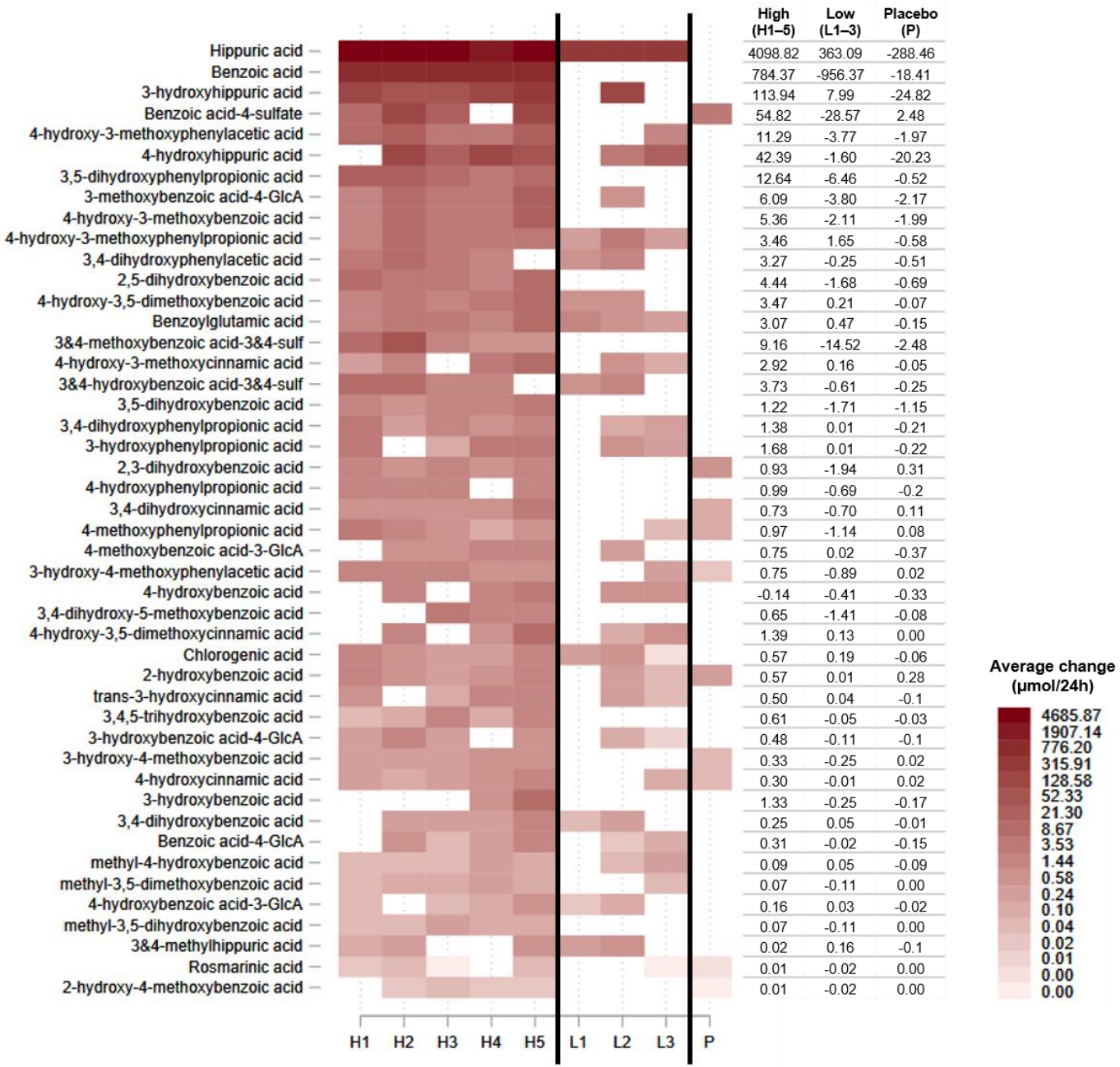


Figure 3-11. Ranked change in metabolites from baseline to 6 mo in the high group

The heatmap shows the ranks of change from baseline for each metabolite (46 metabolites total) per individual in the high/high group (H1 – 5) in the order of their combined rank. Darker colours denote higher ranks (i.e. greater change), lighter colours denote lower ranks (i.e. smaller change).



	High (H1–5)	Low (L1–3)	Placebo (P)
Hippuric acid	4098.82	363.09	-288.46
Benzoic acid	784.37	-956.37	-18.41
3-hydroxyhippuric acid	113.94	7.99	-24.82
Benzoic acid-4-sulfate	54.82	-28.57	2.48
4-hydroxy-3-methoxyphenylacetic acid	11.29	-3.77	-1.97
4-hydroxyhippuric acid	42.39	-1.60	-20.23
3,5-dihydroxyphenylpropionic acid	12.64	-6.46	-0.52
3-methoxybenzoic acid-4-GlcA	6.09	-3.80	-2.17
4-hydroxy-3-methoxybenzoic acid	5.36	-2.11	-1.99
4-hydroxy-3-methoxyphenylpropionic acid	3.46	1.65	-0.58
3,4-dihydroxyphenylacetic acid	3.27	-0.25	-0.51
2,5-dihydroxybenzoic acid	4.44	-1.68	-0.69
4-hydroxy-3,5-dimethoxybenzoic acid	3.47	0.21	-0.07
Benzoylglutamic acid	3.07	0.47	-0.15
3&4-methoxybenzoic acid-3&4-sulf	9.16	-14.52	-2.48
4-hydroxy-3-methoxycinnamic acid	2.92	0.16	-0.05
3&4-hydroxybenzoic acid-3&4-sulf	3.73	-0.61	-0.25
3,5-dihydroxybenzoic acid	1.22	-1.71	-1.15
3,4-dihydroxyphenylpropionic acid	1.38	0.01	-0.21
3-hydroxyphenylpropionic acid	1.68	0.01	-0.22
2,3-dihydroxybenzoic acid	0.93	-1.94	0.31
4-hydroxyphenylpropionic acid	0.99	-0.69	-0.2
3,4-dihydroxycinnamic acid	0.73	-0.70	0.11
4-methoxyphenylpropionic acid	0.97	-1.14	0.08
4-methoxybenzoic acid-3-GlcA	0.75	0.02	-0.37
3-hydroxy-4-methoxyphenylacetic acid	0.75	-0.89	0.02
4-hydroxybenzoic acid	-0.14	-0.41	-0.33
3,4-dihydroxy-5-methoxybenzoic acid	0.65	-1.41	-0.08
4-hydroxy-3,5-dimethoxycinnamic acid	1.39	0.13	0.00
Chlorogenic acid	0.57	0.19	-0.06
2-hydroxybenzoic acid	0.57	0.01	0.28
trans-3-hydroxycinnamic acid	0.50	0.04	-0.1
3,4,5-trihydroxybenzoic acid	0.61	-0.05	-0.03
3-hydroxybenzoic acid-4-GlcA	0.48	-0.11	-0.1
3-hydroxy-4-methoxybenzoic acid	0.33	-0.25	0.02
4-hydroxycinnamic acid	0.30	-0.01	0.02
3-hydroxybenzoic acid	1.33	-0.25	-0.17
3,4-dihydroxybenzoic acid	0.25	0.05	-0.01
Benzoic acid-4-GlcA	0.31	-0.02	-0.15
methyl-4-hydroxybenzoic acid	0.09	0.05	-0.09
methyl-3,5-dimethoxybenzoic acid	0.07	-0.11	0.00
4-hydroxybenzoic acid-3-GlcA	0.16	0.03	-0.02
methyl-3,5-dihydroxybenzoic acid	0.07	-0.11	0.00
3&4-methylhippuric acid	0.02	0.16	-0.1
Rosmarinic acid	0.01	-0.02	0.00
2-hydroxy-4-methoxybenzoic acid	0.01	-0.02	0.00

Figure 3-12. Absolute change (µmol/24h) from baseline to 6 mo for 46 metabolites in high and low metabolisers, compared against placebo

Heatmap:
Absolute change in µmol/24h is shown for the five participants in the high/high group (H1-5), three participants in the low/low group (L1-L3), and the mean of the placebo group (P, n = 30). Data was log transformed for visibility due to the large difference in range. Empty fields denote a negative change.

Table:
Values are the **mean change** per metabolite across participants H1 – 5 (high), L1 – 3 (low), and placebo (n = 30). Data was missing for one

Placebo group

In the placebo group (n = 30) the vast majority of metabolites (78%) did not change over the course of the study. For example, most metabolites decreased marginally, i.e., within 0.2 µmol/24h of the baseline values. Whereas four metabolites decreased by more notable amounts, namely hippuric acid (-288.46 ± 907.57 µmol/24h), 3-hydroxyhippuric acid (-24.82 ± 64.20 µmol/24h), 4-hydroxyhippuric acid (-20.23 ± 60.45 µmol/24h), and benzoic acid (-18.41 ± 592.93 µmol/24h). In general, these data confirm the good adherence to background diet restrictions which were applicable across all intervention groups in the study.

Brief summary on the monitoring of dietary intake during the study

Many of the measured phenolic acids may be naturally present in a multitude of foods (derivatives of hydroxycinnamic acids and to a lesser degree hydroxybenzoic acids [64] or are common catabolites to many flavonoids and other polyphenols. Hence habitual dietary intake is a key factor in the formation of metabolites. In the data presented in this chapter, it is acknowledged that strong variations in diet (across any group, i.e., intervention or placebo groups) could have influenced metabolite profiles; to account for this possibility, the study used a randomised controlled trial design and implemented dietary restrictions across all participants. To recap, in this free-living intervention study, participants were given a list of restricted food items (i.e., to avoid blueberries and minimise anthocyanin intake) for the three weeks before, and throughout the six-month study period. A list of alternative replacement foods was suggested, which included other fruits and vegetables. As this was a long-term, pragmatic study, only blueberries were entirely restricted, whilst other anthocyanin-rich foods such as other berries or red wine were allowed in limited proportions, i.e., one portion per week. Up to four cups of tea and coffee were allowed per day, as well as dark chocolate up to two portions per week. In an attempt to further standardise metabolite profiles, 24h before the study visit, participants were further instructed to consume a low-polyphenol diet and given a more extensive and stricter list of foods to avoid. Finally, a standardised, low-polyphenol evening meal was provided by the research team and consumed by the participant on the evening prior to each assessment visit. It is however acknowledged that phenolic metabolites have been observed in urine 48h after intake of flavonoids[140,221,239], which may have had a bearing on the urine collections made in the 24h prior to each assessment visit.

Of importance to this chapter, whilst the replacement alternative foods (i.e., fruits, vegetables, and beverages (such as bananas, oranges, peppers, cauliflower, beer, fruit juice)) were low in anthocyanins, they were not necessarily low in polyphenols. Therefore, it is plausible that changes in metabolite profiles (pre- to post-intervention) may have reflected changes in habitual dietary intakes as a consequence of adherence to the dietary restrictions. To explore

this, self-reported food frequency questionnaires (FFQ) taken at baseline and endpoint were evaluated[159]. The FFQs reflected the habitual intake of fruits and vegetables and other foods which was reported in the week before the study visit. Anthocyanin intake was estimated using polyphenol composition databases of the US Department of Agriculture (USDA).[71,72]

The FFQ used has been validated (i.e. the assessment of intake has been assessed for accuracy using 24h dietary recalls) and previously used successfully to assess habitual dietary and flavonoid intakes in the context of large epidemiological studies.[98,99] It included 131 items and allowed for the estimation of intake by asking for the consumption frequency of food servings (frequency categories ranged from 'never or less than once a month' to '6+ times per day'; servings were units of food items, such as one apple, or household measures, such as one cup). However, it is important to note that this method is subject to measurement errors and bias which may lead to over- or underestimation of true intake. Participants may under-report true food intake or change in diet during the period of the survey. Also, polyphenol content of some foods may be missing in the databases and it is difficult to extrapolate values from related species, as the polyphenol profiles vary between similar foods. Furthermore, polyphenol content in foods and beverages does not only depend on plant species, but also on the season, year, and processing, which is not always fully documented in dietary records. Finally, intra- and inter-individual variability in polyphenol metabolism and uptake may obscure actual systematic polyphenol exposure.[48]

Therefore, the data derived from the FFQ in the following is understood as an estimation rather than an accurate calculation of dietary intakes. While exact numbers may not be accurate, large differences and comparison of ratios between groups and timepoints are nevertheless able to give an indication of variation between groups or timepoints.

Metaboliser profiles were not a result of changes to habitual diet between baseline and 6 months

As shown in Table 3-11, the FFQ data gave no indication that there were discrepancies in dietary intake which may underpin the metabolite groupings. For example, the low metaboliser group did not have unusually high intakes of polyphenol-rich foods at baseline, compared with the average for the blueberry intervention group. Likewise, the high metaboliser group did not consume particularly low amounts of polyphenol-rich foods at baseline. Neither group significantly reduced, or increased, their intake of anthocyanins, fruits, vegetables, tea, and coffee between baseline and endpoint which provides reassurance that the change in metabolite profiles were likely to be attributable to the blueberry study material. Indeed, between the two timepoints, the low group had increased their intake of fruits and vegetables, whereas the high group had slightly decreased their intake. Intuitively, if any change due to

other fruits and vegetables was expected, we would have expected an increase in phenolic metabolites for the low group and a decrease for the high group, rather than the opposite as observed here.

To further explore whether change in habitual dietary intake may have influenced the metaboliser profiles, efforts were made to identify commonly consumed alternate food sources with common metabolite profiles and subsequently determine the change in their FFQ intakes.

In the high metaboliser group, the increase was most evident for hippuric acid and benzoic acid, both known metabolites in the catabolism of phenolic acids[124]. Especially hippuric acid has been observed frequently as a metabolite of polyphenols, such as from chlorogenic acid in coffee and tea[239], and even proposed as a biomarker[240]. In fact, hippuric acid has many dietary sources and originates from other molecules, for example quinic acid and aromatic amino acids in proteins[222]. Macronutrient intake was comparable between all groups, and although tea and coffee intake greatly differed between groups, the amounts consumed per group were consistent between both timepoints. Therefore, while baseline levels in metabolites derived from tea and coffee were expected to vary among the groups, the change in metabolites between timepoints should not be affected.

It was notable that the 7th highest ranked metabolite to change from pre- to post-intervention was 3,5-dihyddroxyphenylpropionic acid (see row 7 in Figure 3-11 and Figure 3-12), a metabolite that contains a resorcinol structure, which has also been a confirmed metabolite of alkylresorcinols.[241] These are found in wholemeal and cereal products, such as wholemeal bread and wheat bran. In assessment of the FFQ data, it was clear that the high metaboliser group had habitually higher intakes of wholemeal products than the low metaboliser and placebo groups, but again the change between pre- and post-intervention was comparable between groups, suggesting that increases of this metabolite in the panel of the high metaboliser group was not driven by increased intake of wholemeal foods. Supporting the likelihood that increased 3,5-dihyddroxyphenylpropionic acid was derived from blueberry, is data from a study which fed raspberry anthocyanins,[242] which showed that resorcinol was observed as a colonic metabolite of the intervention material. It is conceivable that 3,5-dihydroxyphenylpropionic acid is converted to resorcinol through further microbial β-oxidation and decarboxylation.

Taken together, the FFQ data indicated that the differences in the metabolite profiles was not due to other food sources. This supported the blueberry-derived nature of metabolites in the high group and a lack of response to blueberry intake in the low group. The low group scored low on both factor 1 and factor 2, which would suggest a low bioavailability of blueberry metabolites for these participants in contrast to a high bioavailability for participants in the high group. Such inter-individual variability could be influenced by a variety of reasons, such as the

transit time through the digestive system, the composition of the microflora, the ability to absorb compounds from the gut, and the metabolisation and excretion rate after absorption. For example, the individuals in the low group may have less active colonic microbiota, producing less microbial metabolites available to absorb, metabolise and excrete via urine. The low increase in hippuric acid in the low group, but not in placebo, seems to support this theory.

Table 3-11. FFQ dietary intake at baseline (0 months) and endpoint (6 months)

	High (n = 5)		Low (n = 3)		Full dose (n = 23)		Placebo (n = 27)	
	0 mo	**6 mo**	**0 mo**	**6 mo**	**0 mo**	**6 mo**	**0 mo**	**6 mo**
ACN (mg/d)	32.1 ± 27.8	16.5 ± 7.3	22.2 ± 6.4	26.5 ± 7.4	16.0 ± 10.3	12.5 ± 8.5	20.1 ± 12.9	15.4 ± 14.5
Fruits (g/d)	270.7 ± 176.1	230.3 ± 126.6	132.0 ± 73.2	192.0 ± 111.7	146.2 ± 119.2	130.2 ± 91.6	229.9 ± 113.9	220.3 ± 135.7
Veg (g/d)	279.3 ± 169.3	241.9 ± 114.3	256.9 ± 59.7	311.7 ± 31.7	207.7 ± 103.1	199.9 ± 106.8	266.5 ± 128.3	218.7 ± 88.2
Red wine (g/d)	90.0 ± 136.6	18.0 ± 21.1	39.0 ± 52.7	24.0 ± 27.5	33.6 ± 42.5	15.6 ± 25.3	32.3 ± 39.6	25.0 ± 63.4
Coffee (g/d)	236.1 ± 228.8	195.4 ± 255.5	36.2 ± 39.2	76.9 ± 98.2	296.3 ± 299.4	233.9 ± 211.1	363.4 ± 304.7	295.1 ± 237.5
Tea (g/d)	760.0 ± 419.5	627.0 ± 371.6	63.3 ± 109.7	63.3 ± 109.7	494.9 ± 313.5	419.8 ± 260.1	390.6 ± 248.9	324.2 ± 220.5
Wholemeal (g/d)	69.0 ± 45.8	88.9 ± 117.5	31.9 ± 9.3	43.0 ± 12.0	36.3 ± 26.8	35.4 ± 28.7	-7.9 ± 28.6	40.8 ± 31.0
Energy (kcal/d)	2111.3 ± 304.6	2172.1 ± 668.2	2330.3 ± 294.7	2080.2 ± 361.6	2132.4 ± 590.4	2057.7 ± 546.2	1874.5 ± 516.9	1783.9 ± 456.5
Carb (%energy)	43.1 ± 2.4	45.7 ± 4.8	41.3 ± 2.8	40.9 ± 4.9	43.0 ± 5.6	41.6 ± 6.0	44.5 ± 5.9	44.6 ± 5.6
Fat (%energy)	33.7 ± 3.2	32.7 ± 5.8	37.3 ± 3.6	33.9 ± 1.5	35.7 ± 4.0	36.0 ± 4.3	34.5 ± 5.1	33.2 ± 4.1
Protein (%energy)	19.6 ± 4.3	20.7 ± 5.9	19.2 ± 2.0	21.9 ± 0.3	18.1 ± 3.0	19.3 ± 2.9	17.9 ± 3.1	18.5 ± 3.4

Values are mean ± SD. Dietary data considered invalid and excluded for n = 4 in the placebo group and n = 2 in the full dose group. 0 mo: 0 months (baseline); 6 mo: 6 months (endpoint); ACN: anthocyanins; Veg: vegetables; Wholemeal: wholemeal bread and pasta, brown bread, cereal; Carb: carbohydrates

In conclusion to this analysis, the categorisation of individuals in the high group were confirmed as high metabolisers in response to blueberry intake, and those in the low group as low metabolisers. The following sections examine the metabolites which were common among the high group to identify a subset of metabolites which were most likely to relate to a high metaboliser profile. In addition, the second aim was to determine their role in mediating vascular health.

3.2.2 Determining a preliminary metabolite panel via identification of common metabolites

Identifying common metabolites among high metabolisers

Common metabolites were defined as such if they ranked within the first ten ranks (20% of total metabolites) for absolute change in µmol/24h for at least three of the five participants in the high group (**Table 3-12)**. This cut-off was chosen because the first ten pooled ranks

represented over 99% of total change and were considered as the most relevant metabolites. Interestingly, most of these metabolites were described by factor 2, which was characterised by metabolites of likely microbial origin. As there was some evidence for an association between factor 2 metabolites with ΔFMD, this could infer that these metabolites were not only the ones in greatest abundance, but indeed played a role in mediating potential vascular benefits gained from eating blueberries. Especially phenylpropionic acids and methoxyphenylacetic acids had the highest loadings on factor 2 and may be of particular importance. In addition to the metabolites in the top ten ranks, two more metabolites were selected for further investigation: syringic acid and gallic acid, as these were the expected initial fragmentation products of the two main anthocyanins in blueberries, malvidin and delphinidin.

Table 3-12. Common metabolites among five participants in the high group

Shaded rows indicate metabolites which formed the preliminary panel after evaluation of all criteria (see text for details).

Combined rank[a]	Metabolite	Mean change ± SD (µmol/24h)	n[b]	Different from placebo	Unequal variances[d]	p (FC)[e]
1	Hippuric acid	4098.82 ± 1974.91	5	yes	yes	0.003
2	Benzoic acid	784.37 ± 83.33	5	no	no	–
3	3-hydroxyhippuric acid	113.94 ± 65.93	5	yes	no	0.148
4	Benzoic acid-4-sulfate	54.82 ± 64.48	4	no	no	–
5	4-hydroxy-3-methoxyphenylacetic acid	11.29 ± 9.24	4	yes	no	0.049
6	4-hydroxyhippuric acid	42.39 ± 39.14	4	yes	no	0.320
7	3,5-dihydroxyphenylpropionic acid	12.64 ± 10.22	5	no	yes	0.003
8	3-methoxybenzoic acid-4-GlcA	6.09 ± 6.33	3	no	no	–
9	4-hydroxy-3-methoxybenzoic acid	5.36 ± 5.15	3	yes	no	0.013
10	4-hydroxy-3-methoxyphenylpropionic acid	3.46 ± 2.33	3	yes	no	0.082
13	4-hydroxy-3,5-dimethoxybenzoic acid	3.47 ± 3.29	1	yes	yes	0.035
33	3,4,5-trihydroxybenzoic acid	0.61 ± 0.74	0	no	no	–

[a] Combined rank calculated as geometric mean of ranks of the five participants.
[b] Number of times the metabolite ranked in the top 10 ranks within the high/high group
[c] Metabolite was considered as derived from blueberry if change in metabolite after the intervention was significantly different between the blueberry treatment group and placebo at a significance level of 0.001 (see Table 3 in [159])
[d] "Yes" denotes greater spread in the blueberry group as tested with the Brown-Forsythe test for the equality of variances between blueberry treatment and placebo groups using the median. Variances were considered unequal for p <= 0.2.
[e] p value of Welch's t-test comparing $\log_2$ fold change between the high and low group. Only those metabolites were tested which passed the previous filtering stages.

The selected metabolites (top ten ranks plus two hydroxybenzoic acids) were then evaluated for inclusion in the final panel based on two further criteria:

(1) their dependence on blueberry intake (**Figure 3-13A**) as confirmed as a significant difference between treatments (full dose versus placebo) in published study data.[159]

(2) the inter-individual variability in metabolism of blueberry polyphenols across all participants in the blueberry treatment group (**Figure 3-13B**): estimated by comparing the spread of change between the blueberry and placebo group. If the spread of the data was greater with blueberry intake, this was considered as a metabolite particularly prone to inter-individual variability in the response to eating blueberries.

Common metabolites (among the first ten ranks in Table 3-12) were selected for further inclusion if either of the two criteria was fulfilled. Using these criteria, the twelve metabolites were reduced to eight metabolites (ranks 2, 4, 8, and 33 were neither significantly different from the placebo group nor did they display a bigger spread of the data in comparison to the placebo group). Of the remaining metabolites, only 3,5-dihydroxyphenylpropionic acid (rank 7) was not shown to have significantly changed between treatments. This was selected for inclusion on the basis of criterium (2) as it displayed a much greater spread in the blueberry group versus placebo (also see Figure 3-13B).

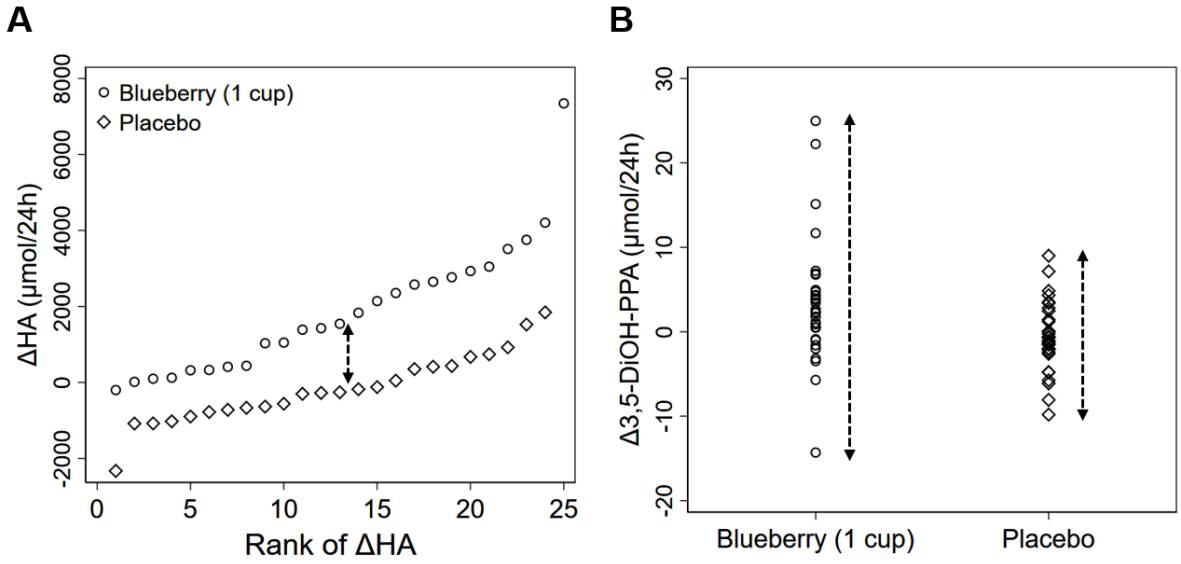


Figure 3-13. Selection criteria for the preliminary metabolite panel

Example graphs to represent the two criteria used to narrow down the selection of twelve metabolites common among high metabolisers to eight metabolites:
(A) a significant difference in change between full dose (Blueberry (1 cup)) and placebo as a marker of blueberry intake
(B) the difference in spread between full dose (Blueberry (1 cup)) and placebo as a marker of inter-individual variability in the metabolism of blueberry polyphenols.
HA: Hippuric acid; 3,5-DiOH-PPA: 3,5-Dihydroxyphenylpropionic acid

Log_2 fold change analysis

To further narrow down the selection of metabolites, the fold change (FC) between low and high groups was compared (**Figure 3-14)**. It was important to look at absolute numbers as well as the FC as it was unclear whether the total amount or the fold change in urinary metabolites

were better indicators of responding to blueberry intake and metabolism. In addition, as mentioned above, many metabolites such as hippuric acid have abundant dietary sources and other foods may have dictated high background levels. Hence although the absolute change may be large, the FC may be small if baseline and endpoint values were large to begin with. FC therefore helped compare the different degrees of change per metabolite in response to eating blueberries between the two groups and was used to determine the utility of the metabolite in distinguishing between both groups (Welch's t-test, Table 3-12, column 'p (FC)'). The rationale for this was that if a metabolite significantly differed between the groups in both absolute and relative values, we could be more certain of this metabolite being able to discriminate between a high and low metaboliser, regardless of whether absolute amount or relative change were better predictors of a suspected metabolism phenotype. FC is presented as the $\log_2$ ratio of endpoint to baseline, which translates to a doubling of the baseline value if $\log_2$ FC = 1 and a quadrupling if $\log_2$ FC = 2. Some values were excluded from the comparison due to small, likely unreliable measurements of less than 0.00005 µmol/24h which inflated the fold change up to ±1000-fold. A total of 20 values were excluded across 19 metabolites and 15 participants. Most of these were in the placebo group and none were pertinent to the analysis in this paragraph.

As shown in **Figure 3-14**, the average $\log_2$ FC from baseline to the 6-month endpoint was greater for all the selected metabolites in the high metaboliser group, compared with the low metaboliser group. In the high group, metabolites varied much less in their $\log_2$ FC (range of means from 0.6 to 1.8), so in contrast to absolute change, none of the metabolites markedly stood out as the one to change the most. This included hippuric acid, which had dominated the absolute amount of change, but was not the highest ranking in terms of FC. For the low group this meant that despite an increase of a notable amount of over 350 µmol/24h, it only increased by 20% in comparison to 220% in the high group. In agreement with low (and sometimes negative) absolute change for the low group, the means of $\log_2$ FC for the low group were mostly around 0 or negative.

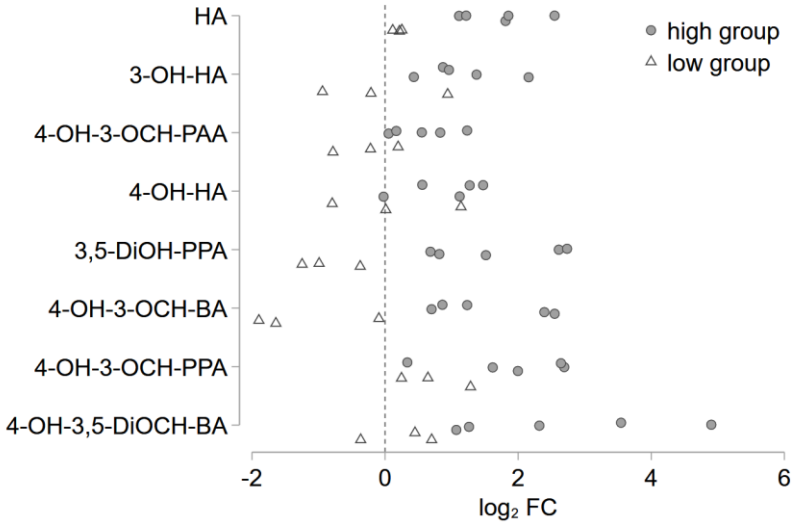


Figure 3-14. Log_2 FC of common metabolites in the high group

Figure compares the log_2 FC in the high and low groups to determine the degree of change from baseline to endpoint per group and how well the groups can be distinguished by the respective metabolite. The high metaboliser group consistently changed to a greater degree and there was a clear distinction between the groups for HA, 4-OH-3-OCH-PAA, 3,5-DiOH-PPA, 4-OH-3-OCH-BA, and 4-OH-3,5-DiOCH-BA.
HA: Hippuric acid; OH: hydroxyl group; OCH: methoxyl group; PPA: phenylpropionic acid; PAA: phenylacetic acid; BA: benzoic acid.
Datapoints are slightly jittered to show overlaying points.

Despite the small group sizes, there was clear distinction between the two groups (low versus high metabolisers) for five of the eight metabolites (hippuric acid, 4-hydroxy-3-methoxyphenylacetic acid (homovanillic acid), 3,5-dihydroxyphenylpropionic acid, 4-hydroxy-3-methoxybenzoic acid (vanillic acid)), and 4-hydroxy-3,5-dimethoxybenzoic acid (syringic acid). In the case of 4-hydroxyhippuric acid and 3-hydroxyhippuric acid, the distinction was not so prominent because the log_2 FC varied greatly in the low group, ranging from a 2-fold decrease to a 2-fold increase, which overlapped with the high metaboliser group. Because the log_2 FC for 4-hydroxyhippuric acid was generally smaller in the high group than 3-hydroxyhippuric acid, the overlap was more pronounced for 4-hydroxyhippuric acid, resulting in a higher p value for the group comparison (see Table 3-12). Although log_2 FC did not show strong support of a significant difference in 3-hydroxyhippuric acid between the high and low groups, the absolute change data confirmed 3-hydroxyhippuric acid as one of the major metabolites upon blueberry intake, increasing three times as much as 4-hydroxyhippuric acid. For this reason, 3-hydroxyhippuric acid remained in the panel, while 4-hydroxyhippuric acid was deemed an uncertain candidate to be reconsidered using the next test. A larger sample size may have been able to confirm or deny a difference in log_2 FC for this metabolite. Similarly, the results for the final metabolite, 4-hydroxy-3-methoxy-phenylpropionic (dihydroferulic acid) were not entirely conclusive either, however because four of the five individuals clearly

increased nearly over 4-fold, this metabolite remained in the closer selection for the panel as well.

Summary of identifying a preliminary metabolite panel

In summary, the metabolite profile of a group of individuals identified previously using factor analysis was studied more closely. Strong differences in metabolite profiles were observed. Because group sizes were small with n = 5 for the high group and n = 3 for the low group, the impact of potential confounders on the high/low concentrations of metabolites was evaluated, and habitual intake of polyphenol-rich foods in the high and low groups and placebo group were compared. No stark differences in intake were found which could explain the observed differences in metabolites between the high and low group. In addition, metabolite patterns in both groups were compared against the placebo group, which received no blueberries and confirmed the good adherence to background diet restrictions which were applicable across all intervention groups in the study. This comparison highlighted the blueberry-derived nature of the metabolites in the high group. Furthermore, a low increase in hippuric acid, the most dominant metabolite, was observed in the low group but not in the placebo group. This seems to support the fact that participants in the low group consumed the blueberry powder but were less responsive to blueberry intake. In conclusion, the categorisation of individuals in the high group were confirmed as high metabolisers in response to blueberry intake, and those in the low group as low metabolisers.

Various criteria were applied to the metabolite profiles and resulted in a selection of eight metabolites which increased with intake of blueberry compounds and mostly differentiated well between high and low metabolisers within the blueberry treatment group. The criteria used were:

1) establishing the most common and relevant metabolites in the high group using absolute change;

2) assessing whether the absolute change in these metabolites was significantly different in the blueberry group compared to the placebo group;

3) assessing whether the spread of absolute change in the blueberry group was greater than in the placebo group, suggesting inter-individual variation in response to blueberry intake; and

4) comparing the fold change between the high and low groups within the blueberry treatment group to ensure a distinction between the two groups.

These metabolites were hippuric acid, 3-hydroxy hippuric acid, 4-hydroxyhippuric acid, 4-hydroxy-3-methoxyphenylacetic acid (homovanillic acid), 3,5-dihydroxyphenylpropionic acid,

4-hydroxy-3-methoxyphenylpropionic acid (dihydroferulic acid), 4-hydroxy-3-methoxybenzoic acid (vanillic acid), and 4-hydroxy-3,5-dimethoxybenzoic acid (syringic acid).

The majority of metabolites was of microbial origin (phenylpropionic, phenylacetic, and hippuric acids). The origin of the two hydroxybenzoic acids could be two-fold. They have previously been detected in blueberries [243] and may also be derived from the initial fragmentation of parent anthocyanins or emerged through various microbial transformations such as dehydroxylation, demethoxylation, and oxidation. The fact that these metabolites, apart from the hydroxybenzoic acids, were all part of factor 2 supports that they seem to describe an underlying anthocyanin metabolism phenotype, which may point to a certain bacterial composition or metabolic activity of the colonic microflora.

The next section considers the association of these metabolites with ∆FMD in the formation of a metabolite panel which is not only descriptive of inter-individual variability to blueberry intake, but also associated with cardiovascular benefits.

3.2.3 Refining the metabolite panel by testing its association with ∆FMD

Formation of a panel by taking the average across the log_2 FC of metabolites

The original hypothesis was that a particular metabolism phenotype which mediates beneficial effects received from eating anthocyanin-rich foods exists as a hidden factor. Coming back to this hypothesis, a scale from the selected metabolites was generated to represent this latent variable. In the first instance, a composite measure was formed from the full selection of metabolites identified in the previous section by taking the average across (1) the absolute change and (2) the log2 FC. The average of log_2 FC was only calculated for those participants who had no missing values for any of the selected metabolites in the panel (n = 3 had missing values for one or two metabolites). Analog to the factor scores computed in the previous section, internal consistency of both averages was determined using Cronbach's alpha to see if one of the averages was more reliable as a composite measure. To remain consistent with the idea that a phenotype would drive differential metabolite levels among people receiving the same treatment, internal consistency was only estimated within the full dose blueberry group. Cronbach's alpha was close to zero for the average of absolute change ($\alpha = 0.06$) and much higher for the average of log_2 FC ($\alpha = 0.84$), suggesting that using fold change of the metabolites for the composite measure was more reliable indicator than absolute change. Hence log_2 FC was used for all further investigations.

Exploring the relationship between the metabolites and ΔFMD

Next, the relationship between the metabolites and ΔFMD using linear regression was assessed. Here, all study participants were used regardless of treatment group, for a total n = 74 with a maximum of three missing values per metabolite (**Table 3-13)**. As a reminder, the three treatment groups were the following: the full dose group received the equivalent of 1 cup blueberries per day, the half dose group received the equivalent of ½ cup of blueberries, and the placebo group received an anthocyanin-free matched placebo. The rationale for including the half dose and placebo at this stage and not before was that if the metabolites truly played a role in mediating vascular function, the origin of the metabolites should not matter, i.e., the placebo group may have experienced similar benefits if other food sources led to high increases in metabolite levels. Moreover, this would indicate a certain transferability of the panel to other flavonoid-rich food sources, given that many flavonoids are degraded to a similar set of catabolites[157].

The $\log_2$ FC of three metabolites (rows 1, 2, and 3 in Table 3-13) was significantly associated with ΔFMD, with each doubling of the metabolite in urine predicting a 0.4 to 0.8% increase in FMD. Four metabolites were moderately associated with ΔFMD (β-coefficient ≈ 0.3), although the effect was not conclusive (rows 4, 5, 6, 7). The final metabolite (row 8) showed the weakest and most doubtful correlation. For this reason, a second panel was formed excluding this last metabolite, vanillic acid. Of all the metabolites, 4-hydroxyhippuric acid had the strongest relationship with ΔFMD. Therefore, it remained in the panel although it had been deemed an uncertain candidate in the previous section.

Table 3-13. Individual regression results for ΔFMD (%)

Row	Metabolite (as independent variable)	n	β (SE)	p	Constant (SE)
1	4-hydroxyhippuric acid	71	0.79 (0.31)	0.014*	0.48 (0.22)*
2	3,5-dihydroxyphenylpropionic acid	72	0.44 (0.21)	0.041*	0.37 (0.23)
3	4-hydroxy-3-methoxyphenylacetic acid (Homovanillic acid)	74	0.61 (0.31)	0.052	0.47 (0.21)*
4	4-hydroxy-3,5-dimethoxybenzoic acid (Syringic acid)	72	0.38 (0.24)	0.125	0.29 (0.25)
5	4-hydroxy-3-methoxyphenylpropionic acid (Dihydroferulic acid)	73	0.26 (0.19)	0.172	0.41 (0.23)
6	Hippuric acid	73	0.32 (0.25)	0.202	0.31 (0.26)
7	3-Hydroxyhippuric acid	74	0.27 (0.22)	0.224	0.40 (0.23)
8	4-hydroxy-3-methoxybenzoic acid (Vanillic acid)	73	0.12 (0.16)	0.463	0.48 (0.22)*
9	Panel 1 *(rows 1 – 8)*	71	0.71 (0.32)	0.032*	0.27 (0.24)
10	Panel 2 *(rows 1 2 3 4 5 6 7)*	71	0.79 (0.33)	0.020*	0.23 (0.25)
11	Panel 3 *(rows 1 2 3 5 6 7)*	72	0.81 (0.33)	0.016*	0.25 (0.24)
12	Panel 4 *(rows 1 2 3)*	72	0.93 (0.33)	0.005**	0.37 (0.22)
13	Panel 5 *(rows 2 3 6 7)*	74	0.7 (0.31)	0.026*	0.25 (0.24)

* $p < 0.05$; ** $p < 0.01$

β Coefficients are the unstandardised linear regression slopes of ΔFMD (%) regressed on the $\log_2$ FC of the independent variables with standard errors in parentheses. Row numbers next to the panels indicate the metabolite row numbers in this table to show the composition of each panel.

Assessing the relationship between the metabolite panels and ∆FMD

Panels 1 and 2:

Both *Panels 1 and 2* were similarly tested for an association with ∆FMD (compare rows 9 and 10 of Table 3-13). The exclusion of vanillic acid strengthened the association, as seen from the increased slope and decreased p value. Further candidates for exclusion were considered in the following.

Panel 3:

Although hippuric acid, 3-hydroxyhippuric acid, and dihydroferulic acid were only moderately correlated with ∆FMD, they were identified among the main metabolites after blueberry intake and it did not seem sensible to exclude these from the panel.

Syringic acid was not ranked among the top ten common metabolites and differed from the other metabolites by structure and probably also by origin as a direct fragment of malvidin glycosides rather than a series of catabolic conversions. The strength of association between syringic acid and ∆FMD was comparable to that of the hippuric acid and 3-hydroxyhippuric acid and in itself gave no reason to exclude syringic acid from the panel. However, a third panel was created excluding syringic acid to test the effect of a panel comprised only of microbial and end-stage metabolites (*Panel 3*, row 11). Panel 3 displayed an even stronger association than Panel 2, suggesting that indeed the microbial and end-stage metabolites were the main effectors. Nevertheless, the fold change analysis showed that syringic acid increased the most relative to baseline and thereby highlighted its status as a blueberry metabolite. Due to its direct, albeit not exclusive, link to malvidin, it was seemingly a strong indicator for the intake of blueberry anthocyanins in comparison to the other metabolites which may have derived from many different flavonoids. Therefore, for future use of this panel as a tool to identify high from low metabolisers, it seemed prudent to keep syringic acid in the panel, accepting that the association with ∆FMD was marginally weaker.

Panel 4:

A trend was becoming apparent that the relationship between the panels and ∆FMD was getting stronger the simpler the panels were. A fourth panel (*Panel 4*, row 12 of Table 3-13) was tested, comprised only of the three metabolites which individually had the strongest effects on ∆FMD (rows 1, 2, and 3). Indeed, the relationship was even stronger for this panel. Each doubling of the panel relative to baseline predicted an almost 1% increase in FMD. While it

may seem that this was the best panel so far in terms of simplicity and effectiveness, there are a few things to bear in mind:

- First, although ranked in the top ten for absolute change, two of the three metabolites ranked in the lowest ten ranks for fold change and in contrast to the other metabolites increased less than half as much relative to baseline. Only the third metabolite, 3,5-dihydroxyphenylpropionic acid, increased by a comparable amount.
- Second, the discrimination between a high and low metaboliser was inconclusive for 4-hydroxyhippuric acid (see Figure 3-14).
- Third, unfortunately 3,5-dihydroxyphenylpropionic acid also appears to be a biomarker for wholewheat and bran [241], hence while this panel was the simplest of all defined by only three metabolites, it was likely too narrow to be a specific marker of blueberry anthocyanin intake and even less so be able to distinguish between metaboliser profiles.

Panel 5:

A final panel was created with only the four microbial metabolites hippuric acid, 3-hydroxyhippuric acid, 3,5-dihydroxyphenylpropionic acid, and homovanillic acid (Panel 5, row 13 of Table 3-13). As explained in methods section (see Section 2.5), this was done because originally, the entire analysis up to now was performed with all available data including those with missed urines and some lower quality ultrasound sequences used to determine FMD measurements. The combination of these factors had increased the total number to n = 35 participants instead of n = 25 participants used in the full dose blueberry group metabolite panel analysis. The metabolite panel which resulted from it included only the four metabolites named above and was the final panel used for the prospective recruitment for the AMP study (see Chapter 5). It is included here to assess validity of its use when a more stringent data quality assessment was applied.

The process of elimination of metabolites used in the initial analysis was similar to the one described above with the following differences: 1) the spread of the data was determined through a visual inspection and not additionally supported by a statistical test and 2) the relationship with ∆FMD was determined in the full dose group only by comparing high and low groups as defined by the top and bottom 15% of the metabolites and composite panel, respectively. Despite the obvious limitation of having included ten potentially invalid observations, the common metabolite analysis and the comparison of the spread between full dose and placebo yielded five of the eight metabolites of the preliminary selection found using the refined approach (the four metabolites in Panel 5 and 4-hydroxyhippuric acid, **Table 3-14)**. The association of each metabolite and the combined panel with ∆FMD was assessed using

Pearson correlation and by comparing the high and low groups (n = 6) using the Mann-Whitney-U test. Although the correlation was moderate, it was significant at a level of 0.1 for all metabolites except 3-hydroxyhippuric acid and 4-hydroxyhippuric acid. Excluding either 3- or 4-hydroxyhippuric acid from the panel showed that excluding 4-hydroxyhippuric acid had no effect on the relationship of the panel with ΔFMD, whereas excluding 3-hydroxyhippuric acid did have a small effect. For this reason, a panel of four metabolites was formed, excluding 4-hydroxyhippuric acid. Table 3-13 clearly showed however, that 4-hydroxyhippuric acid was actually the metabolite with the greatest association with ΔFMD. The inclusion of likely invalid data and the analysis being underpowered by only testing within the full dose group therefore led to the false exclusion of 4-hydroxyhippuric acid in the initial analysis. To test the validity of the panel with all invalid or lower quality data removed, the regression analysis of ΔFMD on Panel 5 including the half dose and placebo groups is repeated here. The results indicate that despite the exclusion of 4-hydroxyhippuric acid, Panel 5 still showed a strong positive relationship with ΔFMD, similar to Panel 1 (see row 13 of Table 3-13).

Table 3-14. Association between ΔFMD and the initial metabolite panel

	Correlation[a] (n = 35)	**p value for high/low comparison[b] (n = 6 each)**
Hippuric acid	0.41 (0.017)**	0.126
Homovanillic acid	0.32 (0.073)*	0.485
3,5-dihydroxyphenylpropionic acid	0.29 (0.098)*	0.485
3-Hydroxyhippuric acid	0.21 (0.246)	0.429
4-Hydroxyhippuric acid	0.09 (0.593)	0.662
Initial Panel (4-OH-HA excluded)	0.41 (0.017)**	0.045

[a] Correlation values are Pearson correlation coefficients with p values in parentheses, n = 35. * p < 0.1; ** p < 0.05
[b] Mann-Whitney-U test

A ROC analysis reveals Panel 2 as the best candidate to detect blueberry intake

To further compare the predictive power of the different panels, a ROC analysis (receiver-operating characteristic) was performed. A ROC graph helps assess the diagnostic ability of binary classifier by determining the true positive rate (TPR, sensitivity) and false positive rate (FPR, 1 – specificity) for different thresholds of the classifier. In clinical testing this is often used to find the best threshold level specificity for a diagnostic test to identify a disease. Such a threshold would have the most appropriate TPR and FPR. A perfect diagnostic test would be able to discriminate perfectly between the two states and have a 100% TPR and 0% FPR (all true positives and no false positives are identified). In a ROC graph this corresponds to the point (0,1) in the top left corner. In contrast if TPR = FPR, then the test has no diagnostic ability to discriminate between the classes and is no better than a classification by pure chance. This case is represented in the graph by the diagonal line through the origin. Values above this line

implicate a better-than-random classification. The area under the curve (AUC) of ROC can be used to estimate the overall performance of a classifier, i.e., that it correctly distinguishes between the two classes, and also compare different screening tests for their diagnostic ability. Generally, the higher the AUC is, the better. A perfect test would have an AUC of 1.0 and a test unable to discriminate between classes would have an AUC of 0.5.

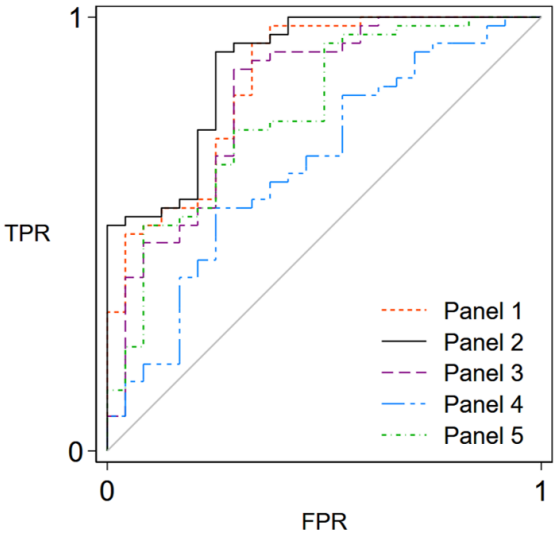
Here, the aim of the analysis was to test the ability of each panel to differentiate between 'blueberry treatment' and 'placebo' (expected AUC > 0.5) and which of the panels performed the best. As both the full dose and half dose group received blueberries, a one-way ANOVA was performed comparing the $\log_2$ FC of the panels between all treatment groups to determine if the full dose and half dose treatment groups could be treated as one 'blueberry' group in the context of the ROC analysis (**Table 3-15**). The results showed that the placebo group was highly significantly different from both the half dose and full dose groups with the exception of Panel 4, which only reached significance when compared to the full dose group. A dose-response trend was visible, however the difference between the half dose and full dose groups was not significant at a level of 0.05 (adjusted p-values > 0.37).

Table 3-15. Comparison of mean $\log_2$ FC of the metabolite panels by treatment group

	Placebo (n=24)	Blueberry half dose (1/2 cup) (n = 25)	Blueberry full dose (1 cup) (n = 25)
Panel 1	-0.43 [-0.78, 0.56][ac]	0.44 [0.22, 0.66]	0.66 [0.41, 0.8]
Panel 2	-0.38 [-0.7, 0.52][ac]	0.48 [0.27, 0.64]	0.71 [0.47, 0.8]
Panel 3	-0.31 [-0.7, 0.69][ac]	0.43 [0.2, 0.68]	0.62 [0.39, 0.74]
Panel 4	-0.35 [-0.93, 1.06][b]	0.14 [-0.12, 0.62]	0.37 [0.13, 0.67]
Panel 5	-0.31 [-0.78, 0.86][ad]	0.47 [0.22, 0.74]	0.66 [0.41, 0.8]

Values are means with 95% CI in square brackets. Letters denote significant differences between groups, p-values were adjusted for multiple comparisons using Tukey method. a = placebo compared with full dose, $p < 0.001$; b = placebo compared with full dose, $p < 0.02$; c = placebo compared with half dose, $p < 0.001$; d = placebo compared with half dose, $p < 0.002$. No significant difference at a level of 0.05 between half and placebo group for Panel 4 (all adjusted p values > 0.37).

For this reason, participants from both the full and half dose group were labelled as 1 for 'blueberry treatment' and those who received the placebo were labelled as 0. For every value of a panel (the classifier), the true positive and false positive rates were calculated and plotted (**Figure 3-15**). Unfortunately, this analysis could not directly answer whether a panel was useful in discriminating between a high and a low metaboliser of blueberry phenolic compounds, as no precedent exists to classify people as such. However, Figure 3-12 illustrated that the low group had very low changes in most metabolites, not so dissimilar to the placebo group. Hence in the extreme sense, assuming low microbial activity and absorption, a low metaboliser may even be approximated by the placebo group. The ability of a metabolite panel to identify metaboliser phenotypes is therefore likely to be somewhat but not enormously weaker to the ability to distinguish between blueberry and placebo intake.



	n	AUC	SE	95% CI
Panel 1	73	0.84	0.051	[0.74, 0.94]
Panel 2	73	0.88	0.044	[0.79, 0.97]
Panel 3	73	0.81	0.059	[0.69, 0.92]
Panel 4	73	0.65	0.071	[0.51, 0.79]
Panel 5	73	0.77	0.061	[0.65, 0.89]

Figure 3-15. ROC analysis of the metabolite panels

A ROC analysis was performed to test the effectiveness of each panel in differentiating between blueberry and placebo intake. This was used as an approximation to distinguish between high and low metabolisers based on the assumption that low metabolisers experience very low changes in metabolites following blueberry intake similar to the placebo group (as illustrated in Figure 3-12). For the analysis, the full and half dose groups were pooled and labelled as '1' for blueberry intake. The placebo group was labelled as '0'.
Of all panels, Panel 2 was the most successful in correctly classifying individuals as having eaten blueberries as indicated by the highest AUC of 0.88.
TPR = true positive rate; FPR = false positive rate; AUC = area under the curve; SE = standard error; CI = confidence interval.

All panels were better than random in classifying individuals as having eaten blueberries. Panel 4 was the least powerful of all panels to differentiate between the two classes while Panel 2 was the best. Due to the parsimonious composition of Panel 4 from just three metabolites, the result was not entirely unexpected and confirmed that Panel 4 was less specific for blueberry intake. The inset table in Figure 3-15 shows that the AUC increased from Panel 4 to Panel 2, suggesting that increasing the number of metabolites also increased the ability of the panel to discriminate between the classes. Interestingly though, Panel 2 which had excluded vanillic acid, performed slightly better than Panel 1 which was comprised of all eight metabolites. Hence, vanillic acid not only reduced the association of the metabolite panel with ΔFMD, but also reduced the classifying ability of the panel. The margin however was quite small and as vanillic acid has been previously reported as one of the main metabolites along with homovanillic acid and dihydroferulic acid after intake of anthocyanin-rich bilberries[122], a berry closely related to blueberries, this metabolite warrants continued investigation in future studies. The results from this dataset, however, suggested that Panel 2 was the strongest panel to use as a marker of blueberry intake and, in extrapolation, to use as a classifier for high and low metaboliser phenotypes. The reduced Panel 5 overall still had an acceptable predicting ability

with an AUC of 0.77, demonstrating its usefulness as a screening tool for the AMP study (Chapter 5) despite the exclusion of the key metabolite 4-hydroxyhippuric acid.

Summary of results from testing different metabolite panels

In summary, five panels were created from the eight metabolites identified as the strongest candidates to indicate a high versus low metaboliser phenotype. Each panel was tested for their ability to correctly identify the blueberry treatment, signifying its value as a marker of blueberry intake. Of the five panels, Panel 2 was selected as the most promising as it showed the highest performance in the ROC analysis and therefore was the most likely to be a marker of good response to blueberry intake. In addition, this panel was positively associated with vascular function as measured by FMD.

3.2.4 Additional exploration of Panel 2

Panel 2 did not significantly differ between the half and full dose groups

As seen in Table 3-15, Panel 2 (i.e., the $\log_2$ FC of Panel 2) was significantly greater in both the half dose and full dose group compared to placebo ($p < 0.001$ for both comparisons). Although the $\log_2$ FC was greatest in every panel for the full dose group, this was not significantly different from the half dose group for any panel at a significance level of 0.05 ($p > 0.37$). In terms of the FMD response, no treatment effect was observed for the half dose (**Figure 3-16A and B**, also see Figure 3 in Curtis et al (2019)[159]). This implies that the equivalent of 1 cup blueberries daily was required to see an effect on endothelial function and further suggested that although Panel 2 increased significantly in the half dose in comparison to placebo, a threshold level of metabolites was reached with 1 cup blueberries per day, but not with ½ cup. This emphasises the potential importance of the change in absolute metabolite amounts from baseline, thus the absolute amounts of Panel 2 metabolites at the 6-month endpoint were compared between half and full dose using Welch's t-test (**Table 3-16**). Interestingly, there was no strong evidence that any of the metabolites was in greater abundance in the full dose group than in the half dose group. Hippuric acid and syringic acid were the most likely to be in higher abundance in the full dose group, however the data was inconclusive (**Figure 3-16C and D**). In both cases, there was an individual in the full dose group who had excreted a much a higher amount than the rest. These participants were part of the group which have been identified as high metabolisers in the previous sections and further studies of these participants could help uncover the underlying factors that characterise high metabolise status. Plausible factors may include differences in microbiome composition,

more efficient absorption from the gut or metabolic adaptations which prolong the circulation of metabolites.

Table 3-16. Comparison of Panel 2 metabolites (absolute amount measured in 24h urine) at 6 months between full and half dose

Metabolite	n (full; half)	Mean (full)	Mean (half)	Mean difference [95% CI]	t statistic	p value
Hippuric acid	(25; 25)	3491.28	2749.36	741.93 [84.55, 1399.31]	1.625	0.113
Syringic acid	(25; 25)	1.64	1.05	0.59 [0.03, 1.15]	1.517	0.140
3,5-Dihydroxyphenylpropionic acid	(25; 25)	9.91	8.50	1.41 [-0.96, 3.77]	0.857	0.396
4-Hydroxyhippuric acid	(25; 25)	59.84	67.27	-7.43 [-24.79, 9.94]	-0.616	0.542
Dihydroferulic acid	(25; 25)	2.79	3.24	-0.45 [-1.55, 0.65]	-0.591	0.557
3-Hydroxyhippuric acid	(25; 25)	135.78	128.73	7.05 [-23.89, 37.98]	0.328	0.744
Homovanillic acid	(25; 25)	18.17	18.81	-0.64 [-4.67, 3.38]	-0.231	0.819

Mean values are in µmol/24h. Means compared using Welch's t-test

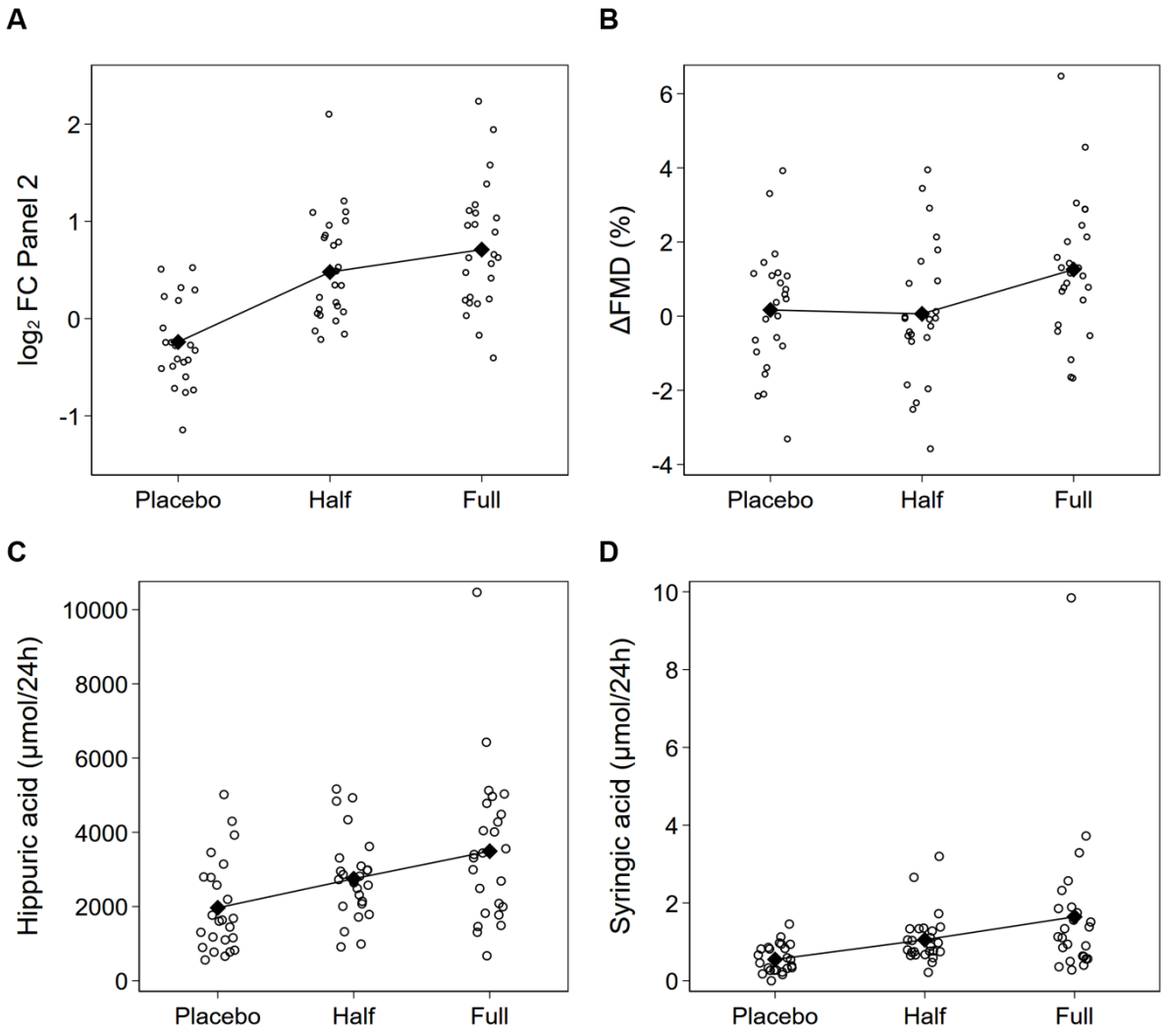


Figure 3-16. Comparison of treatment groups

Figures compare the (A) $\log_2$ FC of Panel 2, (B) ΔFMD, (C) hippuric acid at six months (endpoint), and (D) syringic acid at six months by treatment group (Placebo = 0 blueberries, Half = ½ cup of blueberries, Full = 1 cup of blueberries daily).
Although there was no significant difference in the $\log_2$ FC of Panel 2 between the half and full dose groups (A), no impact on FMD response was observed in the half dose group (B). This suggested that a threshold level of metabolites was reached with 1 cup of blueberries but not with ½ cup. However, there was no strong evidence that any of the Panel 2 metabolites was in greater abundance in the full dose group than in the half dose group at 6 months (C and D and Table 3-16).

$\log_2$ FC(Panel 2) is associated with cGMP

To further investigate the association with ΔFMD, the effect of $\log_2$ FC(Panel 2) on several biomarkers related to endothelial function was explored (**Table 3-17**). These include NO availability and cGMP as mediators of vascular smooth muscle cell relaxation and the inflammatory markers IL-6, TNF-α, CRP, soluble intercellular adhesion molecule 1 (sICAM-1), and soluble vascular cell adhesion molecule 1 (sVCAM-1), which are suspected to be involved

in the development of endothelial dysfunction[39]. Of the indirect markers for NO assessed in this study, only nitrite was included because nitrite, but not nitrate, was shown to reflect activity of the enzyme which produces NO (endothelial nitric oxide synthase (eNOS)).[244]

Table 3-17. Effect of Panel 2 on markers underlying endothelial function

Dependent variable	n	β (SE)	p	constant
ΔcGMP (pmol/ml)[a]	71	21.62 (8.6)	0.014*	-11.36 (4.47)*
ΔNitrite (μM)[a]	71	0.13 (0.06)	0.044*	-0.04 (0.05)
ΔIL-6 (pg/ml)[a]	61	0.01 (0.04)	0.803	0.01 (0.04)
ΔTNF-α (pg/ml)[b]	67	-0.09 (0.06)	0.106	-0.001 (0.04)
ΔCRP (μg/ml)[a]	71	-0.23 (0.15)	0.136	0.04 (0.09)
ΔsICAM-1 (μg/ml)[a]	71	0.001 (0.02)	0.951	0.01 (0.01)
ΔsVCAM-1 (μg/ml)[a]	71	-0.01 (0.02)	0.497	0.01 (0.01)

* $p < 0.05$
[a] robust regression using an MM-estimator due to presence of outliers [184]
[b] ordinary least-squares linear regression
β Coefficients are the unstandardised regression slopes of the dependent variable regressed on the $\log_2$ FC of Panel 2 with standard errors in parentheses.

A significant positive association was observed between Panel 2 and the change in cyclic GMP (cGMP) (β = 21.62, p = 0.014). A positive association was also seen between the panel and change in nitrite levels (β = 0.13, p = 0.044), which together with the positive relationship with cGMP, could suggest an increased bioavailability of NO with a greater fold-change in Panel 2 metabolites. The associations for ΔcGMP and Δnitrite are illustrated in **Figure 3-17**. There was some support for a slight reduction in TNF-α and CRP, although statistically inconclusive. However, this finding was in alignment with the previous observation in Section 3.1.6, where

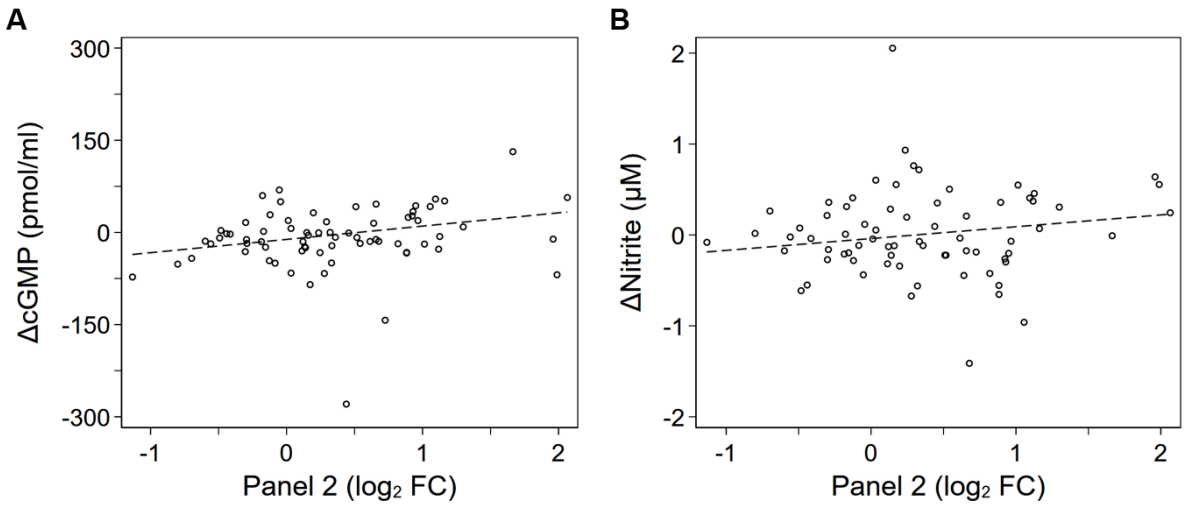


Figure 3-17. Robust regression analysis of ΔcGMP and Δnitrite regressed on Panel 2

Figure shows the association between ΔcGMP and Panel 2 (A) and Δnitrite and Panel 2 (B). Both ΔcGMP (β = 21.62, p = 0.014) and Δnitrite (β = 0.13, p = 0.044) were significantly positively associated with Panel 2 at a significance level of 0.05. Taken together, the positive relationships could suggest that a greater fold-change in Panel 2 metabolites is aligned with an increased bioavailability of NO. cGMP = cyclic guanosine monophosphate; NO = nitric oxide.

high metabolisers seemingly had lower levels of TNF-α, CRP, and IL-6. Conversely, no evidence for an effect was observed here for IL-6, sICAM-1, or sVCAM-1.

3.3 Discussion

It has been established that a large inter-individual variability in response to flavonoid intake exists, meaning individuals are exposed to different circulatory levels of bioactive compounds, potentially affecting the magnitude of the health benefit they receive[204]. Inter-individual variability has been observed in randomised controlled trials for a range of dietary flavonoids and phenolic compounds, e.g, anthocyanins[219,245], flavanones[246], flavan-3-ols[158], isoflavones[205], ellagitannins[247], and lignans[248]. In the literature, individuals exhibiting high levels of flavonoids and phenolic catabolites following an intervention are sometimes termed 'responders', and those with low levels 'non-responders'. Whilst the metabolism of some flavonoids results in unique catabolites (e.g., equol from isoflavones) which allows a clear-cut stratification between participants as producers or non-producers, the metabolites from berries are generally small molecule phenolic acids also common to other flavonoids. Therefore, a response to berry consumption rather reflects a significant increase in metabolites which are already present at baseline[249]. The underlying hypothesis of the study presented in this chapter was that inter-individual variability in response to blueberry anthocyanins mediates the potential cardiovascular benefits a person receives from eating blueberries. The aim of the assessment was to firstly identify a group of metabolites which may be representative of a certain metabolic phenotype ('metabotype'), then secondly, to identify high metabolisers (i.e., responders) and low metabolisers (i.e., non-responders) among those who received the blueberry treatment, and finally, to investigate whether the metabotype mediated differential clinical responses.

The key findings from this exploratory analysis were:

(1) the metabolites appeared to correlate by absorption location;

(2) there was a differential metabolic response among those who received the full dose blueberry treatment;

(3) an improvement in vascular function (as measured by FMD) to blueberry intake was seemingly more associated with colonic microbial phenolic metabolites than metabolites absorbed in the small intestine;

(4) a greater improvement in FMD was observed for high metabolisers versus low metabolisers, however this was statistically inconclusive;

(5) a metabolite panel which characterised a high metaboliser consisted of seven metabolites (4-hydroxyhippuric acid, 3-hydroxyhippuric acid, hippuric acid, syringic acid, homovanillic acid, dihydroferulic acid, 3,5-dihydroxyphenylpropionic acid) and was significantly associated with change in FMD across the three intervention groups;

(6) the association of the metabolite panel with FMD was potentially mediated by an increase in cGMP levels and NO availability.

In metabolomics, subgroups of metabolically similar individuals are frequently identified through a combination of different multivariate analyses including dimension reduction, clustering and classification techniques.[186] Commonly, the first step is a dimension reduction of many metabolites to a smaller set of variables. The technique used here was exploratory factor analysis (EFA), a technique similar to the widely used PCA, but with a fundamental distinction. PCA identifies components which are uncorrelated linear composites of the observed variables and aims to preserve the maximum variance in the data. It is therefore useful if dimension reduction is the only goal. In contrast, EFA uses only the shared variance to identify an underlying, unobserved variable which cannot be measured directly. Therefore, EFA was specifically chosen, as it was better suited for the interpretation of the metabolite pattern based on an underlying factor, which in this case was hypothesised to be a metabolic phenotype.

Using EFA, two factors were extracted. The first factor (factor 1) included cinnamic and benzoic acid derivatives and their glucuronide and sulfate conjugates, while the second factor (factor 2) was comprised of phenylpropionic acids, phenylacetic acids, benzoic acid and hippuric acids. As detailed in Section 3.1.4, these factors were postulated to loosely reflect absorption location. Many of the factor 1 metabolites, such as ferulic acid and vanillic acid have been observed to appear shortly after the ingestion of anthocyanins which is suggestive of absorption from the small intestine[128]. Factor 1 was therefore termed ‘early-stage’ and could be considered as a group of metabolites which consisted mainly of dietary phenolic acids (such as ferulic acid, chlorogenic acid, and vanillic acid) and in part initial degradation products of anthocyanins (such as protocatechuic acid and syringic acid), as well as their phase II metabolites (sulfated, glucuronidated, methylated conjugates). A substantial proportion of anthocyanins pass the small intestine unchanged (40-85% depending on the sugar moiety) and reach the colon unchanged where they are subject to the metabolism of the colonic microbiota.[130,250,251] This involves deglycosylation, cleavage of the C-ring, dehydroxylation, hydrogenation, demethylation, α- and β-oxidation, resulting in various lower molecular weight phenolic catabolites derived from the A- and B-rings. Although some microbial transformation of metabolites occurs also in the small intestine, the major site of microbial degradation of phenolic compounds is in the colon.[157] Factor 2 was therefore hypothesised to represent

metabolites of colonic microbial origin and termed ‘late-stage’ to reflect the absorption time of >5 h after ingestion.[128]

The two factors were correlated (r = 0.5) which is likely due to their connection through metabolic pathways. Indeed, the classification as early or late should be qualified as multiple phenolic acids appear in a biphasic pattern after ingestion of anthocyanins. This can be illustrated through a chain of reactions in the colon which includes a progressive shortening of the carbon sidechain by the microbiota. Hydroxycinnamic acids, which fall under factor 1, are converted to hydroxyphenylpropionic acids (factor 2) through microbial hydrogenation, which are in turn progressively converted to hydroxybenzoic acids (factor 1 metabolites) via α- or β-oxidation.[128] Absorbed hydroxybenzoic acids are then glycinated to hydroxyhippuric acids (factor 2) and excreted via urine.

This close relationship between the factors can be further explained by the EFA model itself. When conducting the EFA, careful considerations had been made regarding data pre-treatment. This included pre-grouping the metabolites by chemical structure, which was both a strength and a limitation. The grouping was necessary for the model to run (without the grouping, the factor analysis did not converge due to the small sample size and large number of metabolites), increased the stability of the factor solution by reducing the impact of sample size, and identified factors likely representing the general absorption site of metabolites. On the other hand, the consequence and limitation was that information regarding the individual metabolites was lost. This caused metabolites such as vanillic acid, which was observed to be a major urinary metabolite following anthocyanin intake in our study and also in other studies,[122,123] to be grouped together with minor metabolites such as 2-hydroxy-4-methoxybenzoic acid, which barely increased following anthocyanin intake. Thus, biological inference from the results could only be made in regard to the groups of metabolites. This limitation was rectified with a detailed follow-up analysis, investigating the metabolite profiles of high and low metaboliser groups identified from a cluster analysis following the EFA, discussed further below.

Importantly, when including change in vascular function as measured by FMD to the factor analysis, FMD clearly associated with factor 2 (see Figure 3-6), which implied an association between FMD and the abundant colonic metabolites. This is in agreement with the hypothesis that lower molecular weight phenolic catabolites are likely the mediators of anthocyanin bioactivity rather than their parent compounds[123] and was further supported by the fact that in our study FMD significantly increased in the full dose group compared with the half and placebo groups, but no parent anthocyanins were detected in urine or serum in our study.[159]

The FMD association with factor 2 was also observed in a correlation analysis between change in FMD and the factor scores (see Figure 3-7). Surprisingly, although in the factor solution FMD

did not load on Factor 1 and in fact had a weak negative loading, a mild correlation with FMD was observed when excluding a participant who scored very highly on factor 1 (Figure 3-7A). This correlation was almost entirely driven by chlorogenic acid, a compound which has been linked to vasoactive effects in animal and *in vitro* studies, possibly by increasing NO bioavailability by decreasing oxidative stress,[252,253] but which has been associated with equivocal findings in human studies. For example, whilst studies in young (mean age 24 years), healthy, adults have shown acute biphasic increases in FMD at 1 and 5 hours after consuming coffee containing 89 or 310 mg doses of chlorogenic acid (increase at 1h: 1.10% and 1.34%; increase at 5h: 0.79% and 1.52%)[254], these favourable findings were not reproduced in older adults (mean age 60 y) with a diverse vascular status.[255] Likewise, 8 weeks of coffee consumption did not affect FMD in healthy adults with a mean age of 38 y [256].

In this chapter, a composite measure which quantified the average amount of change across grouped metabolites (within a factor) was generated. This assessment showed that factor 2 metabolites were in much greater abundance than factor 1 metabolites. The absolute amount in µmol was 280-fold higher for factor 2 metabolites than factor 1, and the relative increase in factor 2 metabolites, compared with baseline levels, was approximately 2-fold greater than for factor 1. This was easily explained when considering that most of the anthocyanins pass to the colon unchanged and the conversion of factor 1 metabolites towards factor 2 metabolites (for example hydroxycinnamic acids to hydroxyphenylpropionic acids and hydroxybenzoic acids to hippuric acids). Hippuric acid contributed the majority of the absolute increase of factor 2 metabolites from baseline, at 90% of the total change observed. The scope of this change was in agreement with a raspberry study by Ludwig et al which fed 300 g of berries, containing 292 µmol anthocyanins, and observed that hippuric acid made up 85% of increase in all anthocyanin and phenolic acid derivatives.[219]

Hippuric acid is endogenously produced and about 1-2 uM is excreted daily in urine.[222] Because it is likely a common end metabolite of multiple polyphenols[257] and can be derived from different endogenous and dietary sources, including quinic acid and aromatic amino acids in proteins, it is difficult to determine the exact dietary origin. However, a ^{13}C-tracer study has confirmed that hippuric acid is indeed a metabolite of anthocyanins[123], although the amount recovered in urine of 30% of total recovered anthocyanins and phenolic acids was substantially lower than observed in our study or in the raspberry study by Ludwig et al mentioned above. On the basis of this data, it seems likely that a meaningful proportion of the hippuric acid observed in our study may have been from alternative sources. Nevertheless, in the present study as well as in the raspberry study, hippuric acid significantly increased with berry intake (over and above the placebo group that followed the same dietary guidance and were expected to have also metabolised hippuric acid from other sources) and has been proposed as a

biomarker for fruit and vegetable intake.[240] It therefore still retains value as a marker of intake and relative increase especially when compared to a placebo, but is less likely to be informative when assessing absolute bioavailability of flavonoids.

Inter-individual variability in urinary excretion of metabolites was observed for both factors 1 and 2. Plotting the ranks of the factor composite measures against one another revealed that individuals clustered in four groups (see Figure 3-10): (1) a group which excreted limited amounts of both factor 1 and factor 2; (2) a group which excreted medium amounts of factor 1 and limited amounts of factor 2; (3) a group which excreted medium amounts of factor 2 and limited amounts of factor 1; (4) a group which excreted high amounts of both factor 1 and 2. Those in group (4) were classified as 'high' metabolisers, while those in group (1) were classified as 'low' metabolisers.

Between-subject variability is affected by the genetic variability in the enzymes, transporters, and carrier proteins as well as the gut microbiota involved in the absorption and metabolisation of phenolic compounds.[204] As the majority of anthocyanins reach the colon intact, the differences in metabolisation observed in the present study could partially be ascribed to variations in composition of the gut microbiota.

A bidirectional interaction between the gut microbiota and polyphenols has been observed in a few human intervention studies. A probiotic containing *Bifidobacterium longum* taken daily for four weeks significantly increased the bioavailability of flavanone metabolites and phenolic acids after consumption of orange juice by 5% and 22%, respectively.[258] Notably, these effects were only observed after chronic intake, but not acute after ingestion of a single bolus of the probiotic. On the other hand, dietary compounds are key factors in the composition of the gut microbiota.[259] Prebiotic effects, i.e., induction of growth of beneficial bacteria, was observed after ingestion of anthocyanins from 272 ml red wine for 20 days.[260] Boto-Ordóñez et al reported a mean increase of 50% from baseline for *Bifidobacterium* in the highest tertile of change and a significant correlation between increases in Bifidobacterium and four anthocyanin metabolites (mean change from baseline in tertile of highest increase of *Bifidobacterium*): 4-hydroxybenzoic acid (18.04 µmol per 24 h), syringic acid (1.37 µmol per 24 h), *p*-coumaric acid (2.05 µmol per 24 h), and homovanillic acid (20.28 µmol per 24 h). Moreover, there is some evidence that an alteration of the gut microbiota is associated with obesity and metabolic syndrome.[259] Moreno-Indias et al found that participants with metabolic syndrome differed in the gut microbial composition from healthy volunteers at baseline, but not after the intervention. Red wine intake for one month increased the levels of beneficial bacteria including *Bifidobacterium* (55% mean increase in $\log_{10}$ copies per gram faeces), *Lactobacillus* (56% increase), *Faecalibacterium prausnitzii* (36% increase), and *Roseburia* (28% increase), and decreased the numbers of bacteria associated with systemic inflammatory marker LPS

(lipopolysaccharide) including *Escherichia coli* and *Enterobacter cloacae* (both about 20% decrease).[261] The authors also reported improvements in biomarkers for metabolic syndrome following red wine intake, including blood pressure (-10 mmHg for DBP and SBP), blood glucose (-32 mg/dl), HDLC (-15.6 mg/dl), and LPS (-0.17 endotoxin units/ml) which could be related to the observed changes in the microbiota.

Importantly, while the composition of the gut microbiota can be modulated by dietary polyphenols, the effect on different people will vary due to variations in the gut microbiota between individuals.[262] In the present study, it could be speculated that changes to the microbiota may plausibly underpin the changes in metabolites in factors 1 and 2 – a hypothesis that requires confirmation outside of this thesis. That accepted, it could be hypothesised that a participant scoring high on both factors could have experienced an increase in certain bacterial species, such as *Bifidobacterium* and *Lactobacillus*, and increased microbial activity along the entirety of the GI tract, while the gut microbial composition of a low metaboliser may be less affected.

In support of this, the metabolites identified as characteristic of a high metaboliser (i.e., high in both factors) included three of the four metabolites which Boto-Ordóñez et al reported to be associated with an increase in *Bifidobacterium*: i.e., 4-hydroxybenzoic acid, syringic acid, and homovanillic acid. In the present study, 4-hydroxybenzoic acid was not observed directly as a major urinary metabolite, however its glycinated phase II metabolite 4-hydroxyhippuric acid was. Also, *p*-coumaric acid (4-hydroxycinnamic acid) was not observed to be a major metabolite in this study, but could potentially have been degraded to 4-hydroxybenzoic acid and excreted as hippuric acid.

In human studies, most metabolites within the high metaboliser profile (i.e., hippuric acid, 4-hydroxyhippuric acid, 3-hydroxyhippuric acid, syringic acid, homovanillic acid, and dihydroferulic acid) have been frequently observed as major metabolites following anthocyanin intake.[122,123,218,219,225,226,245,263,264] In contrast, 3,5-dihydroxyphenylpropionic acid is reported less frequently, however in the cited publications above, this compound was also not assessed. Related compounds, however, have been reported. 3,5-dihydroxybenzoic acid was observed in urine following anthocyanin intake from elderberries [245] and *in vitro* fermentation of raspberry anthocyanins found a significant accumulation of resorcinol (benzene-1,3-diol)[242]. This is relevant because 3,5-dihydroxyphenylpropioinic acid could potentially be derived from the dehydroxylation of the anthocyanin delphinidin, which has a trihydroxylated B- ring[265] and subsequently converted to 3,5-dihydroxybenzoic acid through microbial β-oxidation and then to resorcinol through decarboxylation.

A positive relationship was observed between change in FMD and the metabolite panel created from the metabolites named above. A mechanistic insight for this relationship may be gained

from the positive association between the metabolite panel and changes in cGMP as well as nitrite. NO bioavailability is a key determinant of endothelial function and arterial vasodilation is largely mediated through the NO/cGMP pathway. NO is produced by the endothelial nitric oxide synthase (eNOS) in neighbouring endothelial cells. Its downstream effects are mediated by cGMP in the vascular smooth muscle cells which include smooth muscle relaxation and thereby blood vessel dilation.[39] Nitrite, a product of NO metabolism, has been found to reflect eNOS activity,[244] therefore the increase in nitrite observed in this study could indicate an increased NO bioavailability in high metabolisers, which in turn would lead to an elevation in cGMP levels and improvement in FMD. These results were in line with another study which gave a similar amount of anthocyanins in purified form (320 mg) for 12 weeks.[90] The authors reported a significant increase in FMD and cGMP in the anthocyanin group compared to placebo (2.9% and 15 pmol/ml, respectively), as well as a positive correlation between FMD and cGMP ($r = 0.428$, $p < 0.001$). In addition, the authors observed in an acute sub-study that the effects of anthocyanins on FMD were blocked by the presence of a NOS inhibitor. In the same paper they also investigated the effect of anthocyanins on vasorelaxation with and without a NOS inhibitor in an *in vitro* model and in sum concluded that anthocyanins act through the NO/cGMP pathway.

Next to affecting eNOS activity, another possible pathway for an increase in NO bioavailability through blueberry anthocyanins is through their antioxidant properties. NO is a highly reactive radical which rapidly reacts with reactive oxygen species (ROS) to form peroxynitrite. Thereby decreased oxidative stress (i.e., decreased levels of ROS) through inactivation of ROS or modulation of enzyme expression and activity increases NO concentration. Kuntz et al reported that the consumption of 330 ml of an anthocyanin-rich fruit juice and smoothies.from red grapes and bilberries containing 840 and 983 mg anthocyanins improved antioxidant status compared to an anthocyanin-low placebo. The authors observed enhanced activities of superoxide dismutase by 6% and of catalase by 21%, enzymes involved in the antioxidant defence of cells.[266] Direct scavenging of ROS has also been shown in in vitro studies for anthocyanins as well as phenolic acids, such as syringic acid [267,268].

Furthermore, there was an indication that the inflammatory markers TNF-α and CRP might be inversely associated with the metabolite panel. As both inflammatory markers affect NO bioavailability by inhibiting eNOS activity, this could be a further indication that NO bioavailability was increased.[238,269] Anti-inflammatory effects of anthocyanins have been demonstrated in several human studies. Three to four weeks of intervention with an anthocyanin-rich product like wine or sweet bing cherries has been reported to reduce CRP levels by 20 – 35 % [261,270,271] and is potentially linked to a reduction in plasma levels of LPS through an increase in *Bifidobacterium*, a species which protects the intestinal barrier.[261] TNF-

α plays a key role in the progression of endothelial dysfunction and induces the gene expression of pro-inflammatory mediators, such as IL-6 and the cell adhesion molecules sICAM-1 and sVCAM-1.[238] There are studies which have observed a reduction in these markers in response to different doses of anthocyanin supplementation. Zhu et al reported a decrease of 11.6% in sVCAM-1 after 12 weeks of 320 mg anthocyanin supplement.[90] Barona et al fed grape powder containing 35 mg anthocyanins for 30 days and observed a 6% reduction in sICAM-1, but no change in sVCAM-1.[198] Vugic et al reported a 50% reduction in IL-6 in individuals with obesity after supplementing 320 mg anthocyanins for 28 days.[272] However, not all studies report an effect. Dohadwala et al did not find any impact on sICAM-1 or CRP in patients with coronary artery disease after four weeks of cranberry juice containing 94 mg anthocyanins.[84] In a study by Stull et al, even a much larger dose of blueberry anthocyanins (668 mg) taken daily over six weeks by men and women with obesity did not affect inflammatory markers.[87] Equally, in our study no association was found between the metabolite panel and IL-6, sICAM-1 or sVCAM-1. This could indicate that the improvement in endothelial function was independent of significant changes to the inflammatory state and more due to increases in NO bioavailability.

Interestingly, while the metabolite panel in both the half dose group (182 mg anthocyanins in ½ cup of fresh blueberries) and the full dose group (364 mg anthocyanins in 1 cup of fresh blueberries) was significantly greater than placebo, the change in the full dose group was only marginally greater than the half dose group. On the other hand, a treatment effect was only seen for the full dose group (see Figure 3-16), which implied that a threshold was reached with 1 cup, but not ½ cup of blueberries. However, at 6 months, none of the individual metabolites in the panel were significantly different between half dose and full dose, only hippuric acid and syringic acid showed an indication of being more abundant in the full dose group (see Table 3-16). As discussed earlier, an increase in *Bifidobacterium* was reported to be associated with an increase in syringic acid. It could therefore be that the observed threshold effect was localized in the gut, with a higher dosage of anthocyanins inducing a greater modulation of the gut microbial composition and thereby impacting on the interaction between the gut and host physiology. Further mechanisms which may explain the connection between the modulation of the gut microbiota and subsequent health effects could include the involvement of the gut microbiota in regulation of metabolic processes such as nutrient absorption, energy storage, regulation of cholesterol metabolism and insulin sensitivity, and inflammatory responses.[259] Indeed, a further interesting finding was that one of the clusters (cluster 3, see Figure 3-10D) had the lowest FMD response of all clusters although these individuals had excreted higher amounts of factor 2 metabolites, which had been observed to be associated with FMD. The fact that cluster 3 exhibited a lack of factor 1 metabolites, but not factor 2, may point to other aspects of metabolism, such as the composition of the microbiota in the small intestine. This

could affect other dietary components absorbed in the small intestine, particularly macronutrients such as lipids,[232] which may have knock-on effects on metabolic processes and vascular function.

Concluding remarks and outlook

In this analysis, a group of metabolites was identified which characterised a high metaboliser of blueberry anthocyanins and which was associated with an improvement in vascular function. Careful considerations regarding data treatment were made to increase the chance of a reliable and meaningful result despite the limited sample number which is reflected in the large agreement between the results of this study and the major urinary anthocyanin metabolites found in other anthocyanin intervention studies. Nevertheless, due to the exploratory nature of the study however, this metabolite panel should be considered as putative and the stability of the metabolite panel as well as its effect on vascular function should be validated in further studies. Also, the population in this study was aged 50 -75, had metabolic syndrome, and was predominantly white and male. The transferability of the metaboliser profile determined in this study to other age groups, females, other ethnicities and health status remains to be confirmed in future studies. Finally, a fairly new realisation is the impact of day-to-day variation in the metabolism of flavonoids. Intra-individual variability was recently reported to be mainly driven by the gut microbiota[273] and should be considered when exploring the existence of metabotype.

To my knowledge, this study was the first assessment of a potential metabotype based on anthocyanin metabolites and could provide an exciting premise for future anthocyanin interventions. The metabolite panel, which is based on change relative to baseline, could be trialled in future studies for prospective recruitment on the basis of metaboliser status, similar to studies over the last few decades which have prospectively recruited participants on the basis of confirmed equol and urolithins producer statuses.

4 Methods pertaining to the assessment of anthocyanin metaboliser profile for the AMP study (Chapter 5)

4.1 Study aims and general study design

The AMP study was an acute, placebo-controlled, dietary cross-over study in adults with overweight or obesity. The principal aim was to test whether differential metaboliser types mediated cardiometabolic responses; thus, a primary aim was to prospectively recruit participants based on their metabolism profiles following a 'blueberry challenge'. Subsequently an energy-dense test meal was consumed with or without blueberries, and cardiometabolic responses assessed over a 48h period. The study aimed to determine to what extent the metaboliser profile mediated cardiometabolic effect; both in the postprandial phase, and over an extended 48 h period. Previously, a ^{13}C-tracer study showed that anthocyanin metabolites remain in circulation up-to 48 h after consumption[140], therefore the AMP study aimed to establish if the presence of these metabolites had cardiometabolic consequences.

The AMP study was split into two phases. The general study design is shown in **Figure 4-1**.

The first phase, the 'blueberry challenge' (see Section 4.2; including challenge regimen and anthocyanin and phenolic composition), was a screening phase to confirm anthocyanin metaboliser status based on the metabolite panel developed from the retrospective analysis of the CIRLCES study data (Chapter 3 and described in detail in Section 3.2.3). For 5 days prior to the blueberry challenge, participants were required to adhere to dietary restrictions to control the extent to which the background diet (including flavonoids) affected the metabolism profile; this was continued throughout the 48h urine sampling period of the 'challenge'.

Those assessed as either 'high-' or 'low-metabolisers' proceeded to the second phase of the AMP study, which consisted of a cross-over designed, double-blind, intervention study (with random allocation to treatment order). Consistent with the blueberry challenge, participants adhered to the same dietary restrictions for a period of 5 days prior to the study assessments (and throughout the 48h assessment period). Additionally, all food and drink were provided for four-day periods (two days before and two days after the energy-dense meal). In each of the two test meal periods, an energy-dense test meal was consumed which was adjusted to an individual's resting BMR (age, sex, and bodyweight component characteristics of the calculation) and estimated physical activity level. 50% daily energy requirements were provided by the test meal, of which 40% of calories came from fat. The energy-dense meal was accompanied by either a blueberry or placebo milkshake (see Section 4.2). Prior to the test meal consumption and subsequently for the 48h following, vascular function and a range

of cardiometabolic markers were observed. Additionally, extensive biological sampling was undertaken, including repeat 24h urine samples (collected for one day before the test meal, and the 48h follow-up period) and blood sampling (via a cannula on assessment days in the clinic, and fasted blood samples at 24 and 48h timepoints). As shown in Figure 4-1, the two assessment periods were separated by a wash-out of at least one week.

NB: only the blueberry challenge and determination of metaboliser groups is part of this thesis; the analysis of the second phase was the topic of another PhD thesis.

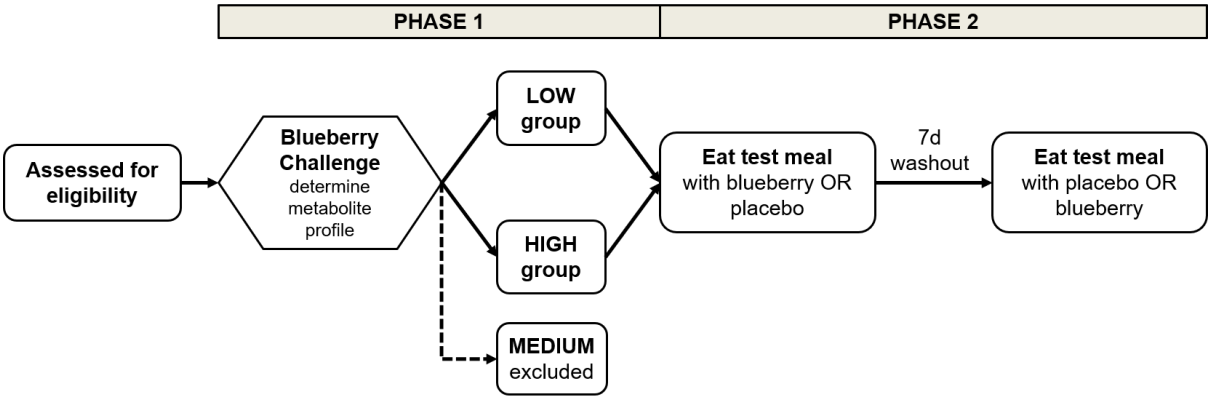


Figure 4-1. General AMP study design

The AMP study was an acute, placebo-controlled, dietary cross-over study in adults with overweight or obesity. It consisted of two phases: participants were prospectively recruited based on their metabolism profiles following the blueberry challenge in phase 1 of the study. In phase 2, an energy-dense test meal was consumed with or without blueberries, and cardiometabolic responses were assessed over a 48h period.

4.1.1 Study population

In total, 121 men and women aged 50 – 80 years completed the blueberry challenge and were assessed for metaboliser profile. Participants were recruited through local advertisement and GP practice involvement and were: non-smokers, generally healthy, but with overweight or obesity (BMI ≥ 25 kg/m^2). The study flow is presented in **Figure 4-2**.

Eligibility to participate in the study was assessed through 1) the administration of a health and lifestyle questionnaire, collected after initial expression of interest to participate, and 2) a fasted clinical screening of general health (assessed through clinical chemistry analysis following blood sampling), which was conducted at the clinical research facility based at the Quadram Institute, Norwich, by study research nurses. Exclusion criteria included existing clinical diagnosis of cardiovascular disease, diabetes or cancer (excluding basal-cell carcinoma), as well as prescribed hypoglycaemic or anti-hypertensive medication, hormone replacement therapy, or the intake of supplements containing flavonoids, nitrate, nitrite or fish oil.

The rationale to recruit generally healthy participants who were at least overweight, was because body size has been reported to be an indicator for cardiovascular disease risk.[274] There is evidence from population-based studies and human intervention trials that dietary flavonoids, and in particular the subclass of anthocyanins, may improve cardiometabolic health in terms of reducing risk for hypertension[70,108] and T2DM[200], reducing arterial stiffness[70,84], improving insulin sensitivity[87,99], increasing HDLC levels[89,159], and improving vascular function[90,91,159] (more details on health effects in the introduction, Section 1.3). Therefore, including more anthocyanins in one's habitual diet may provide a simple strategy to address the cardiometabolic comorbidities of both overweight and obesity.

Data from n = 2 participants was excluded (n = 1 due to violation of dietary restrictions, and n = 1 due to intolerance to the milkshake), leaving a sample size of n=119 described in this chapter (see Figure 4-2).

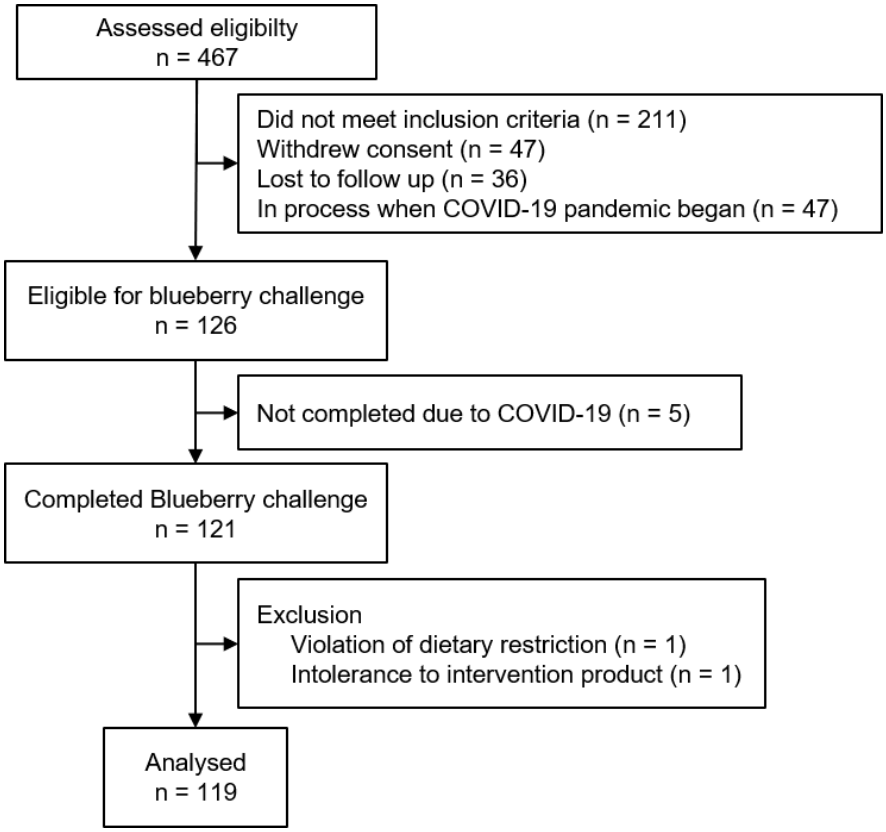


Figure 4-2. Study flow diagram

Flowchart shows the recruitment and retention in the study.

4.2 Design of the blueberry challenge and intervention drink

The study flow of the blueberry challenge is shown in **Figure 4-3**. The blueberry challenge was designed to deliver a single dose of blueberries, provided in the form of a milkshake, with the anthocyanin metabolite profile assessed from urinary analysis. In broad terms, the metaboliser profile was assimilated by comparing the cumulative excretion of urinary metabolites from a previously determined metabolite panel for 48h following the blueberry milkshake, versus a baseline 24h-urine sample collected prior to milkshake intake. From these data, participants were characterised as 'Low', 'Medium', or 'High' metabolisers based on their profile. The metabolites and method used to determine metaboliser groups is described below.

The milkshake contained 36 g of freeze-dried blueberry powder (equivalent to 216 g or 1.5 cups of fresh blueberries) mixed with 500 g of semi-skimmed milk, providing 382 mg of anthocyanins and 1116 mg total phenolic acids. The milkshake was prepared immediately prior to consumption in an opaque container and consumed under supervision (in the cross-over test meal phase, this was done by an unblinded researcher who did not undertake any cardiometabolic assessments). The container was rinsed twice with water to ensure all of the intervention product had been consumed. 500 g of semi-skimmed milk was used to dissolve the powder to make the drink more palatable and smoother for drinking. The same milkshake composition was used within the test meal in the second phase of the study (not covered in this thesis).

The intervention dose for the blueberry challenge (36 g freeze-dried blueberries; providing 382 mg anthocyanin) was chosen to provide a similar amount of anthocyanins as used in the CIRCLES study (i.e., 364 mg anthocyanins, see Section 2.1) as this was the basis for the metabolism profile assessment. Consequently, this amount of anthocyanins was anticipated to induce a comparable excretion of metabolites as in the CIRCLES study. Of note, the serving size was smaller in the CIRCLES study due to differences in anthocyanin abundance between the two harvested crops prior to homogenisation as a freeze-dried material. In terms of the appropriateness of this anthocyanin dose and potential cardiometabolic responses (in the second phase of the AMP study), other single-dose intervention studies have reported vascular improvements as measured by FMD with similar doses of anthocyanins. Zhu et al observed in hypercholesterolemic men and women with an average BMI of 26.6 kg/m^2 an increase in FMD from 8.3% at baseline to 11.0% at 1h after intake and 10.1% at 2h after intake of 320 mg purified anthocyanin supplement.[90] In healthy, slightly leaner men, Rodriguez et al reported increases in FMD of similar magnitude (2.4% at 1h, and 1.5% at 2h) following the consumption of a blueberry drink containing 310 mg anthocyanins.[91]

As summarised earlier, throughout the blueberry challenge (five days prior and during the 48h after blueberry milkshake intake (seven days in total)) participants were instructed to adhere to dietary restrictions to a) control the intake of anthocyanins, and b) to reduce the background levels of food sources which are metabolised into similar phenolic acids as anthocyanins. The dietary restrictions prohibited anthocyanin-rich fruits, vegetables, and beverages or foods containing anthocyanins (such as strawberry jam), and extended to other food sources, which are commonly consumed and rich in polyphenols (see **Table 4-1**). As indicated by the Phenol Explorer on polyphenol content in foods[199], this included, for example, chocolate and tea (rich sources of flavan-3-ols); citrus fruits (rich source of flavanones); and coffee (rich source of chlorogenic acids).

Compliance to dietary restrictions in the five days prior to the blueberry milkshake was assessed at the clinical research facility on the day of the intervention by a face-to-face questionnaire covering all foods on the restriction list. Adherence to dietary restriction during the urine collection days was monitored using a 3-day food diary. These were thoroughly checked for any foods on the restricted list. Compliance to dietary restrictions was very high; only n = 2 participants violated the dietary restrictions, of which n = 1 participant chose to repeat the blueberry challenge.

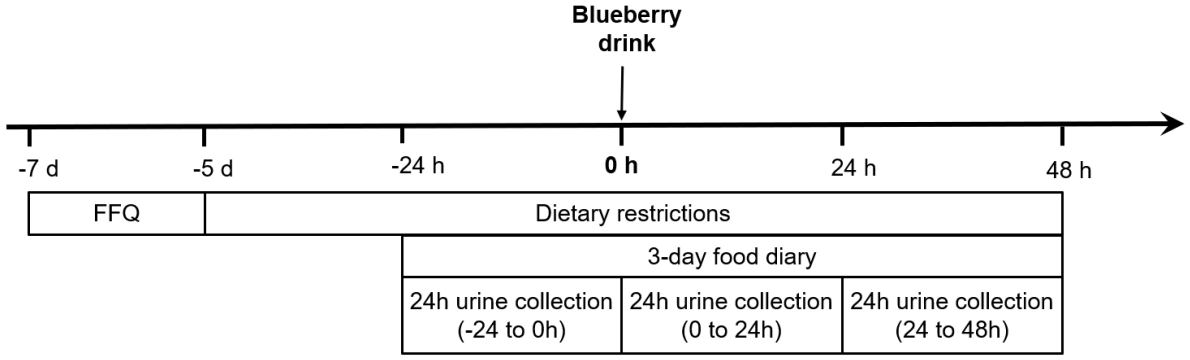


Figure 4-3. Design of the blueberry challenge

The blueberry challenge consisted of five components: a blueberry drink on the day of the study visit; a food frequency questionnaire, assessed one week before the study visit (see Section 4.6); dietary restrictions spanning from five days before the study visit to two days after; urine collection for three days (one day before the blueberry drink and for two days after); and a three-day food diary on the urine collection days.
FFQ: food frequency questionnaire.

Table 4-1. Table of dietary restrictions as given to participants

	Foods to avoid	**Alternatives examples**
Any drinks, salads, sandwiches and products (e.g., ready meals, pizza, casserole) containing:		
Fruit	**Berry fruits:** Blueberries, blackberries, raspberries, cranberries, strawberries, black currants, gooseberries, elderberries, aronia (chokeberries), acai, lingonberries **Other:** Cherries, plums, red/purple grapes, citrus fruits (oranges, grapefruit, lemon, lime)	Bananas, apples, green grapes, figs, kiwi, melon, pineapple, peaches, avocado, pears
Vegetables	**Red / purple vegetables:** Beetroot, aubergine, red onion, radishes, red cabbage, black/red beans **Other**: Potatoes	White onion, courgette, peppers, tomatoes, peas, sweetcorn, mushrooms, green beans, white beans, asparagus Sweet potato
Other	**Dark chocolate and milk chocolate** **Foods containing dark or milk chocolate**	White chocolate Custard cream, Rich tea biscuits
Drinks	**Red wine** **Juice/smoothies of fruits and vegetables named above** **Drinks flavoured with fruits or vegetables (e.g., fruit squash/cordial, fizzy drinks such as lemonade, Fanta, Sprite)** **Coffee and Tea (black, green, oolong, fruit and herbal teas)** **Drinks containing chocolate (e.g., chocolate milk, Ovaltine)**	Water White wine other alcoholic drinks (e.g., lager/cider/spirits) Coke, Soda water, Ginger ale Milk

4.3 Definition of high and low anthocyanin metabolisers

The urinary metabolites which formed the metabolite panel used to categorise high and low metabolisers included hippuric acid, 3-hydroxyhippuric acid, 4-hydroxy-3-methoxyphenylacetic acid, and 3,5-dihydroxyphenylpropionic acid.

These metabolites correspond to Panel 5 described in Section 3.2.3 and were the result of the original analysis performed in Chapter 3 with all available data (n = 35). As described in Section 2.5 in the note to examiners, this panel differed from the final panel presented in Chapter 3 (i.e. Panel 2 in Section 3.2.3), because the analysis shown in Chapter 3 used a refined approach with more stringent data quality assessments applied. This analysis (with n = 25 instead of n = 35) was performed after the recruitment for the AMP study was already completed and Panel 5 implemented.

Because the refined analysis resulted in seven metabolites in the final panel, the validity to use a metabolite panel formed of only four metabolites was confirmed through the following:

1) these metabolites were among the most abundant urinary metabolites in the CIRCLES study following the blueberry intervention;

2) the 7-metabolite panel and the 4-metabolite panel strongly correlated ($r = 0.90$, $p < 0.0001$); and

3) the 4-metabolite panel had a strong positive relationship with change in FMD even with all data of poorer quality removed ($\beta = 0.70$, $p = 0.026$), which was only slightly weaker than the 7-metabolite panel ($\beta = 0.79$, $p = 0.020$) (see Table 3-13).

For the AMP study prospective recruitment, the metabolite panel was calculated as the average of the $\log_2$ fold change ($\log_2$ FC) of each metabolite from 0 to 48h. The value at 0h equals the amount measured in the urine sample collected from -24 to 0h and the value at 48h equals the cumulative amount measured from 0 to 24h and 24 to 48h. An assessment over 48 h was chosen because previous research has shown that the anthocyanin metabolites assessed in this study are elevated over at least a 48 h period in urine following the intake of a single dose of anthocyanins.[122,123,219]

Based on the results from Chapter 3, high metabolisers were categorised as those in the top 15 % of the distribution and low metabolisers as those in the bottom 15%. This was because the points at which the $\log_2$ FC started to deviate from the average was at approximately the 15^{th} and 85^{th} percentiles (indicated in black in **Figure 4-4**). The 70% in the middle were classified as medium metabolisers. Because of the rolling recruitment to the study, the first categorisation was done after at 24 people had completed the blueberry challenge.

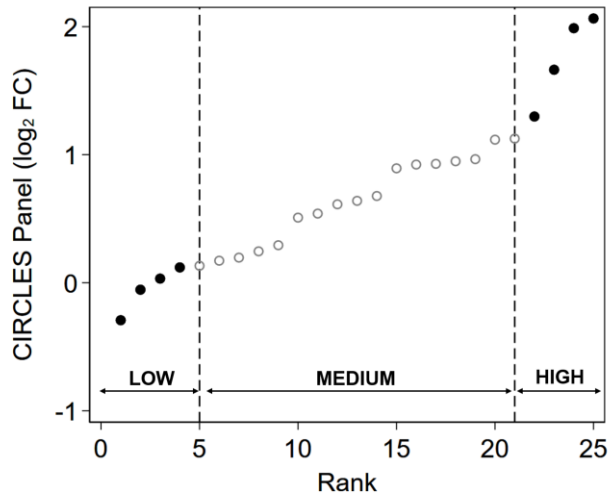


Figure 4-4. Distribution of the CIRCLES metabolite panel

Figure shows the categorisation of low, medium, and high metabolisers in the CIRCLES study by log_2 FC of Panel 2 (see Chapter 3). Grouping into low (lower 15%), medium (middle 70%), and high (upper 15%) metabolisers is indicated through the dashed lines at the 15th and 85th percentiles. Low and high metabolisers are coloured in black, medium in grey.

4.4 Biological sampling and quantification of urinary anthocyanin metabolites using LCMS

To assess metaboliser status, participants were asked to provide 24h pooled urine samples on three consecutive days. 24h before the blueberry milkshake provided the baseline measurement and two subsequent samples at 0 to 24h and 24 to 48h following the blueberry drink. Bacteria in urine may grow and metabolise compounds of interest, such as hippuric acid, thereby altering their concentrations.[275] Compounds could further degrade over time due to oxidation. To prohibit bacterial growth and breakdown of metabolites, the collection containers contained 15 g of boric acid (bacteriostatic recommended for urine collections[276]) and 100 mg of ascorbic acid (antioxidant[140]) as preservatives to prohibit bacterial growth and breakdown of metabolites. The first urine sample (-24h to 0h) was received on the day of the blueberry milkshake. The +24h and +48h pooled urine samples were handed in on a second visit to the facility two days after the milkshake (i.e., the 48h mark). Each urine collection was mixed well, and a 10 ml aliquot was centrifuged for 15 min at 1300 rcf at 4 °C. Three aliquots of 500 µl each were acidified with 95% formic acid (reagent grade, Sigma Aldrich, Dorset UK) using 32 µL/mL and stored at -80 °C until analysis.

The samples were analysed for blueberry anthocyanin metabolites using liquid-chromatography tandem mass spectrometry (LCMS; the combination of HPLC (high performance liquid chromatography) and ESI-MS/MS (tandem mass spectrometry using

electrospray ionisation) based on a published method developed in our group[183] and further adaptions made for the CIRCLES study by Dr. Vera van der Velpen (published in Curtis et al (2019)[159]). Due to breakdown of the instrument during the study, sample analysis was performed on two different instruments in the same lab. Samples were analysed across seven runs, of which five runs (samples of n = 82 participants) were analysed on an Agilent 1200 HPLC connected to a SCIEX 3200 series Q-trap MS/MS (SCIEX, Warrington, UK); and two runs (samples of n = 37 participants) were analysed on a Shimadzu UFLCXR coupled to a SCIEX 4000 MS/MS (SCIEX, Warrington, UK).

Described in the following are the solid phase extraction, external standard curve and LCMS methods (including slight adjustments to the method described in Section 2.4), as well as an overview of method development and validation of the LCMS method for the second instrument.

Solid phase extraction (SPE)

100 µl urine were purified using a 96 well SPE plate (Strata™-X Polymeric Reversed Phase, microelution 2 mg/Well, Phenomenex, Cheshire, UK). Plates were preconditioned with 200 µl of 1% formic acid (reagent grade, Sigma Aldrich, Dorset UK) in methanol (LCMS grade, Fisher Scientific, Loughborough, UK) followed by 200 µl 1% formic acid in water (MilliQ grade, 18.2 MΩ cm^{-1}). Samples were then loaded onto the plate, washed twice with 200 µl 1% formic acid in water and drained under gravity. The plate was dried using a vacuum manifold, applying first 15 inHg and then 5 inHg for 30 min. Samples were eluted with 1% formic acid in methanol. For the elution, the plate was first allowed to soak with the eluent for 5 min before slowly eluting by applying mild vacuum (2 - 4 inHg) for 5 minutes. Phlorizin was added after the SPE at 10 µM final concentration as an internal standard to correct for instrument drift. The final sample purified volume was 100 µl.

External standard curve

Matrix-matched standard curves with 14 points were prepared from 4x-diluted pooled urine (from -24 to 0h urine samples of 10 randomly chosen participants and purified through SPE) ranging from 0 to 100 µM except for hippuric acid. Due to the known high concentration of hippuric acid, the standard curve ranged from 0 to 2000 µM. Linearity of standard curves across the seven runs ranged from 0.954 to 0.999, with 61% of standard curves > 0.993.

LCMS method Agilent 1200/SCIEX 3200

The LCMS method (HPLC gradient and parameter settings on the SCIEX 3200 mass spectrometer was the same as described in Section 2.4, except that a reduced number of analytes was acquired, the internal standard phlorizin was added, and MRM (multiple reaction

monitoring) with 50 ms dwell time was used instead of sMRM (scheduled MRM). The compound parameters of the relevant analytes were reassessed and individually optimised using syringe infusion from purchased standards (Sigma, Dorset, UK). The metabolites were identified based on retention time and three to four transitions (see **Table 2-1** below and **Figure 4-5A**).

LCMS method on Shimadzu UFLCXR /SCIEX 4000

Following the SPE, 1 µl of sample was injected onto a Kinetex polyfluorphenol HPLC column (2.6µM, 100 x 2.1mm; Phenomenex, Cheshire, UK) at a column temperature of 37°C. The HPLC method used a mobile phase of 0.1% formic acid in water and 0.1% formic acid with a flow rate between 0.3 and 0.45 ml/min using a stepped gradient from 1% to 99% acetonitrile over a total run time of 13 min. The source parameter settings were: curtain gas: 25 psi; ion source gas 1: 35 psi; ion source gas 2: 50 psi; ion spray voltage: -4500; temperature: 650°C. Compound specific parameters for each analyte were individually optimised using syringe infusion from purchased standards (Sigma, Dorset, UK). Metabolite identification was based on retention time and two to three transitions (see Table 2-1 below and Figure 4-5B).

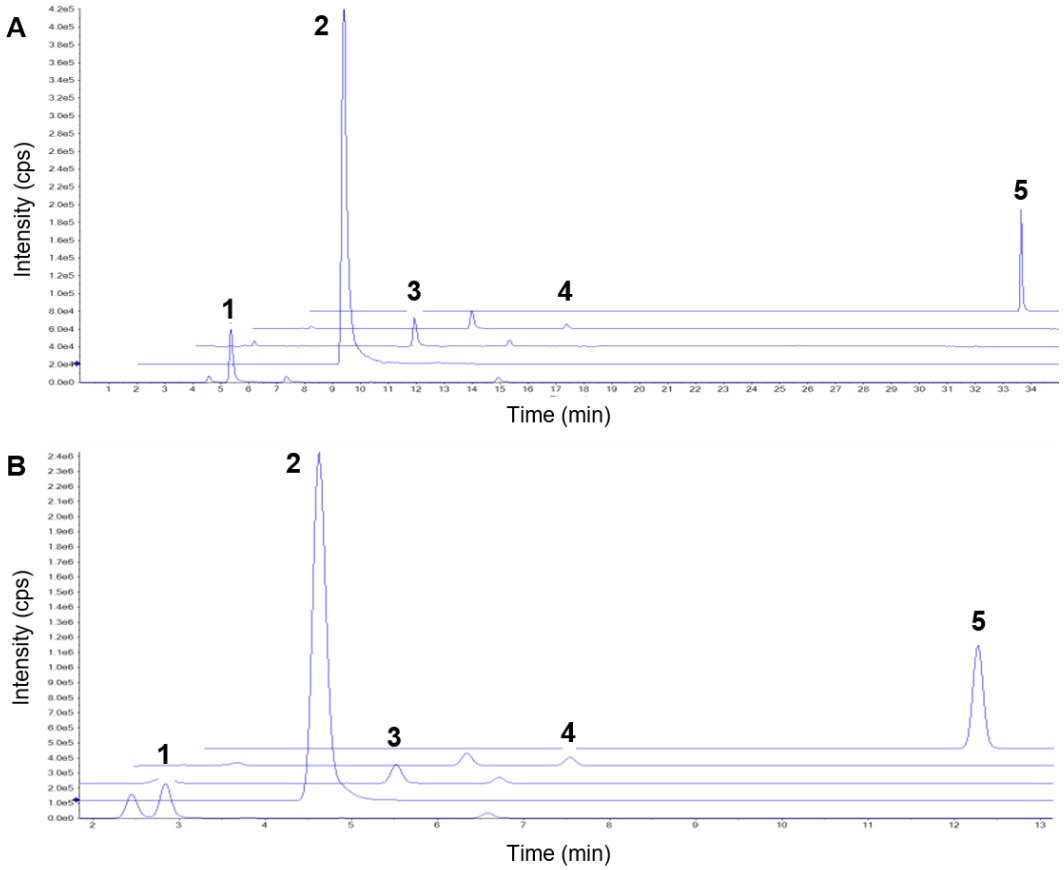


Figure 4-5. Multiple reaction monitoring (MRM) chromatogram of standard compounds for (A) SCIEX 3000 and (B) SCIEX 4000

Compounds: **1** 3-hydroxyhippuric acid; **2** hippuric acid; **3** 3,5-dihydroxyphenylpropionic acid; **4** 4-hydroxy-3-methoxyphenylacetic acid; **5** phlorizin

Method validation for the HPLC-MS/MS method on the Shimadzu/SCIEX 4000 instrument

The method was validated in terms of precision and accuracy as described in the EMA Guideline for Bioanalytical Validation [277].

Precision (the degree of scatter between repeated measurements) and accuracy (closeness of determined value to true value) were assessed using n = 8 urine samples taken from the AMP study and spiked with three different concentrations across the quantitative range (low, medium, high). The results are shown in **Table 4-2**. Across the four metabolites average precision for the low, medium, and high concentrations were 9.8%, 4.7%, and 3.5% (total range 2.2 – 15.8 %). Average accuracies were 105.2%, 85.5%, and 80.6% (total range 72.8 – 117.7%). Generally, the measurement was more precise at high concentrations, and more accurate at mid to low concentrations, but was considered acceptable overall. The precision of the internal standard phlorizin was also determined with 40 repeated injections across different runs and was at 4.4%.

Although the accuracy of three metabolites at mid and high concentrations were slightly below the recommended value of 80%, good precision levels showed that the level of accuracy was consistent throughout repeated testing at varying concentrations. Because the change to the second instrument happened at the beginning of the COVID-19 pandemic, this level of accuracy was accepted to progress with the analysis of samples before the laboratory closed.

Table 4-2. Precision and accuracy of the method used on Shimadzu/SCIEX 4000 system

Compound	low (µM)[a]	medium (µM)[a]	high (µM)[a]	Low Accuracy (Precision) in %	Med Accuracy (Precision) in %	High Accuracy (Precision) in %
3,5-dihydroxyphenylpropionic acid	1	37.5	75	84.4 (9.4)	85.5 (3.8)	77.6 (2.5)
3-hydroxyhippuric acid	10	37.5	75	116 (9.0)	76.4 (4.9)	80 (3.6)
4-hydroxy-3-methoxyphenylacetic acid (homovanillic acid)	5	37.5	75	117.7 (5.2)	74.8 (2.2)	72.8 (2.9)
Hippuric acid	200	750	1500	102.7 (15.8)	105.4 (7.7)	92.1 (5.1)

[a] Concentration of standard compound spiked into n = 8 random urine samples taken from the AMP study

Data analysis

Peak areas were identified based on two to four transitions and quantified relative to the most abundant transition. SCIEX OS 1.4 software (AB Sciex) was used to determine analyte peak areas, which were normalised to the peak area of the internal standard phlorizin, and concentrations. The concentrations were adjusted to the background level contained in the matrix used to generate the external standard curve, i.e., the quantified values represent the amount of metabolite measured above the background present in the matrix (shown in **Table 4-3**).

Table 4-3. Background concentration of metabolites present in the matrix of the standard curve

Metabolite	Concentration (µM)
3,5-Dihydroxyphenylpropionic acid	2.2
3-Hydroxyhippuric acid	13.9
4-Hydroxy-3-methoxyphenylacetic acid (homovanillic acid)	3.0
Hippuric acid	736.6

Further data processing was handled in Excel. This included adjustment of the concentrations to the total volume of the 24h pooled sample, giving the total metabolite amount in µmol/24h to account for variations in urine volume between participants. Any metabolites which had a detectable peak, but quantified below the blank, were adjusted to the LOD.

LOD for both instruments was estimated following guidelines from Agilent Technologies [278] (Table 2-1). A standard close to the expected limit of detection and a blank were injected in triplicate. The average area of the blank value was used to subtract the background from each

analyte. Using a one-sided t-test, the LOD was then determined as the amount of analyte that gives a signal statistically greater than 0 with a known confidence level. The equation used was $LOD = \frac{t \times CV \times c_{std}}{100\%}$ with t = 0.292 for 2 degrees of freedom and 95% confidence, CV = coefficient of variation of the replicate area measurement, and c_{std} = amount of standard.

Table 4-4. HPLC-MS/MS identification and LOD of compounds on both instruments (SCIEX 3200 and SCIEX API 4000)

		SCIEX 3200			SCIEX 4000		
Compound	**Mode**	**R_t (min)**	**ion transitions (m/z)**	**LOD[a] (µM)**	**R_t (min)**	**ion transitions (m/z)**	**LOD[a] (µM)**
3,5-Dihydroxyphenylpropionic acid	-	8.4	181/137, 122, 95	0.74	3.9	181/137, 95	0.36
3-Hydroxyhippuric acid	-	5.8	194/150, 148, 93	5.16	2.9	194/150, 93	1.81
4-Hydroxy-3-methoxyphenylacetic acid	-	11.8	181/137, 166, 122, 79	1.72	5.1	181/137, 122, 78	0.20
Hippuric acid	-	7.9	178/134, 132, 77	20.22	3.8	178/134, 77	15.23
Phlorizin[b]	-	25.4	435/273, 167, 123	ND	8.9	435/273, 167, 123	ND

[a] LOD was calculated following guidelines from Agilent Technologies [278] using triplicate measurements and 95% confidence.
[b] Phlorizin added as internal standard at 10 µM concentration post-SPE; LOD not determined.

4.5 Assessment of anthropometric measures

As part of the screening procedures for study participation, anthropometric measures were assessed in the fasted state. The clinical screen was conducted within 3 months of the blueberry challenge. Participants were measured for weight (kg) and height (m) (average of two measurements), and BMI was calculated using the standard equation (kg/m^2).

4.6 Assessment of habitual diet and flavonoid intake

Dietary intake was assessed using a 131-item validated food frequency questionnaire (FFQ) from the EPIC study[165], recorded one week before the study visit, in n = 108 participants. Assistance with the collation of these data was provided by Dr. Amy Jennings and Veronica Bion. Of note, FFQ data was not assessed for n = 11 participants from my dataset. At the time of the first phase of FFQ data analysis (mid-February 2020), the blueberry challenge assessment had not yet been completed for these volunteers and the plan was to repeat a second phase of FFQ analysis in this small sub-set of participants (n=11) after the blueberry challenge had been completed. However, the emergence of the COVID-19 pandemic and the subsequent shut down of the university resulted in this activity being cancelled, hence the missing data.

In quality assurance steps, dietary data was excluded in the dataset if energy intake was deemed implausible (i.e., greater than ± 2 SD from the mean of the ratio of energy intake to estimated energy requirement; n = 2) or more than ten food items were left blank (n = 0).

Nutrient and flavonoid content was calculated as described by Welch et al [279]. Briefly, flavonoid values were assigned to each of the foods in the FFQ. The amount of individual compounds was determined as the sum of the content in each food of the specified portion size multiplied by the frequency of consumption. Flavonoid intakes of the six main subclasses, namely anthocyanins (cyanidin, delphinidin, malvidin, pelargonidin, peonidin, petunidin), flavanones (eriodictyol, hesperitin, naringenin), flavan-3-ols (catechins, epicatechins), flavonols (isorhamnetin, kaempferol, myricetin, quercetin), flavones (apigenin, luteolin), and oligomer and polymer flavonoids (theaflavins, thearubigins, proanthocyanidins) were derived from the flavonoid databases of the US Department of Agriculture (USDA)[71,72]. This method been previously used successfully to estimate flavonoid intakes for association studies[98,99].

4.7 Statistical analysis

All analysis was performed in Stata 16.0 (Stata Corp, College Station, Texas, USA). Medians of metaboliser groups at baseline were compared using quantile regression[280].

For the FFQ data, the difference between the three metaboliser groups were compared using the Kruskal-Wallis test, followed up by pairwise comparisons using Dunn's method.

The relationship between gender and metaboliser group was assessed using the chi-squared test.

A comparison of the metaboliser groups between the CIRCLES and AMP study was done using the Mann-Whitney-U test.

Sample size calculation

The main sample size calculation for the AMP study was based on the requirements for the principal statistical analysis of phase 2 of the study. A separate consideration was made for the minimum sample size requirement for the blueberry challenge to detect a difference between low and high metabolisers after blueberry intake. Both are described in the following.

Minimum sample size requirement for the blueberry challenge:

The AMP study was the first study to prospectively recruit participants based on the metaboliser panel established in Chapter 3 (see also Section 4.3 on the definition of high and low metabolisers). As such there was no precedent to estimate the group difference and SD following a single exposure to blueberries. As an approximation, the sample size calculation was based on the difference in $\log_2$ FC in Panel 5 metabolites between the blueberry (1 cup) and placebo groups in the CIRCLES study (Section 3.2.3, Table 3-15). Assuming a difference of 0.97 between the groups and an SD of 0.94, at a significance level of 0.05 and 80% power, the calculation generated a requirement for a sample size of n = 16 per group. Thus, due to

the definition of high and low metabolisers as those in the top and bottom 15 % of the distribution of the metabolite panel, a minimum of n = 107 participants were required of which n = 32 would be assessed as high or low metabolisers.

Sample size calculation for phase 2 of the AMP study

Phase 2 of the AMP study involved a cross-over intervention design. The primary statistical analysis aim of phase 2 was to determine time by treatment effects in two metaboliser groups (high and low) for change in %FMD following an energy-dense test meal with or without blueberries. A power calculation was performed based on the FMD analysis in the CIRCLES study, published in Curtis (2019)[159] and used the blueberry and placebo groups as approximations for high and low metabolisers. Using a difference of 1.09 %FMD between blueberry and placebo intake and an overall population SD of the difference of 1.87, the calculation suggested a requirement of n = 35 per group to have 90% power to detect a difference between the two groups at a significance level of p = 0.05. To account for dropouts between random allocation to treatment and completion of the study, the recruitment aim was 15% higher, increasing the group size to 41 per group. Thus, the total sample size for phase 2 of the AMP study was n = 82, of which 70 were required to complete the study. For the blueberry challenge, this translated to a total recruitment target of n = 274 due to the definition of high and low metabolisers as those in the top and bottom 15 % of the distribution of the metabolite panel (see Section 4.3 below).

Due to the emergence of the COVID-19 pandemic, the study was stopped prematurely. At the time of termination, n = 121 completed the blueberry challenge, and n = 26 had completed at least one arm of the cross-over trial.

5 AMP study – Prospective recruitment by anthocyanin metaboliser profile

5.1 Introduction and study objective

Although flavonoids, and in particular anthocyanins, have been associated with cardiometabolic health benefits, results from randomised controlled trials remain equivocal. This may partly be due to the inter-individual variability in response to flavonoid intake, resulting in differences in absorption, metabolism and exposure of target tissues to flavonoids and their metabolites.

The study objective of the blueberry challenge was to determine whether individuals could be grouped as low, medium, or high metabolisers after a single dose of blueberries using a panel of urinary anthocyanin metabolites which was identified in the previous study through an exploratory analysis (see Chapter 3). The panel was found to be a set of likely colonic microbial metabolites that putatively characterised a high metaboliser of blueberry anthocyanins and was associated with an improvement in vascular function. Although these metabolites have been observed to be present at a basal level due to endogenous production and other dietary sources, the intake of blueberry anthocyanins is expected to increase the urinary excretion of these metabolites beyond their basal level. Importantly, however, the metabolite panel reflected a chronic intake of blueberry anthocyanins for six months, whereas in this study, its transferability to a single exposure is tested.

5.2 Results

Baseline data of participants

In total, n = 119 volunteers provided valid samples after participating in the blueberry challenge (n = 2 had invalid data due to violation of dietary restrictions and intolerance to the intervention drink). Their baseline characteristics are shown in **Table 5-1**. The majority of participants were male (61%) and average BMI was 29.1 kg/m^2.

Table 5-1 Participant characteristics[1]

Characteristic	All (n = 119)
Age (y)	61.8 ± 7.4
Gender (n (%) female)	46 (39%)
BMI (kg/m^2)	29.1 ± 3.4

[1] Values for are mean ± SD

Classification of participants into three distinct groups using the metabolite panel

The metabolite panel was assessed as the log_2 FC from baseline to 48h. Using the 15^{th} and 85^{th} percentiles as cut-offs, the participants were classified into three distinct groups with n = 19 in the low group, n = 78 in the medium group, and n = 22 in the high group (see **Figure 5-1A**).

The number of participants in the low and high group were not equal because the analysis was spread over seven runs while recruitment and the blueberry challenge were still ongoing. This meant that the percentile cut-offs changed with every new plate that was analysed, whereas the final 15^{th} and 85^{th} percentiles included all n = 119 participants. These are represented by the two dashed lines in Figure 5-1A. Participants, who in earlier analysis runs were classified as high or low metabolisers but were grouped as medium in the final analysis, appear in Figure 5-1A as high or low metaboliser, but between the two dashed lines.

Figure 5-1A illustrates that the log_2 FC ranged from -3.2 to 6.6 across all participants. The mean in the low group was 0.1, demonstrating no to very limited change in excretion of metabolites over 48h. The medium group had a mean of 1.8, which translates to a 3.5-fold increase in urinary metabolites. On average, the high group increased nearly 16-fold in urinary metabolites. Across all participants (irrespective of metaboliser group), the mean log_2 FC was 1.93.

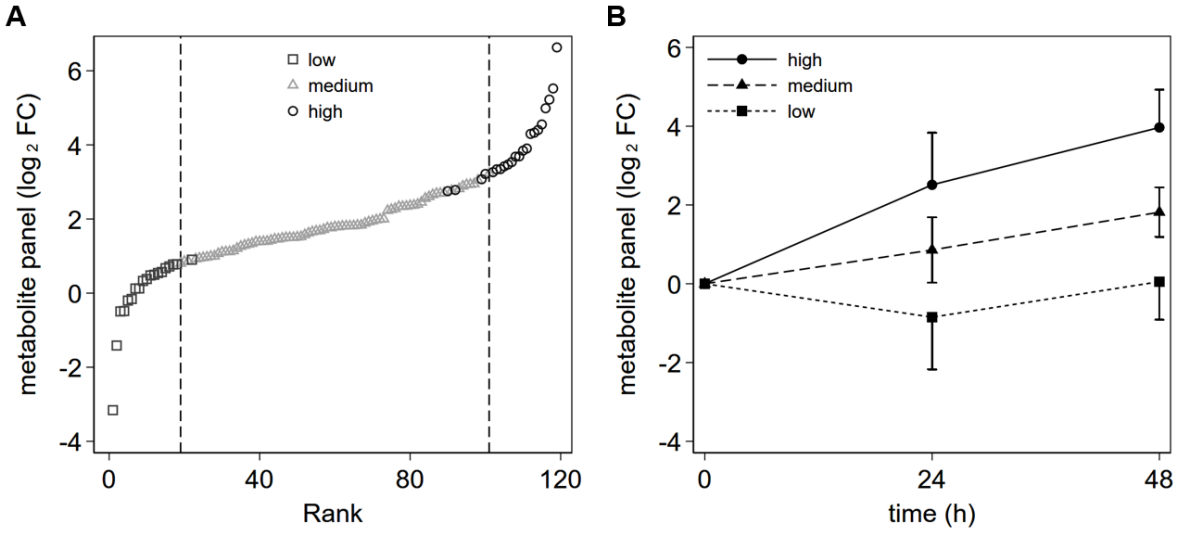


Figure 5-1. The distribution of the metabolite panel (log2 FC) at 48h (A) and its change over time (B)

(A) The distribution of the metabolite panel (log_2 FC) at 48h. Participants below the 15th percentile (left dashed line) were classified as low, while participants above the 85th percentile (right dashed line) were classified as high metabolisers. Participants in between were classified as medium. Some participants, who were classified as high or low appear between the two dashed lines due to rolling recruitment and classification of participants before the analysis of the entire sample of n = 119 had been completed. The dashed lines represent the final 15th and 85th percentiles including all n = 119 participants.
(B) The log_2 FC of the metabolite panel plotted over time by group (low, medium, and high metaboliser) as determined in (A). Datapoints show cumulative mean ± SD.

Figure 5-1B shows the cumulative change of the metabolite panel over time. In both the medium and the high group, the excretion of metabolites increased from baseline to 48h, with the sharpest increase being in the high group for both timepoints. In contrast, the low group at first decreased in excretion of metabolites from 0 to 24h and then returned to baseline level at 48h.

The absolute means measured at each time point are given in **Table 5-2** per group and for all participants. Across the entire sample ('All' in Table 5-2), the maximum amount of change from baseline was observed at 0 to 24h for all metabolites except for 3,5-dihydroxyphenylpropionic acid, which reached a maximum amount at 24 to 48h. As expected, the main metabolite was hippuric acid with an average maximum amount of 2222.3 ± 1444.6 µmol/24h at 0 – 24h in comparison to 6 – 20 µmol/24h for the other metabolites. Unexpectedly, the excretion of two metabolites changed little over time in comparison to the results in the previous chapter and other single-dose studies: the maximum difference from baseline for 3,5-dihydroxybenzoic acid was 1.3 µmol/24h observed at 24 to 48h (21% increase); for 4-hydroxy-3-methoxyphenylacetic acid the maximum difference was 0.6 µmol/24h at 0 to 24h (4% increase). In contrast, de Ferrars et al reported a 600% increase for 3,5-dihydroxy benzoic acid within the first three

hours.[245] Part of the discrepancy could be a reflection of differing doses and urine storage methods. De Ferrars et al fed 500 mg of anthocyanins from elderberry and did not add any preservatives to the collected urine prior to storage, which may have affected the metabolite concentration. For 4-hydroxy-3-methoxyphenylacetic acid, Nurmi et al reported a maximal increase of 13% at 0 to 24h over a total observation period of 48h.[122] Although the increase was three times the amount of what was observed in our study, the value is better aligned with our findings as the anthocyanin dosage Nurmi et al used was nearly double with 650 mg anthocyanins from bilberries and lingonberries.

Interestingly, the low group had the highest baselines for all metabolites, which may explain the lack of response in these metabolites as well as the initial decrease observed at 0 to 24h for this group.

Table 5-2. Urinary excretion of panel metabolites up to 48h after blueberry drink by group

Group	Time	3,5-DiOH-PPA	3-OH-HA	4-OH-3-OCH-PAA	HA
Low (n = 19)	-24 to 0 h	11.0 ± 13.6	31.3 ± 54.1	17.5 ± 11.6	1708.7 ± 1461.0
	0 to 24h	5.4 ± 10.7	18.4 ± 37.0	12.4 ± 9.3	2139.3 ± 1529.8
	24 to 48h	8.5 ± 11.8	15.3 ± 35.5	13.2 ± 10.9	759.3 ± 871.0
Medium (n = 78)	-24 to 0 h	6.1 ± 10.6	12.3 ± 44.4	16.7 ± 10.0	841.8 ± 1122.6
	0 to 24h	6.4 ± 8.6	17.0 ± 27.1	17.3 ± 12.8	2249.8 ± 1406.2
	24 to 48h	6.3 ± 7.2	14.0 ± 35.3	16.4 ± 10.1	1060.4 ± 1320.1
High (n = 22)	-24 to 0 h	2.7 ± 4.5	7.5 ± 9.7	12.3 ± 9.0	448.2 ± 896.8
	0 to 24h	6.4 ± 6.3	25.0 ± 24.8	17.4 ± 12.2	2196.5 ± 1568.3
	24 to 48h	11.6 ± 10.8	27.7 ± 35.3	17.6 ± 11.6	1148.0 ± 1573.9
All (n = 119)	-24 to 0 h	6.3 ± 10.5	14.4 ± 42.5	16.0 ± 10.2	907.5 ± 1198.6
	0 to 24h	6.2 ± 8.5	18.7 ± 28.4	16.6 ± 12.3	2222.3 ± 1444.6
	24 to 48h	7.6 ± 9	16.8 ± 35.4	16.1 ± 10.5	1028.5 ± 1307.2

Values are mean ± SD in µmol/24h. 3,5-DiOH-PPA: 3,5-dihydroxyphenylpropionic acid; 3-OH-HA: 3-hydroxyhippuric acid; 4-OH-3-OCH-PAA: 4-hydroxy-3-methoxyphenylacetic acid; HA: Hippuric acid

Comparison of the baseline excretion of individual metabolites by group

The difference in baseline values between the groups was further evaluated. Because a large variability within the groups was indicated through the large SD values in in Table 5-2, the medians rather than the means were compared between the groups as a more robust measure against extreme values.

When comparing the median excretion of the individual metabolites at baseline, it was confirmed that for three of four metabolites (i.e., 3,5-dihydroxyphenylpropionic acid, 4-hydroxy-3-methoxyphenylacetic acid, and hippuric acid) the median baseline value was highest in the low group and lowest in the high group (**Figure 5-2)**. Using quantile regression (see **Table 5-3** for results), these differences were significant at a level of 0.05 for 3,5-dihydroxyphenylpropionic acid ($p = 0.048$) and hippuric acid ($p < 0.001$), but not 3-

hydroxyhippuric acid (p = 0.551) or 4-hydroxy-3-methoxyphenylacetic acid (p = 0.307). In a pairwise follow-up test with Bonferroni correction for multiple comparisons, the median baseline value of the low group was significantly higher than the high group for 3,5-dihydroxyphenylpropionic acid (p = 0.017) and hippuric acid (p < 0.001). It was also higher in the low group in comparison to the medium group for hippuric acid (p = 0.003). The medians between the medium and the high group did not differ significantly.

Furthermore, the large variability was also confirmed through the wide interquartile ranges. This was particularly apparent in the low group for 3-hydroxyhippuric acid and hippuric acid (metabolites 2 and 4 in Figure 5-2).

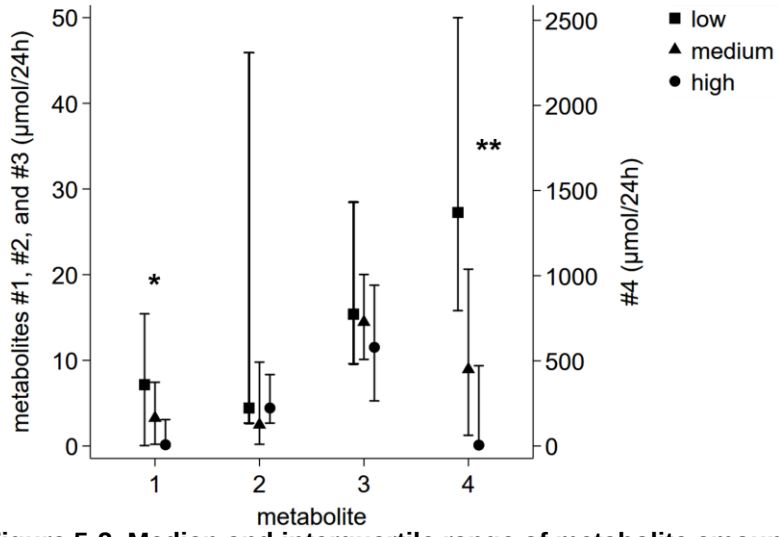


Figure 5-2. Median and interquartile range of metabolite amounts at baseline per metaboliser group

Plot shows the median amounts at baseline (-24 to 0 h sample) per metaboliser group. Bars represent the interquartile range. Metabolites 1 – 4 are as follows: 1 = 3,5-dihydroxyphenypropionic acid; 2 = 3-hydroxyhippuric acid; 3 = 4-hydroxy-3-methoxyphenylacetic acid; 4 = hippuric acid. Difference in medians between the three groups was tested using the quantile regression followed by pairwise comparison of medians with Bonferroni correction. The medians for metabolites 1 and 4 were significantly different at a level of 0.05. * p = 0.017 (low versus high); ** p = 0.003 (low versus medium) and p < 0.001 (low versus high).

Table 5-3. Median and interquartile range at baseline

Group	Baseline	3,5-DiOH-PPA	3-OH-HA	4-OH-3-OCH-PAA	HA
Low (n = 19)	median	7.15[a]	4.43	15.39	1371.36[bc]
	IQR	0.05 – 15.43	2.66 – 45.92	9.58 – 28.46	794.94 – 2515.66
Medium (n = 78)	median	3.26	2.48	14.47	449.89
	IQR	0.2 – 7.43	0.2 – 9.78	10.1 – 20.02	62.5 – 1037.96
High (n = 22)	median	0.14	4.43	11.52	5.05
	IQR	0.04 – 3.08	2.66 – 8.33	5.26 – 18.77	2.28 – 471.13
All (n = 119)	median	2.98	4.43	14.08	540.84
	IQR	0.17 – 7.65	0.2 – 9.87	9.58 – 20.26	62.5 – 1278.08
Quantile regression	p (trend)	0.007	0.551	0.307	< 0.001

Values are median and interquartile range in µmol/24h for each metaboliser group. Quantile regression was used to compare medians between groups. Significant differences at a level of 0.05 indicated below.
[a] p = 0.017 for low versus high
[b] p = 0.003 for low versus medium
[c] p < 0.001 for low versus high
3,5-DiOH-PPA: 3,5-dihydroxyphenylpropionic acid; 3-OH-HA: 3-hydroxyhippuric acid; 4-OH-3-OCH-PAA: 4-hydroxy-3-methoxyphenylacetic acid (homovanillic acid); HA: Hippuric acid

Comparison of habitual food intake by group

To determine whether habitual intake of certain foods influenced the metaboliser profile, a validated food frequency questionnaire which estimates the habitual intake in the last year was assessed. The groups were compared for differences in habitual dietary intake of polyphenol-rich foods and fibre using the Kruskal-Wallis test. Fibre was included as there is strong evidence that it is associated with a healthy gut microbiome and may influence bacterial derived metabolites through modulation of the microbial diversity and activity.[55]

A trend for higher intake in the low group could be observed for several foods and nutrients (including fruits, red wine, total flavonoids, and fibre), but only anthocyanins and wholemeal intake were significantly different between the groups at a significance level of 0.05 (p = 0.031 and 0.051). Pairwise follow-up tests showed that for both anthocyanins and wholemeal, the habitual intake in the low group was significantly higher than the medium and the high group (p for anthocyanins: 0.006 (low versus medium) and 0.013 (low versus high); p for wholemeal: 0.029 (low versus medium) and 0.008 (low versus high)). There was no statistical difference between the medium and the high group (anthocyanins: p = 0.433; wholemeal: 0.146).

Table 5-4. FFQ dietary intake data

	All (104)	Low (n = 18)	Medium (n = 64)	High (n = 22)	p
Energy (kcal)	2165.7 (1807.9 – 2448.9)	2231.2 (1824.8 – 2420.2)	2134.1 (1818.0 – 2437.0)	2151.3 (1768.9 – 2477.2)	0.854
Protein (% energy)	18.0 (16.5 – 19.9)	18.0 (16.7 – 19.3)	18.2 (16.5 – 20.1)	17.4 (15.4 – 19.1)	0.350
Fat (% energy)	15.5 (14.3 – 16.7)	14.8 (13.7 – 15.9)	15.6 (14.4 – 16.6)	15.8 (15.1 – 16.9)	0.174
Carbohydrate (% energy)	46.2 (41.9 – 49.7)	47.6 (41.3 – 50.2)	46.6 (42.3 – 50.3)	44.7 (39.3 – 48.3)	0.374
Total flavonoids (mg/d)	838.3 (623.7 – 1245.1)	884.8 (725.4 – 1353.8)	834.3 (500.8 – 1245.1)	844.9 (732.6 – 1109.9)	0.34
Anthocyanins (mg/d)	20.7 (13.3 – 30.7)	27 (19.3 – 37.8)	20.5 (12.6 – 26.4)	14.7 (10.2 – 34.2)	0.031
Fruits (g/d)	195.7 (126.3 – 292.7)	213.4 (129.7 – 468.4)	191 (115.2 – 282.5)	189.1 (142.5 – 275.8)	0.636
Vegetables (g/d)	220.2 (168.6 – 298.7)	213.5 (145.9 – 297.5)	225.7 (174.8 – 301.6)	212.7 (144.1 – 298.6)	0.544
Coffee (g/d)	475 (81.4 – 475.0)	332.5 (81.4 – 475.0)	475 (27.1 – 475.0)	475 (190.0 – 475.0)	0.589
Tea (g/d)	475 (190.0 – 855.0)	475 (475.0 – 855.0)	475 (149.3 – 855.0)	475 (475.0 – 475.0)	0.371
Red wine (g/d)	18 (0.0 – 57.6)	55.8 (9.0 – 99.0)	18 (0.0 – 54.0)	19.8 (0.0 – 75.4)	0.200
Beer (g/d)	41.4 (20.7 – 124.3)	41.4 (20.7 – 124.3)	41.4 (20.7 – 124.3)	82.9 (41.4 – 124.3)	0.426
Wholemeal (g/d)	40.5 (19.3 – 65.5)	65.2 (32.9 – 111.1)	40.7 (17.7 – 58.6)	27.9 (19.3 – 40.7)	0.051
Fibre (AOAC) (g/d)	18.5 (14.9 – 24.2)	20.7 (16.9 – 27.4)	19.4 (14.8 – 23.5)	18.1 (13.0 – 22.6)	0.236

Values are median the with interquartile range in parentheses. Group differences (low, medium, high) tested using Kruskal-Wallis test. Wholemeal: wholemeal bread and pasta, brown bread, cereal. AOAC: fibre content (non-digestible carbohydrates) estimated using the American Association of Analytical Chemists methods; FFQ: food frequency questionnaire.

Comparison of the metabolite panel at 24h and 48h

Previous studies have observed that some anthocyanin metabolites are excreted over at least 48h, however the majority of metabolites seems to be excreted over the first 24h.[122,123,219] The metabolite panel at 24h and 48h was compared to examine at which timepoint the classification of participants was more reliable. **Figure 5-3** shows distribution of the metabolite panel at 24h and at 48h. In both graphs (for 24h and 48h), the classification of participants into low, medium, and high metabolisers is based on the 15th and 85th percentiles at 48h, which are represented by the dashed lines in both graphs. The graph at 48h is the same as in Figure 5-1 and is shown here for an easier direct comparison.

At 24h, 76% of participants were already classified in their final groups. Of the 24% which were not in their final groups: at 24h, n = 11 were in the medium group, but were classified as low at 48h. Further n = 9 were in the low group, but were classified as medium at 48h. N = 5 participants, which at 24h were grouped as medium or even low in one case (arrows in Figure 5-3, left graph), strongly increased in excretion of the measured metabolites between 24 and 48h and were classified as high metabolisers at 48h.

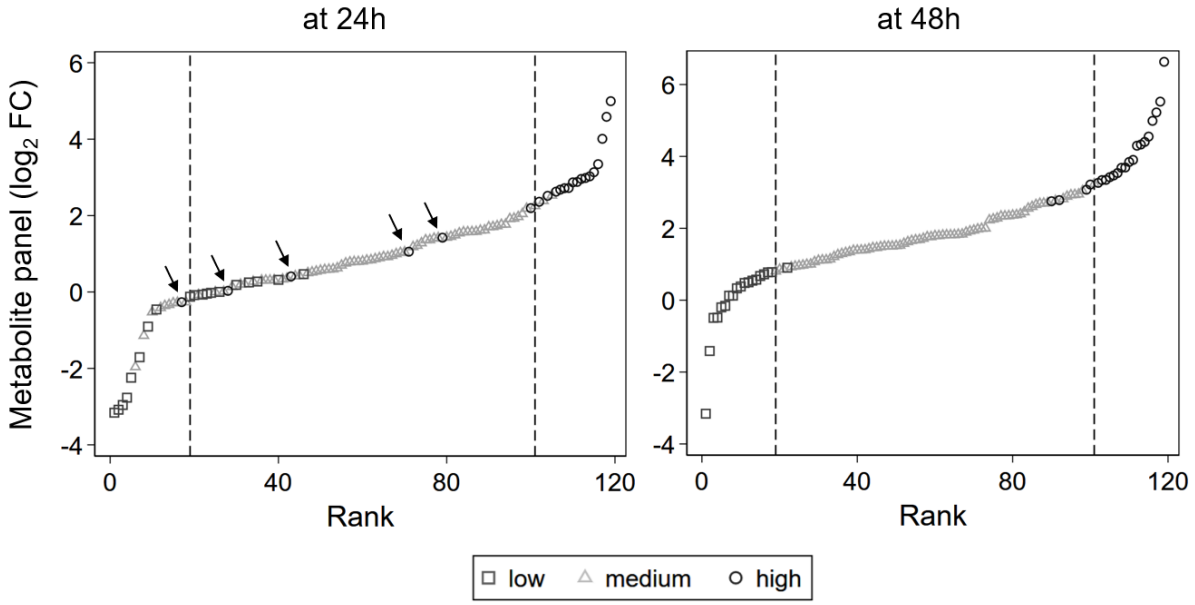


Figure 5-3. Distribution of the metabolite panel at 24 and 48h

Graphs show the log_2 FC of the metabolite panel against the rank of participants at 24h and 48h. Dashed lines show the group cut-offs at the 15th and 85th percentiles (based on distribution of the metabolite panel at 48h). The arrows in the left graph point to participants whose urinary metabolites strongly increased between 24 – 48h so that they were grouped as 'high' at 48h.

Comparison of the metabolite panel at 48h by gender

Because the low group consisted of uneven numbers of men and women (n = 13 men and n = 6 women), while the high group had even numbers of men and women (n = 11 men and n = 11 women), the association between gender and metaboliser group was assessed. However, a chi-square test did not reveal a significant relationship between gender and metaboliser group ($\chi^2(2; n = 119) = 1.67$, $p = 0.435$) (see **Figure 5-4**).

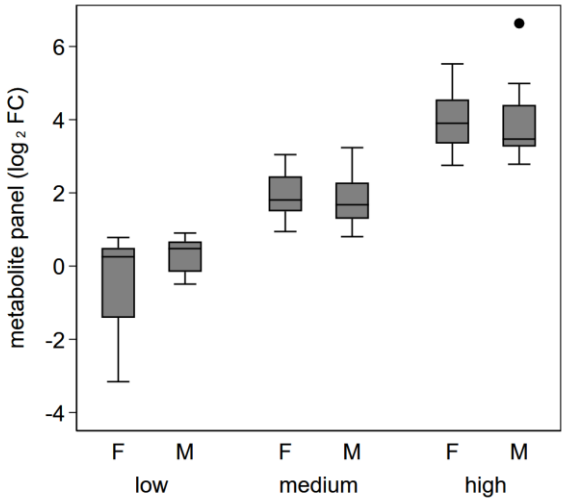


Figure 5-4. Boxplot of the metabolite panel (log2 FC) by gender

Figure shows the distribution of the metabolite panel by gender by group to assess whether gender was associated with metaboliser group. No significant effect of gender on metaboliser groups was observed (chi-square test $\chi^2(2; n = 119) = 1.67$, $p = 0.435$).
Metaboliser groups: low, medium, high. Number of men and women per group (group (men; women): low (13; 6), medium (49; 29), high (11; 11).
F = female; M = male.

Comparison of the panel between the CIRCLES and the AMP study at 24h

To assess how the metabolite response from a single dose of blueberries differed from the response after repeated intake for a longer period of time, the $\log_2$ FC of the metabolite panel in the CIRCLES and the AMP study were compared (**Figure 5-5**). Because the CIRCLES study had only collected urine samples up to 24h, the AMP profile used in this comparison was also at the 24h timepoint. Equally, the metabolite panel for the CIRCLES study was calculated for the four metabolites assessed in AMP for comparability. Only those participants who received the full dose in CIRCLES and thereby a similar amount of anthocyanins were used for this comparison.

Figure 5-5 shows that the mean change in metabolites for the total sample was similar for both studies and that per metaboliser group the general direction of the change in metabolites was the same between the studies. The magnitude of change, however, was more pronounced in the AMP study (i.e., single exposure). The $\log_2$ FC in the low group was significantly greater in the negative direction for AMP (mean difference = -1.05, $p = 0.01$, as tested using the Mann-Whitney-U test), and significantly greater in the positive direction in the high group (mean difference = 1.43, $p < 0.001$). No significant differences were found in the medium group ($p = 0.225$) or across the entire sample ($p = 0.345$).

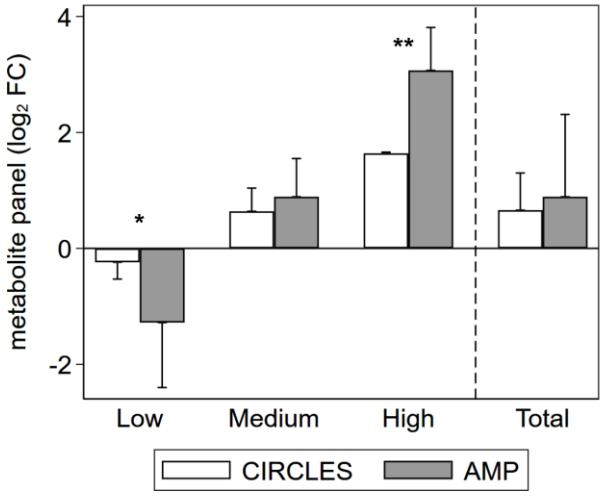


Figure 5-5. Comparison of the metabolite between the CIRCLES and AMP study

Plot shows the mean of the metabolite panel (log_2 FC) in the CIRCLES and AMP studies. For comparability, the CIRCLES panel was calculated with the four metabolites assessed for the AMP study. Equally, because the CIRCLES study had only collected urine samples up to 24h, the AMP profile corresponds to the 24h timepoint. Notably, this comparison contrasts a single exposure (AMP) with a prolonged daily exposure over six months (CIRCLES).
Error bars show standard deviation.
* p = 0.019; ** p < 0.001 (Mann-Whitney-U test).

5.3 Discussion

The aim of the first stage of the AMP study, the blueberry challenge, was to determine the metabolism profile of participants following the consumption of a blueberry milkshake. The profile was assessed from the excretion of metabolites over 48h in urine using the metabolite panel determined through retrospective analysis of a six-month blueberry intervention study (Panel 5, Section 3.2.3). The panel was calculated as the average log_2 FC from baseline to 48h of four metabolites (3,5-dihydroxphenylpropionic acid, 3-hydroxyhippuric acid, 4-hydroxy-3-methoxyphenylacetic acid, and hippuric acid).

The key finding in this analysis was that the metabolite panel was able to classify participants into three distinct groups after consumption of blueberry using the 15th and 85th percentiles as group cut-offs. Due to the rolling recruitment of the AMP study, these cut-offs changed as more participants were analysed. However, although the first analysis only included n = 24 participants, and each additional run added about 15 to 30 participants to the sample size, only 4% of participants were classified differently across all analysis runs. Furthermore, these participants were close to the final percentile cut-offs, suggesting that even in small sample sizes, the 15th and 85th percentiles appear to be quite accurate. The final 15th and 85th percentiles translated into the following thresholds: 0.78 log_2 FC for the lower cut-off and 3.26

log_2 FC for the upper. The total range at 48h varied between -3.2 to 6.6 log_2 FC; clearly demonstrating a wide variability in response between individuals.

Each of the three metaboliser groups had distinct profiles (see Figure 5-1B). The high group increased the most, with the steepest increases between both timepoints among the three groups, while the medium group increased steadily between 0 and 48h. Interestingly, the low group initially decreased from 0 to 24h for the metabolite panel before reaching baseline again at 48h. These differences in profiles could point to group specific differences in metabolization (in the gut and after absorption) and transit time. The fact that the low group increased in hippuric acid from 0 to 24h and decreased again between 24 to 48h suggested that they did absorb some of the anthocyanin metabolites, however overall, the excretion of metabolites did not increase substantially from baseline, which suggests very limited absorption.

A possible confounding factor to the observed decrease in metabolites in the low group was the high baseline levels observed. These were significantly higher in the low group than the high group in the case of 3,5-dihydroxyphenylpropionic acid and hippuric acid (see Figure 5-2 and Table 5-3). A certain level at baseline was expected, based on the findings from the previous chapter and knowledge of other sources for the metabolites: Hippuric acid is endogenously produced from aromatic amino acids and is a definitive catabolite of many polyphenols.[123,222,281] It has been suggested as a potential biomarker of fruit and vegetable intake[240] and in observational studies linking metabolomics with epidemiological data has been inversely associated with systolic and diastolic blood pressure (-1.6 and -0.7 mmHg per 3.55 mmol hippuric acid per 24h)[282], BMI[283], and reduced odds of developing MetS (odds ratio: 0.795; p = 0.026)[284], with hippuric acid acting as a potential marker of polyphenol intake as well as microbial diversity. 3,5-dihydroxyphenylpropionic acid has been reported to be a human urinary metabolite of alkylresorcinols found in wholewheat products.[241] Although currently not experimentally confirmed, it is not unlikely that 3,5-dihydroxyphenylpropionic acid could also be derived from flavonoids with a trihydroxylated B-ring through dehydroxylation at position 4' and microbiotic ring fission to produce a phenylpropionic acid. One of the two main anthocyanidins in blueberries, delphinidin, is a trihydroxylated flavonoid.

Thus, a reason for the higher baseline in the low group in comparison to the high group could be a higher intake of polyphenol-containing foods and wholemeal products prior to the study visit. The analysis of a food frequency questionnaire revealed that the low group habitually ate nearly twice as many anthocyanins (mainly from fruit and red wine) and wholemeal products (wholemeal bread, pasta, brown bread, oats, bran) (see Table 5-4) than the high group. Throughout the study, participants were asked to adhere to dietary restrictions which mainly prohibited the intake of anthocyanin-rich foods. However, participants were allowed to choose from other fruits and vegetables, such as apples, white onions, and tomatoes, as well as

wholemeal products. It may be that the background diet contributed to increased baseline levels in the low group in comparison to the other groups. However, this cannot fully explain the decrease observed in the low group following the blueberry milkshake, if the diet was comparable between the three assessment days. Nevertheless, for the assessment of an anthocyanin metabolic profile in the future, it would be beneficial to ask participants to adhere to a low-polyphenol diet (in comparison to a low-anthocyanin diet) to reduce the influence on the baseline. Such a low-polyphenol diet might exclude all fruits and vegetables, nuts and seeds, wholewheat and other high-fibre products, and beverages (tea, coffee and fruit juices), and alcohol (wine, beer, spirits) as done in a study by Ludwig et al [219].

A further important aspect is the metabolite panel used for the analysis in this chapter. The final metabolite panel presented in Chapter 3 consisted of seven metabolites, whereas the panel used here was the result of the original analysis and consisted only of four metabolites. The metabolites which were not included in this chapter are: syringic acid, dihydroferulic acid, and 4-hydroxyhippuric acid. Although berry anthocyanins are metabolised to a set of catabolites common among many polyphenols and therefore not unique[249], the metabolites assessed in this chapter are frequently reported to increase upon blueberry consumption. Particularly syringic acid could be of particular value as it is likely (albeit not exclusively) derived from malvidin, one of the major anthocyanins in blueberries and in Chapter 3 shown to increase the most relative to baseline following blueberry intake. If the metabolite panel is seen as a marker of blueberry intake, the ROC analysis in the previous chapter (see Figure 3-15) revealed that the reduced panel (four metabolites) was able to correctly predict blueberry intake in 77% of cases, while the predicting capability for the full panel (seven metabolites) was at 88%. Therefore, using the full panel instead of the reduced panel should increase specificity towards blueberry intake and reduce the impact of baseline variation due to other foods.

Another factor to consider is the effect of the food matrix the blueberry anthocyanins were in. The freeze-dried blueberry powder was dissolved in 500 g semi-skimmed milk, of which the proteins as well as the fat may have had an impact on the bioavailability of anthocyanins. Consumption of phenolics in conjunction with milk has been investigated in a few human cross-over studies with six to twenty participants for different foods and beverages, including strawberries[285], tea[286,287], coffee[288], and jujube juice[289]. Although the beverages used in the studies did not necessarily contain anthocyanins, they contained phenolic acids similar to the metabolites of anthocyanins or also present in blueberries, and included chlorogenic acids, hydroxycinnamic acids and hydroxybenzoic acids. Findings from the studies were equivocal: the consumption of 200g strawberries containing 96 mg pelargonidin-3-glucoside with or without cream did not significantly reduce the maximum plasma concentration (C_{max}) of the

anthocyanin, but it did significantly delay the absorption time to reach the peak concentration (t_{max}) by over 1h.[285] In contrast, Reddy et al found that catechin absorption from a cup of black tea (200 mg catechins) was dampened by 17%, but not delayed.[286] This finding was not observed by van Het Hof et al, who did not find any impact of milk on the absorption magnitude or timing of black tea catechins (dose per tea preparation was 300 mg).[287] Similarly, a study investigating the impact of milk and sugar plus non-dairy creamer on the bioavailability of downstream metabolites of chlorogenic acids (caffeic acid, ferulic acid, and isoferulic acid) from coffee containing 332 mg chlorogenic acids, reported no significant impact of milk addition to coffee.[288] However, the authors observed that sugar and non-dairy creamer delayed t_{max} of the phenolic acids by 45 to 120 minutes. A further study conducted by Zhang et al tested the absorption of phenolic acids from jujube juice in the presence of just juice, whole milk, skimmed milk, and milk fat.[289] This allowed them to consider the impact of milk protein (skimmed milk) and milk fat separately. The authors found that only whole milk reduced C_{max} by 25 – 35% and AUC by 30 – 44% of most phenolic acids, while skimmed milk delayed t_{max} by about 1 – 1,5h. Milk fat by itself did not have a significant impact on the pharmacokinetic parameters. The authors concluded that interactions between milk proteins and phenolic acids were the cause of the impaired absorption, but added that both milk protein and fat were required for this inhibitory effect. It seems that milk proteins may bind phenolic acids in the stomach and could thereby delay the transit time through the GI tract, while the fat contained in milk may reduce the absorption of phenolic acids.

The contradictory findings across studies could be due to differences in phenolic compounds and the type of milk product used, as the binding of milk proteins to the phenolic compounds likely depends on their structure[288] and the reduction in absorption and delay in transit time seems to mostly rely on the fat content. Due to the use of semi-skimmed milk (i.e., reduced fat) for the blueberry milkshake in the present study, it seemed likely that the milk itself had little effect on the bioavailability, although it may have slightly delayed absorption time. However, due to the collection of 24h pooled urines, a delay in absorption time would not affect the results. In sum, based on the above, the milk base of the blueberry drink was not likely the cause for a low bioavailability of anthocyanin metabolites in the low group. In this study, a milk-based drink was chosen to provide the same intervention drink for the blueberry challenge as well as the second phase of the study. Future studies assessing an anthocyanin metabolite profile may benefit from using a neutral matrix such as water to prevent any doubts of matrix effects on absorption and metabolism of the compounds of interest.

Moreover, gender is thought to potentially influence the bioavailability of flavonoids. While some gender-differences in phase I and II metabolism of flavonoids in the liver have been

observed, the evidence is limited.[150] Likewise, no effect of gender on metaboliser was observed in this study (see Figure 5-4).

The classification into metaboliser groups was designed *a priori* to be assessed at 48h because phenolic metabolites of anthocyanins were observed in urine for over 48h.[122,140,219] The classification at 24h and 48h was compared to determine if the additional day of urine collection had any benefits over a 24h assessment (see Figure 5-3). Although at 24h the majority of participants (76%) were classified as they have been at 48h, there were some considerable changes to groupings at 48h. The most striking were five participants who at 24h were ranked much lower and were grouped as medium and in one case even in the low group. These participants significantly increased in the urinary excretion of the metabolite panel and at 48h were much higher ranked and classified as high metabolisers. In addition, 20 further participants changed groups between 24 and 48h, moving from low to medium or medium to low. It seemed therefore that the additional day of urine collection provided information about a late increase in metabolite abundance in some participants which would have been missed at the 24h timepoint.

The metabolite panel used in this analysis to determine the metaboliser status was based on findings from a chronic study in which participants were repeatedly exposed to a dose of blueberries every day for six months (described in Section 2.1). It was unclear whether the panel expressed as $\log_2$ FC from baseline would be transferable to an acute study due to the differences in study design. The comparison can be seen in Figure 5-5. In theory, the assessment of metaboliser profile following the consumption of a dose of blueberries should be comparable between the chronic and acute studies due to the rapid metabolisation and subsequent elimination of anthocyanins and phenolic compounds. This was confirmed in the direct comparison of chronic intake (CIRCLES study) and single dose (AMP study), which showed that there was no significant difference when comparing the entire sample (see 'total' in Figure 5-5). However, the magnitude of change was stronger in the low and high groups in the AMP study in comparison to the CIRCLES study. These differences could be due to the differences in study design. While the dietary restrictions in CIRCLES were similar, the length of run-in time was much longer with 21 days versus 5 days in AMP. This may have affected the baseline measurement and led to a reduced response in the low group in CIRCLES. Furthermore, while the amount of anthocyanins was similar in the CIRCLES and AMP studies, differences in serving size (26 and 36 g respectively) may have provided significant differences in other components of blueberries such as phenolic acid and fibre content, which may have led to the increased response in the high group in AMP. Also, the study population was of similar age and weight, however CIRCLES was conducted in participants with metabolic syndrome, while participants in AMP were generally healthy. This may have had an impact on

the metabolisation and absorption of polyphenols. Finally, the data from CIRCLES was the result of daily intake of blueberries for six months, whereas in AMP, participants consumed blueberries only once. The long-term intake in CIRCLES would have sustained elevated levels of metabolites in the gut and in circulation which could have implications on the long-term modulation of the gut microbiome. These implications would not be present in an acute, single-dose setting.

As discussed in the previous chapter, the polyphenol dietary interventions may function as prebiotics and enhance proliferation of certain bacterial species, such as *Bifidobacterium*[261], which was associated with a greater abundance of some anthocyanin metabolites[260]. Such changes in the gut microbial composition in response to dietary interventions can vary highly between individuals due to the variability in the gut microbial composition between people and is likely affected by the presence and abundance of the bacterial species prior to the intervention[262]. It may therefore be speculated that with or without a prebiotic effect of blueberries on the gut microbiota after a single or repeated dose, the relative abundance of microbial metabolites may be a reflection of the gut microbial composition pre-intervention and differentiate a high from a low metaboliser irrespective of a long-term chronic intake or a single acute exposure.

The results from the chronic CIRCLES study further provided the basis for the calculation of a minimum sample size to detect distinct metaboliser groups using the metabolite Panel 5. The minimum target of n = 107 was achieved and exceeded, allowing the definition of distinct groups. Indeed, it is likely that the original target overestimated the actual required minimum sample size due to key differences between the AMP and CIRCLES studies:

- The estimate for the high metaboliser group used for the sample size calculation was derived from the blueberry treatment group and included a mix of high, medium, and low metabolisers. This means that this estimate for the high group was likely smaller than the actual value. Consequently, the estimate for the group difference between high and low was likely smaller, resulting in a greater sample size in the calculation.
- The urine collection window in CIRCLES was 24h, whereas in the AMP study urine was collected up to 48h. As visible from Figure 5-5, the magnitude of change was already greater for high metabolisers in the AMP study compared to the CIRCLES study and would likely be even greater at 48h. Figure 5-1 shows that while the high group continued to increase between 24 and 48h, while the low group did not, further driving up the group difference. This means again, that the actual group difference was very likely underestimated, resulting in a greater sample size in the calculation.
- The AMP intervention product, while similar in anthocyanin content to the CIRCLES product, contained greater amounts of other phenolic acids, which may cause an

increased response in the high metaboliser group in the AMP study and therefore potentially a greater group difference.

In summary, these differences likely led to an underestimation of the difference between high and low metaboliser groups and overestimation of the required minimum sample size for the AMP study to detect two distinct groups. The mean difference between high and low metaboliser groups requires further confirmation in validation and reproducibility studies, ideally using the full panel (Panel 2) instead of the reduced panel (Panel 5) to increase specificity towards blueberry intake

Conclusion

Overall, in this study, participants were classified into three distinct anthocyanin metabolic profiles following the intake of a single dose of blueberries using the metabolite panel consisting of the four metabolites identified in Section 3.2.3. Those identified as high and low metabolisers were prospectively recruited to the cross-over phase of the AMP study, which assessed cardiometabolic responses in each group after consumption of a test meal with or without blueberries. The outcome of this analysis (Phase 2) will be of great interest and may possibly provide some insight whether the grouping by metabolic profile has clinical significance.

6 General discussion and future perspectives

This thesis set out to assess the inter-individual variability in response to intake of anthocyanins from blueberries and to address the question of whether differential phenotypes for the metabolisation of anthocyanins existed and could be characterised based on the urinary excretion of a single or a group of metabolites. It also set out to determine the relationship between such a urinary metabolite profile and cardiometabolic markers, including vascular function. These objectives were addressed in Chapter 3 using retrospective data from a six-month long intervention study known as the 'CIRCLES study', in which participants with metabolic syndrome received a daily bolus of 26 g freeze-dried blueberry powder (providing 364 mg anthocyanins). The final aim of this thesis was to assess whether the same urinary metabolite profile could be applied in an ongoing dietary intervention study to prospectively recruit individuals as high, medium, or low anthocyanin metabolisers. This objective was addressed in Chapter 5, in which participants of the AMP study received a single dose of blueberries (providing 382 mg of anthocyanins) as a 'blueberry challenge'.

Inter-individual variability in response to flavonoid intake relates to differences in the absorption and metabolisation of flavonoids between individuals and affects the bioavailability of bioactive compounds for the individual. It is thought to be partly responsible for the contradictory evidence on health benefits obtained from human clinical trials. As such, there is a well-established body of research which highlights the importance of addressing this issue in research.[151,156,202–208] The potential benefits of understanding more about inter-individual variability are considered to be wide-ranging, including a greater clarification for effective dietary recommendations for flavonoids and other plant bioactives, the potential to apply personalised nutritional strategies (unique guidelines tailored to the individual) based on response status, improvements in healthy dietary choices, advancement in food technologies to enhance bioavailability of polyphenols from foods and supplements, and the creation of databases which aggregate knowledge on the metabolism of plant bioactives in humans.[290]

Wide inter-individual variability in the response to anthocyanin intake has been reported previously in human studies giving a variety of anthocyanin-rich foods and beverages, such as a bilberry and lingonberry puree (650 mg anthocyanins)[122], 500 mg ^{13}C-labelled cyanidin-3-glucoside[123,140], wild blueberry beverage (350 mg anthocyanins)[218], raspberry supplement (182 mg anthocyanins)[219], elderberry extract (500 mg anthocyanins)[245], black currant juice (1029 mg anthocyanins).[264] In addition, the major urinary metabolites which are observed in response to anthocyanin intake are fairly consistent across literature. Commonly reported major urinary metabolites include hippuric acid, 3-hydroxyhippuric acid, 4-hydroxyhippuric acid, vanillic acid, homovanillic acid, ferulic acid, and 3,4-Dihydroxyphenylacetic acid. However, no clear clusters

of metabolites which might differentiate individuals by their metabolic phenotype have been described for anthocyanins.

In Chapter 3, a mixed approach of multivariate and univariate data analysis was utilized to identify an aggregating group of anthocyanin metabolites, which in combination are proposed to describe a 'high metaboliser', with the expectation that such an individual may benefit more from the consumption of anthocyanin-rich blueberries than lower metabolisers. In the first, multivariate analysis, an exploratory factor analysis was performed, summarising the measured metabolite variables with two new variables (factor scores 1 and 2). Subsequently, by plotting the two factor scores against each other, clusters of participants were identified, demonstrating the variability in metabolic profiles (see Figure 3-10). Of the four participant clusters, two were diametrically opposed in their metabolite profile. It is these two groups that were of further interest: those who excreted high amounts of metabolites associated with both factors 1 and 2 were classified as high metabolisers and those who excreted limited amounts of metabolites associated with both factors were classified as low metabolisers. In the follow-up, univariate analysis, a detailed comparison of the metabolic profiles of high and low metabolisers was conducted. This resulted in a selection of seven metabolites which could be used for future research as a marker of a high metaboliser profile (see Table 3-13): 4-hydroxyhippuric acid, 3-hydroxyhippuric acid, hippuric acid, syringic acid, homovanillic acid (4-Hydroxy-3-methoxyphenylacetic acid), dihydroferulic acid (4-hydroxy-3-methoxyphenylpropionic acid), and 3,5-dihydroxyphenylpropionic acid.

The proposed characterisation of a high metaboliser, using this metabolite panel, is supported by the evidence that these metabolites have been frequently observed as the major urinary metabolites following anthocyanin intake in other human studies.[122,123,218,219,225,226,245,263,264] Only 3,5-dihydroxyphenylpropionic acid is reported less frequently, however the reason for this may be that this compound is less frequently determined in metabolite analyses, including in those publications cited. This could be because the six most common anthocyanins have a hydroxyl group at the 4'-position of the B-ring, and the heritage of 3,5-dihydroxyphenylpropionic acid is not obvious. Related compounds, however, have been reported. 3,5-dihydroxybenzoic acid was observed as a urinary metabolite in an acute human study following anthocyanin intake from elderberry extract[245] and *in vitro* fermentation of raspberry anthocyanins found a significant accumulation of resorcinol (benzene-1,3-diol)[242]. This is relevant because 3,5-dihydroxyphenylpropioinic acid could potentially be derived from the dehydroxylation of the anthocyanin delphinidin, which has a trihydroxylated B- ring[265] and subsequently converted to 3,5-dihydroxybenzoic acid through microbial β-oxidation and then to resorcinol through decarboxylation. It should be noted, however, that 3,5-dihydroxybenzoic acid and resorcinol may also be derived from the A-ring. As 3,5-dihydroxyphenylpropionic acid

may be a metabolite of alkylresorcinols present in wholegrain products,[241] this metabolite requires additional validation in future studies.

A key finding in Chapter 3 was that change in FMD, a measure of vascular function, when added to the exploratory factor analysis model, was associated with factor 2, but not with factor 1. On the basis of known metabolic pathways of anthocyanins and other flavonoids, the two factors were hypothesised to represent the following: factor 1) 'early' metabolites absorbed in the small intestine (dietary phenolic acids, initial degradation products, and their phase I and II metabolites); and factor 2) 'late' metabolites, which were likely low molecular weight products of the colonic gut microbial metabolism (see Section 3.1.4). Moreover, when comparing the high and low metabolisers from the cluster analysis, i.e., high versus low excretors of factor 1 and 2 metabolites, the high metaboliser group exhibited a greater increase in FMD than the low metabolisers, although the group difference was not statistically significant at a level of 0.05 (mean difference = 1.85, p = 0.235) (see Figure 3-10D). Together, these findings suggested that the change in FMD was aligned with the accumulation of colonic microbial metabolites in urine. This result supported the hypothesis that the low molecular weight phenolic catabolites play a key role as the bioactive mediators of clinical effects rather than the parent compounds[123]. This result was further supported by the fact that in our study, FMD significantly increased with the blueberry treatment after six months in comparison to the placebo treatment, but no parent anthocyanins were detected in urine or serum.[159] In addition, the bioactivity of phenolic metabolites was shown in an *in vitro* study, in which anthocyanin metabolites, but not the parent compounds, reduced palmitate-induced endothelial inflammation in human aortic endothelial cells.[143] The mixture of metabolites included hydroxyhippuric acid, hippuric acid, benzoic acid-4-sulfate, isovanillic acid-3-sulfate, and vanillic acid-4-sulfate, and reduced monocyte binding and expression of inflammatory chemokines and adhesion molecules, up to 50%. Palmitate-induced endothelial dysfunction in aortic segments of mice was also reduced and vasodilation recovered by about 50% when aortic segment was treated with the blueberry metabolites. Notably, the experiment was performed with metabolites in physiologically relevant concentrations.

The clinical significance of the metabolic profile was further assessed by creating a composite measure of the metabolite panel. The panel was expressed as the average $\log_2$ fold change from baseline, calculated for all participants in the CIRCLES study irrespective of treatment group, i.e., including the full dose group (364 mg anthocyanins per day), the half dose group (182 mg anthocyanins per day), and the placebo group (0 mg anthocyanins per day). The rationale to include the other treatment groups was that if the metabolites truly played a role in mediating vascular function, as speculated in the previous paragraph, the origin of the metabolites should not matter, i.e., the placebo group may have experienced similar benefits

if other food sources led to high increases in metabolite levels. Moreover, this would indicate a certain transferability of the panel to other flavonoid-rich food sources, given that many flavonoids are degraded to a similar set of catabolites.[157] The $\log_2$ fold change of the panel was significantly associated with change in FMD as well as change in cGMP and nitrite. Particularly the association with cGMP and nitrite could potentially provide some mechanistic insight into the bioactivity of the phenolic acids, as vasodilation is largely mediated through the NO/cGMP pathway. Nitrite is a product of NO metabolism and through eNOS stimulation and inhibition experiments has been found to reflect eNOS activity.[244] Therefore, the increase in nitrite observed in this study could indicate an increased NO bioavailability in high metabolisers, which in turn would lead to an elevation in cGMP levels and improvement in FMD. These results were in line with another study which gave a similar amount of bilberry and black currant anthocyanins in purified form (320 mg) for 12 weeks.[90] The authors reported a significant increase in FMD and cGMP in the anthocyanin group compared to placebo (increase from baseline by 28.4% and 12.6%, respectively) as well as a positive correlation between change in FMD and change in cGMP ($r = 0.428$, $p < 0.001$). In addition, the authors observed in an acute sub-study that the effects of anthocyanins on FMD were blocked by the presence of a NOS inhibitor. Similarly, in the same paper they also found that the effect of anthocyanins on vasorelaxation was abolished with a NOS inhibitor in an *in vitro* model of rat thoracic aortic rings, and in sum concluded that anthocyanins act through the NO/cGMP pathway. Furthermore, there was an indication that an increase in the panel may be associated with a reduction in the inflammatory markers TNF-α and CRP. TNF-α has been linked to a decrease in NO bioavailability in two ways: by inhibiting NO synthesis and by promoting the removal of NO. TNF-α inhibits eNOS activity, potentially via downregulation of eNOS expression, and stimulates ROS production through activation of the NADPH oxidase, which in turn react with NO.[238] Similarly, CRP also inhibits eNOS function by downregulating eNOS expression and mediates uncoupling of eNOS.[269] The uncoupling converts eNOS from producing NO to producing superoxide radicals, thereby not only is NO bioavailability reduced, but also oxidative stress is promoted which contributes to a further reduction of available NO and promotes endothelial dysfunction. Therefore, a potential reduction in inflammatory markers could point to an increase in NO bioavailability and improvement in endothelial function.

Importantly, the clinical effects of blueberry consumption may not only be mediated through action of phenolic acids after absorption from the gut, but may also be mediated through prebiotic effects on the gut microbiota, i.e., proliferation of beneficial bacteria, promoting an improved gut health. Clinical effects may therefore also be the result of other (non-polyphenolic) bacterial-derived bioactive metabolites. Increasing evidence suggests that the gut microbiota and their metabolites such as short-chain fatty acids (SCFA, a product of fibre fermentation) and lipopolysaccharides (LPS, also called endotoxin) could be linked with

cardiovascular disease.[291] SCFA, such as butyrate, are the main energy source for colonocytes and are critical for the maintenance of the gut mucosal barrier. A reduction in bacterial species which produce butyrate can lead to a dysfunctional gut barrier, leading to increased permeability of the gut and increased plasma LPS levels, which trigger inflammatory processes. A study in which participants with metabolic syndrome consumed red wine polyphenols (272 mL per day) for one month observed an increase in the levels of beneficial bacteria including intestinal barrier protectors *Bifidobacterium* and *Lactobacillus* (increase by 50%), and butyrate-producers *Faecalibacterium prausnitzii* and *Roseburia* (increase by 36% and 28%, respectively).[261] They also observed decreased numbers of bacteria associated with systemic inflammatory marker LPS (lipopolysaccharide) including *Escherichia coli* and *Enterobacter cloacae* (decrease by 21% and 23%, respectively). The authors also reported significant improvements in biomarkers for metabolic syndrome following red wine intake which could be related to the observed changes in the microbiota, including blood pressure (SBP: -10.5 mmHg; DBP: -9.2 mmHg), blood glucose (-29.7 mg/dl), HDLC (+9.4 mg/dl), and LPS (-0.17 endotoxin units/ml). Of note, a different study reported that the increase in *Bifidobacterium* after ingestion of anthocyanins from 272 ml red wine for 20 days was associated with increases in four anthocyanin metabolites: *p*-coumaric acid, syringic acid, and in particular homovanillic acid and 4-hydroxybenzoic acid.[260] In the highest tertile of change in *Bifidobacterium* metabolites significantly increased by 2.05, 1.37, 20.28, and 18.04 µmol/24h, respectively. Two of these metabolites (syringic acid and homovanillic acid) were part of the metabolite panel identified in Chapter 3. Although 4-hydroxybenzoic acid was not directly part of the panel, its glycinated phase II metabolite 4-hydroxyhippuric acid was. Also, *p*-coumaric acid (4-hydroxycinnamic acid) was not observed to be a major metabolite in this study, but as Boto-Ordóñez propose, *p*-coumaric acid could potentially have been degraded to 4-hydroxybenzoic acid and subsequently excreted as 4-hydroxyhippuric acid.

Chapter 5 describes the first phase of the AMP study, the blueberry challenge, in which the urinary metabolite panel identified in Chapter 3 (expressed as the $\log_2$ fold change from baseline) was used to classify individuals as high, medium, or low metabolisers for the prospective recruitment to a subsequent dietary cross-over intervention (phase 2 of the AMP study). In contrast to the CIRCLES study, in which participants were exposed to a daily dose of blueberries for six months, the blueberry challenge of the AMP study involved a single exposure and urine collections at baseline (prior to the blueberry bolus) and for 48h after consumption. It should be noted that a reduced version of the metabolite panel with four instead of seven metabolites was used on the basis of an initial analysis (see Section 2.5, note to examiners for details). The panel presented in Chapter 3 represents the result of a refined analysis performed after recruitment to the AMP study was already completed.

In contrast to the metabolism of some flavonoids which result in unique microbial metabolites, such as equol from the isoflavone daidzein[292] or urolithins from ellagitannins[247], berry anthocyanin metabolites are generally small molecule phenolic acids also common to other flavonoids. A metabolic response to the consumption of berry anthocyanins therefore rather reflects a significant increase in metabolites which are already present at baseline[249] and stratification between participants as 'responders' and 'non-responders' to anthocyanin intake is not clear-cut. Hence, in Chapter 5 participants were stratified by taking the 15th and 85th percentiles of the distribution of the panel as cut-offs for low, medium, and high metabolisers. This resulted in three groups with distinct responses to anthocyanin intake over 48h. The steepest increase for the panel was observed for the high group, whereas the low group first decreased slightly at 24h and then returned to baseline at 48h. The use of log_2 fold change as a measure has the advantage of weighting each metabolite in relation to its baseline value rather than the absolute amount. This makes metabolites which vary greatly in their abundance comparable. For example, mean baseline excretion of hippuric acid was about 50 times more abundant than homovanillic acid. The disadvantage, however, is that the log_2 FC will become very large if the baseline value is small, even when the absolute change from baseline is also very small. The method of assessing change from baseline for the stratification of participants should therefore be evaluated individually per study due to variability in baselines. For example, in the AMP study as well as in the CIRCLES study, the mean baseline value for homovanillic acid was similar at 16 µmol/24h. However, Ludwig et al did not detect any homovanillic acid in urine at baseline.[219] In this case log_2 fold change would not be applicable, as the logarithm of 0 is undefined.

Furthermore, a high inter-individual variation for the metabolites was observed at baseline. Such a variation is not surprising and in large parts due to differences in the individual genetics, gut microbial composition, and environmental factors such as lifestyle and diet.[293] A study design in which individuals are repeatedly measured in a cross-over design (intervention and placebo) can be helpful in accounting for not only inter-individual variation, but intra-individual variation by assessing the reproducibility of the metabolic profile after repeated exposures. These type of studies are termed 'N-of-1' studies and focus on participant-specific factors and the variation observed within each participant rather than across participants.[293]

In refinement of the current approach, it would be of interest to assess whether the metabotype assessed at baseline was still the same after a longer period of time or whether the intervention had, for example, influenced the gut microbial composition which then led to changes in the metabolic profile. The justification to pursue this line of enquiry is based on a recent assessment, where the reproducibility of differences in the gut microbiota was demonstrated in a placebo-controlled cross-over study, in which participants consumed pomegranate extract

in two different doses (providing 160 mg or 640 mg phenolics).[156] In this study, urolithin producer metabotype was assessed at the beginning of each intervention period (three times in total; at baseline and once after each dose). Three participants, who at baseline were of urolithin metabotype 0, changed their metabotype after the second dose to a urolithin metabotype A. The authors also assessed the modulation of the gut microbiota through the pomegranate extract, focusing on the bacteria relevant for urolithin production. They found that those of urolithin metabotype 0 had lower levels of *Gordonibacter* than the other metabotypes. However, the three participants who switched from metabotype 0 to metabotype A presented higher baseline levels of *Gordonibacter* than the other participants who were of metabotype 0. The consumption of the higher dose of pomegranate extract in the 'metabotype 0 responders' was associated with a significant increase in growth of *Gordonibacter*.

Of course, of particular interest is the clinical significance of the stratification by metaboliser profile. In the AMP study, those who were classified as high or low metabolisers were subsequently invited to take part in the second, cross-over phase of the AMP study. The outcome of this second phase, which assessed cardiometabolic markers after the consumption of a test meal with and without blueberries, will be of great interest and may possibly provide some insight whether the grouping by the metabolic profile as assessed in Chapter 5 was associated with clinical effects. This analysis is ongoing and not part of this thesis. However, from an aligned example from literature, the pomegranate study mentioned above highlighted how inter-individual variability in response to flavonoid intervention can clarify a potential clinical effect of bioactive phenolics.[156] The study showed that the functional stratification by urolithin producer metabotypes mediated the beneficial changes observed for blood lipids (e.g., reduction in small LDLC and oxidized LDLC). Significant reductions were only observed for those of urolithin metabotype B, but not metabotype A (small LDLC: -47%; oxLDLC: -24%). However, when viewed across the entire sample irrespective of metabotype, no significant effect was observed.

From a public health perspective, the stratification of participants into high and low metabolisers or responders and non-responders begs the question how those who are not high metabolisers can benefit from these findings. Determining what factors determine a low and high metaboliser status is the essence of future efforts in disentangling inter-individual variability. The interplay between gut microbiota and metabolism of polyphenols is clearly relevant for the bioavailability and bioactivity of phenolic acids and the health benefits an individual receives. Also, for anthocyanins it seems rather well established that microbial metabolites are likely the bioactive molecules. However, the clinical significance of a high versus low metaboliser of anthocyanins remains to be assessed. How continued habitual intake, different doses, and even probiotics may influence the anthocyanin metabolic profile of

individuals is an exciting topic for the future, whether to develop personalized nutritional strategies or improve the granularity of dietary recommendations for the public.

7 References

1. Witkamp RF, van Norren K. Let thy food be thy medicine….when possible. *Eur J Pharmacol*. 2018;836(May):102-114. doi:10.1016/j.ejphar.2018.06.026

2. Afshin A, Sur PJ, Fay KA, et al. Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2019;393(10184):1958-1972. doi:10.1016/S0140-6736(19)30041-8

3. World Health Organisation. Noncommunicable diseases.

4. Abbafati C, Abbas KM, Abbasi-Kangevari M, et al. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*. 2020;396(10258):1223-1249. doi:10.1016/S0140-6736(20)30752-2

5. Institue for Health Metrics and Evaluation (University of Washington). GBP compare - Global Burden of Disease 2019 Vizhub. Published 2021. https://vizhub.healthdata.org/gbd-compare/#

6. Katan MB. Health claims for functional foods. 2004;328(January):180-181.

7. Seidman MA, Mitchell RN, Stone JR. Pathophysiology of Atherosclerosis. *Cell Mol Pathobiol Cardiovasc Dis*. Published online 2014:221-237. doi:10.1016/B978-0-12-405206-2.00012-0

8. Davignon J, Ganz P. Role of Endothelial Dysfunction in Atherosclerosis. *Circulation*. 2004;109(23_suppl_1). doi:10.1161/01.CIR.0000131515.03336.f8

9. Eckel RH, Kahn R, Robertson RM, Rizza RA. Preventing cardiovascular disease and diabetes: A call to action from the American Diabetes Association and the American Heart Association. *Circulation*. 2006;113(25):2943-2946. doi:10.1161/CIRCULATIONAHA.106.176583

10. Eckel RH, Grundy SM, Zimmet PZ. The metabolic syndrome. *Lancet*. 2005;365(9468):1415-1428. doi:10.1016/S0140-6736(05)66378-7

11. Gregor MF, Hotamisligil GS. Inflammatory mechanisms in obesity. *Annu Rev Immunol*. 2011;29:415-445. doi:10.1146/annurev-immunol-031210-101322

12. Health Survey England 2019. Accessed December 7, 2021. http://healthsurvey.hscic.gov.uk/support-guidance/public-health/health-survey-for-england-2019/key-findings.aspx

13. Després J-P, Moorjani S, Lupien PJ, Tremblay A, Nadeau A, Bouchard C. Regional Distribution of Body Fat,m Plkasma Lipoproteins, and Cardiovascular Disease. *Arterioscler Thromb Vasc Biol*. 1990;10(4):497-511. doi:10.1161/ATVBAHA.111.241919

14. Björntorp P. " Portal " Adipose Tissue as a Generator of Risk Factors for Cardiovascular Disease and Diabetes. *Arter Thromb Vasc Biol*. 1990;10:493-497. doi:10.1161/01.ATV.10.4.493

15. Carey DG, Jenkins AB, Campbell L V, Freund J, Chisholm DJ. Abdominal Fat and Insulin Resistance in Normal and Overweight Women: Direct Measurements Reveal a Strong Relationship in Subjects at Both Low and High Risk of NIDDM. *Diabetes*. 1996;45(5):633 LP - 638. http://diabetes.diabetesjournals.org/content/45/5/633.abstract

16. Coutinho T, Goel K, Corrêa De S D, et al. Central obesity and survival in subjects with

coronary artery disease: A systematic review of the literature and collaborative analysis with individual subject data. *J Am Coll Cardiol*. 2011;57(19):1877-1886. doi:10.1016/j.jacc.2010.11.058

17. Dallongeville J, Bhatt DL, Steg PH, et al. Relation between body mass index, waist circumference, and cardiovascular outcomes in 19,579 diabetic patients with established vascular disease: the REACH Registry. *Eur J Prev Cardiol*. 2012;19(2):241-249. doi:10.1177/1741826710394305

18. Rexrode KM, Carey VJ, Hennekens CH, et al. Abdominal adiposity and coronary heart disease in women. *J Am Med Assoc*. 1998;280(21):1843-1848. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed4&NEWS=N&AN=1998405511

19. Sam S, Haffner S, Davidson MH, et al. Relation of abdominal fat depots to systemic markers of inflammation in type 2 diabetes. *Diabetes Care*. 2009;32(5):932-937. doi:10.2337/dc08-1856

20. Sam S, Haffner S, Davidson MH, et al. Relationship of Abdominal Visceral and Subcutaneous Adipose Tissue With Lipoprotein Particle Number and Size in Type 2 Diabetes. *Diabetes*. 2008;57(August):2022-2027. doi:10.2337/db08-0157.

21. Boden G. Obesity, insulin resistance and free fatty acids. *Curr Opin Endocrinol Diabetes Obes*. 2003;18(2):139-143. doi:10.1097/MED.0b013e3283444b09.45Obesity

22. Pereira S, Park E, Mori Y, et al. FFA-induced Hepatic Insulin Resistance in vivo is mediated by PKC-δ, NADPH Oxidase, and Oxidative Stress. *Am J Physiol Endocrinol Metab*. 2014;307(1):E34-46. doi:10.1152/ajpendo.00436.2013

23. Santomauro AT, Boden G, Silva ME, et al. Overnight lowering of free fatty acids with Acipimox improves insulin resistance and glucose tolerance in obese diabetic and nondiabetic subjects. *Diabetes*. 1999;48(9):1836-1841.

24. Lage R, Diéguez C, Vidal-Puig A, López M. AMPK: a metabolic gauge regulating whole-body energy homeostasis. *Trends Mol Med*. 2008;14(12):539-549. doi:10.1016/j.molmed.2008.09.007

25. Steinberg GR, McAinch AJ, Chen MB, et al. The suppressor of cytokine signaling 3 inhibits leptin activation of AMP-kinase in cultured skeletal muscle of obese humans. *J Clin Endocrinol Metab*. 2006;91(9):3592-3597. doi:10.1210/jc.2006-0638

26. Moraes-Vieira PM, Yore MM, Dwyer PM, Syed I, Aryal P, Kahn BB. RBP4 activates antigen-presenting cells, leading to adipose tissue inflammation and systemic insulin resistance. *Cell Metab*. 2014;19(3):512-526. doi:10.1016/j.cmet.2014.01.018

27. Norseen J, Hosooka T, Hammarstedt A, et al. Retinol-binding protein 4 inhibits insulin signaling in adipocytes by inducing proinflammatory cytokines in macrophages through a c-Jun N-terminal kinase- and toll-like receptor 4-dependent and retinol-independent mechanism. *Mol Cell Biol*. 2012;32(10):2010-2019. doi:10.1128/MCB.06193-11

28. Shi H, Kokoeva M V, Inouye K, Tzameli I, Yin H, Flier JS. TLR4 links innate immunity and fatty acid – induced insulin resistance. *J Clin Invest*. 2006;116(11):3015-3025. doi:10.1172/JCI28898.TLRs

29. Scheller J, Chalaris A, Schmidt-Arras D, Rose-John S. The pro- and anti-inflammatory properties of the cytokine interleukin-6. *Biochim Biophys Acta - Mol Cell Res*. 2011;1813(5):878-888. doi:10.1016/j.bbamcr.2011.01.034

30. Mauer J, Chaurasia B, Goldau J, et al. Signaling by IL-6 promotes alternative activation of macrophages to limit endotoxemia and obesity-associated resistance to

insulin. *Cell Metab*. 2014;15(3):403-416. doi:10.1038/ni.2865

31. Kim HJ, Higashimori T, Park SY, et al. Differential effects of interleukin-6 and-10 on skeletal muscle and liver insulin action in vivo. *Diabetes*. 2004;53(4):1060-1067. doi:10.2337/diabetes.53.4.1060

32. Sears B, Perry M. The role of fatty acids in insulin resistance. *Lipids Health Dis*. 2015;14(1):121. doi:10.1186/s12944-015-0123-1

33. Thaler JP, Guyenet SJ, Dorfman MD, Wisse BE, Schwartz MW. Hypothalamic inflammation: Marker or mechanism of obesity pathogenesis? *Diabetes*. 2013;62(8):2629-2634. doi:10.2337/db12-1605

34. Feingold KR, Grunfeld C. Tumor necrosis factor-alpha stimulates hepatic lipogenesis in the rat in vivo. *J Clin Invest*. 1987;80(1):184-190. doi:10.1172/JCI113046

35. Hoogeveen RC, Gaubatz JW, Sun W, et al. Small dense low-density lipoprotein-cholesterol concentrations predict risk for coronary heart disease: The Atherosclerosis Risk in Communities (ARIC) study. *Arterioscler Thromb Vasc Biol*. 2014;34(5):1069-1077. doi:10.1161/ATVBAHA.114.303284

36. Ali KM, Wonnerth A, Huber K, Wojta J. Cardiovascular disease risk reduction by raising HDL cholesterol - Current therapies and future opportunities. *Br J Pharmacol*. 2012;167(6):1177-1194. doi:10.1111/j.1476-5381.2012.02081.x

37. Vane JR, Änggård EE, Botting RM. Regulatory Functions of the Vascular Endothelium. *N Engl J Med*. 1990;323(1):27-36. doi:10.1056/NEJM199007053230106

38. Dhaun N, Goddard J, Kohan DE, Pollock DM, Schiffrin EL, Webb DJ. Role of endothelin-1 in clinical hypertension: 20 years on. *Hypertension*. 2008;52(3):452-459. doi:10.1161/HYPERTENSIONAHA.108.117366

39. Münzel T, Sinning C, Post F, Warnholtz A, Schulz E. Pathophysiology, diagnosis and prognostic implications of endothelial dysfunction. *Ann Med*. 2008;40(3):180-196. doi:10.1080/07853890701854702

40. Esper RJ, Nordaby RA, Vilariño JO, Paragano A, Cacharrón JL, Machado RA. Endothelial dysfunction: A comprehensive appraisal. *Cardiovasc Diabetol*. 2006;5:1-18. doi:10.1186/1475-2840-5-4

41. Korkmaz H, Onalan O. Evaluation of Endothelial Dysfunction: Flow-Mediated Dilation. *Endothelium*. 2008;15(4):157-163. doi:10.1080/10623320802228872

42. Hadi HAR, Carr CS, Al Suwaidi J. Endothelial dysfunction: cardiovascular risk factors, therapy, and outcome. *Vasc Health Risk Manag*. 2005;1(3):183-198.

43. Barac A, Campia U, Panza JA. Methods for evaluating endothelial function in humans. *Hypertension*. 2007;49(4):748-760. doi:10.1161/01.HYP.0000259601.38807.a6

44. Urso C, Caimi G. [Oxidative stress and endothelial dysfunction]. *Minerva Med*. 2011;102(1):59—77.

45. Teixeira BC, Lopes AL, Macedo RCO, et al. Inflammatory markers, endothelial function and cardiovascular risk. *J Vasc Bras*. 2014;13:108-115.

46. Tran V, De Silva TM, Sobey CG, et al. The Vascular Consequences of Metabolic Syndrome: Rodent Models, Endothelial Dysfunction, and Current Therapies. *Front Pharmacol*. 2020;11(March):1-10. doi:10.3389/fphar.2020.00148

47. Muraki I, Imamura F, Manson JE, et al. Fruit consumption and risk of type 2 diabetes: results from three prospective longitudinal cohort studies. *BMJ*. 2013;347. doi:10.1136/bmj.f5001

48. Zamora-Ros R, Touillaud M, Rothwell JA, Romieu I, Scalbert A. Measuring exposure to the polyphenol metabolome in observational epidemiologic studies: Current tools and applications and their limits. *Am J Clin Nutr*. 2014;100(1):11-26. doi:10.3945/ajcn.113.077743

49. Salvini S, Hunter DJ, Sampson L, et al. Food-based validation of a dietary questionnaire: The effects of week-to-week variation in food consumption. *Int J Epidemiol*. 1989;18(4):858-867. doi:10.1093/ije/18.4.858

50. Feskanich D, Rimm EB, Giovannucci EL, et al. Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire. *J Am Diet Assoc*. 1993;93(7):790-796. doi:https://doi.org/10.1016/0002-8223(93)91754-E

51. Hu FB, Rimm E, Smith-Warner SA, et al. Reproducibility and validity of dietary patterns assessed with a food- frequency questionnaire. *Am J Clin Nutr*. 1999;69(2):243-249. doi:10.1093/ajcn/69.2.243

52. Jager KJ, Zoccali C, MacLeod A, Dekker FW. Confounding: What it is and how to deal with it. *Kidney Int*. 2008;73(3):256-260. doi:10.1038/sj.ki.5002650

53. Agardh E, Allebeck P, Hallqvist J, Moradi T, Sidorchuk A. Type 2 diabetes incidence and socio-economic position: A systematic review and meta-analysis. *Int J Epidemiol*. 2011;40(3):804-818. doi:10.1093/ije/dyr029

54. Bertoia ML, Mukamal KJ, Cahill LE, et al. Changes in Intake of Fruits and Vegetables and Weight Change in United States Men and Women Followed for Up to 24 Years: Analysis from Three Prospective Cohort Studies. *PLoS Med*. 2015;12(9):1-20. doi:10.1371/journal.pmed.1001878

55. Valdes AM, Walter J, Segal E, Spector TD. Role of the gut microbiota in nutrition and health. *BMJ*. 2018;361:k2179. doi:10.1136/bmj.k2179

56. Makki K, Deehan EC, Walter J, Bäckhed F. The Impact of Dietary Fiber on Gut Microbiota in Host Health and Disease. *Cell Host Microbe*. 2018;23(6):705-715. doi:10.1016/j.chom.2018.05.012

57. US Department of Agriculture. USDA National Nutrient Database for Standard Reference, Legacy Release. Published 2019. Accessed December 23, 2021. https://fdc.nal.usda.gov/index.html

58. Siracusa L, Ruberto G. Chapter 2 - Plant Polyphenol Profiles as a Tool for Traceability and Valuable Support to Biodiversity. In: Watson RRBT-P in P, ed. Academic Press; 2014:15-33. doi:https://doi.org/10.1016/B978-0-12-397934-6.00002-4

59. Bertoia ML, Rimm EB, Mukamal KJ, Hu FB, Willett WC, Cassidy A. Dietary flavonoid intake and weight maintenance: three prospective cohorts of 124 086 US men and women followed for up to 24 years. *BMJ*. 2016;352:i17. doi:10.1136/bmj.i17

60. Fernandez de Simon B, Perez-Ilzarbe J, Hernandez T, Gomez-Cordoves C, Estrella I. Importance of phenolic compounds for the characterization of fruit juices. *J Agric Food Chem*. 1992;40(9):1531-1535. doi:10.1021/jf00021a012

61. Croteau R, Kutchan TM, Lewis NG. Natural Products (Secondary Metabolites). In: Buchanan B, Gruissem W, Jones R, eds. *Biochemistry Molecular Biology of Plants*. American Society of Plant Physiologists; 2000:1250-1318. doi:10.1016/j.phytochem.2011.10.011

62. Bravo L. Polyphenols: Chemistry, dietary sources, metabolism, and nutritional significance. *Nutr Rev*. 1998;56(11):317-333. doi:10.1111/j.1753-4887.1998.tb01670.x

63. Wursch P, Del Vendo S, Rosset J, Smiley M. The tannin granules from ripe carob pod.

Food Sci + Technol + Technol Aliment. 1984;17(6):351-354.

64. Manach C, Scalbert A, Morand C, Rémésy C, Jiménez L. Polyphenols: Food sources and bioavailability. *Am J Clin Nutr*. 2004;79(5):727-747. doi:10.1093/ajcn/79.5.727

65. Falcone Ferreyra ML, Rius SP, Casati P. Flavonoids: biosynthesis, biological functions, and biotechnological applications. *Front Plant Sci*. 2012;3(September):222. doi:10.3389/fpls.2012.00222

66. Manach C, Williamson G, Morand C, Scalbert A, Rémésy C. Bioavailability and bioefficacy of polyphenols in humans.I.Review of 97 bioavailability studies. *Am J Clin Nutr*. 2005;81:230S-242S.

67. Williamson G, Manach C. Bioavailability and bioefficacy of polyphenols in humans.II.Review of 93 intervention studies. *Am J Clin Nutr*. 2005;81:243S-255S.

68. Cassidy A, Mukamal KJ, Liu L, Franz M, Eliassen AH, Rimm EB. High anthocyanin intake is associated with a reduced risk of myocardial infarction in young and middle-aged women. *Circulation*. 2013;127(2):188-196. doi:10.1161/CIRCULATIONAHA.112.122408

69. Cassidy A, Bertoia M, Chiuve S, Flint A, Forman J, Rimm EB. Habitual intake of anthocyanins and flavanones and risk of cardiovascular disease in men. *Am J Clin Nutr*. 2016;104(3):587-594. doi:10.3945/ajcn.116.133132

70. Jennings A, Welch AA, Fairweather-tait SJ, et al. Higher anthocyanin intake is associated with lower arterial stiffness and central blood pressure in women. *Am J Clin Nutr*. 2012;96:781-788. doi:10.3945/ajcn.112.042036.

71. Bhagwat S, Haytowitz D, Holden J. USDA Database for the Flavonoid Content of Selected Foods, Release 3.1. Published online 2014.

72. Bhagwat S, Haytowitz DB. USDA Database for the Proanthocyanidin Content of Selected Foods, Release 2 (2015). *US Dep Agric Agric Res Serv*. Published online 2015:69. https://doi.org/10.15482/USDA.ADC/1324621.

73. Andersen OM, Jordheim M. Basic Anthocyanin Chemistry and Dietary Sources. In: Wallace TC, Giusti MM, eds. *Anthocyanins in Health and Disease*. Taylor & Francis Group; 2014:13-90.

74. Wallace TC, Giusti MM. Anthocyanins. *Adv Nutr*. 2015;6(5):620-622. doi:10.3945/an.115.009233

75. Kong J-M, Chia L-S, Goh N-K, Chia T-F, Brouillard R. Analysis and biological activities of anthocyanins. *Phytochemistry*. 2003;64(5):923-933. doi:10.1016/S0031-9422(03)00438-2

76. Andersen ØM, Jordheim M. Anthocyanins. *Encycl Life Sci*. 2010;(October). doi:10.1002/9780470015902.a0001909.pub2

77. Miguel MG. Anthocyanins: Antioxidant and/or anti-inflammatory activities. *J Appl Pharm Sci*. 2011;1(6):7-15.

78. García-Alonso M, Rimbach G, Rivas-Gonzalo JC, de Pascual-Teresa S. Antioxidant and Cellular Activities of Anthocyanins and Their Corresponding Vitisins AStudies in Platelets, Monocytes, and Human Endothelial Cells. *J Agric Food Chem*. 2004;52(11):3378-3384. doi:10.1021/jf035360v

79. Rossetto M, Vanzani P, Mattivi F, Lunelli M, Scarpa M, Rigo A. Synergistic antioxidant effect of catechin and malvidin 3-glucoside on free radical-initiated peroxidation of linoleic acid in micelles. *Arch Biochem Biophys*. 2002;408(2):239-245. doi:10.1016/S0003-9861(02)00561-1

80. Ziberna L, Tramer F, Moze S, Vrhovsek U, Mattivi F, Passamonti S. Transport and bioactivity of cyanidin 3-glucoside into the vascular endothelium. *Free Radic Biol Med*. 2012;52(9):1750-1759. doi:10.1016/j.freeradbiomed.2012.02.027

81. Blacker BC, Snyder SM, Eggett DL, Parker TL. Consumption of blueberries with a high-carbohydrate, low-fat breakfast decreases postprandial serum markers of oxidation. *Br J Nutr*. 2013;109(9):1670-1677. doi:10.1017/S0007114512003650

82. Velika B, Kron I. Antioxidant properties of benzoic acid derivatives against Superoxide radical. *Free Radicals Antioxidants*. 2012;2(4):62-67. doi:10.5530/ax.2012.4.11

83. Razzaghi-Asl N, Garrido J, Khazreaei H, Borges F, Firuzi O. Antioxidant Properties of Hydroxycinnamic Acids: A Review of Structure- Activity Relationships. *Curr Med Chem*. 2013;20(36):4436-4450 (15).

84. Dohadwala MM, Holbrook M, Hamburg NM, et al. Effects of cranberry juice consumption on vascular function in patients with coronary artery disease. *Am J Clin Nutr*. Published online 2011:934-940. doi:10.3945/ajcn.110.004242.

85. Johnson SA, Figueroa A, Navaei N, et al. Daily blueberry consumption improves blood pressure and arterial stiffness in postmenopausal women with pre- and stage 1-hypertension: a randomized, double-blind, placebo-controlled clinical trial. *J Acad Nutr Diet*. 2015;115(3):369-377. doi:10.1016/j.jand.2014.11.001

86. Li D, Zhang Y, Liu Y, Sun R, Xia M. Purified anthocyanin supplementation reduces dyslipidemia, enhances antioxidant capacity, and prevents insulin resistance in diabetic patients. *J Nutr*. 2015;145(4):742-748. doi:10.3945/jn.114.205674

87. Stull AJ, Cash KC, Johnson WD, Champagne CM, Cefalu WT. Bioactives in Blueberries Improve Insulin Sensitivity in Obese, Insulin-Resistant Men and Women. *J Nutr*. Published online 2010. doi:10.3945/jn.110.125336

88. Lynn A, Hamadeh H, Leung WC, Russell JM, Barker ME. Effects of Pomegranate Juice Supplementation on Pulse Wave Velocity and Blood Pressure in Healthy Young and Middle-aged Men and Women. *Plant Foods Hum Nutr*. 2012;67(3):309-314. doi:10.1007/s11130-012-0295-z

89. Erlund I, Koli R, Alfthan G, et al. Favorable effects of berry consumption on platelet function, blood pressure, and HDL cholesterol. *Am J Clin Nutr*. 2008;87(2):323-331. doi:87/2/323 [pii]

90. Zhu Y, Xia M, Yang Y, et al. Purified Anthocyanin Supplementation Improves Endothelial Function via NO-cGMP Activation in Hypercholesterolemic Individuals. *Clin Chem*. 2011;57(11):1524-1533. doi:10.1373/clinchem.2011.167361

91. Rodriguez-Mateos A, Rendeiro C, Bergillos-meca T, et al. Intake and time dependence of blueberry flavonoid – induced improvements in vascular function : a randomized , controlled , double-blind , crossover intervention study with mechanistic insights into biological activity 1 – 3. *Am J Clin Nutr*. 2013;98(Cvd):1179-1191. doi:10.3945/ajcn.113.066639.

92. Hooper L, Kroon PA, Rimm EB, et al. Flavonoids, flavonoid-rich foods, and cardiovascular risk: a meta-analysis of randomized controlled trials. *Am J Clin Nutr*. 2008;88(1):38-50. doi:10.1093/ajcn/88.1.38

93. Kay CD, Hooper L, Kroon PA, Rimm EB, Cassidy A. Relative impact of flavonoid composition, dose and structure on vascular function: A systematic review of randomised controlled trials of flavonoid-rich food products. *Mol Nutr Food Res*. 2012;56(11):1605-1616. doi:10.1002/mnfr.201200363

94. Warner EF, Smith MJ, Zhang Q, et al. Signatures of anthocyanin metabolites identified

in humans inhibit biomarkers of vascular inflammation in human endothelial cells. *Mol Nutr Food Res*. 2017;61(9):1-10. doi:10.1002/mnfr.201700053

95. Edirisinghe I, Banaszewski K, Cappozzo J, et al. Strawberry anthocyanin and its association with postprandial inflammation and insulin. *Br J Nutr*. 2011;106(6):913-922. doi:10.1017/S0007114511001176

96. DeFuria J, Bennett G, Strissel KJ, et al. Dietary blueberry attenuates whole-body insulin resistance in high fat-fed mice by reducing adipocyte death and its inflammatory sequelae. *J Nutr*. 2009;139(8):1510-1516. doi:10.3945/jn.109.105155

97. Qin Y, Xia M, Ma J, Hao Y, Liu J. Anthocyanin supplementation improves serum LDL- and HDL-cholesterol concentrations associated with the inhibition of cholesteryl ester transfer protein in dyslipidemic subjects. *Am J Clin Nutr*. 2009;90(2):485-492. doi:10.3945/ajcn.2009.27814.It

98. Jennings A, MacGregor A, Spector T, Cassidy A. Higher dietary flavonoid intakes are associated with lower objectively measured body composition in women: evidence from discordant monozygotic twins. *Am J Clin Nutr*. 2017;105(3):626-634. doi:10.3945/ajcn.116.144394

99. Jennings A, Welch AA, Spector T, Macgregor A, Cassidy A. Intakes of Anthocyanins and Flavones Are Associated with Biomarkers of Insulin Resistance and Inflammation in Women. *J Nutr*. 2014;144(16):202-208. doi:10.3945/jn.113.184358

100. Liu Y-J, Zhan J, Liu X-L, Wang Y, Ji J, He Q-Q. Dietary flavonoids intake and risk of type 2 diabetes: a meta-analysis of prospective cohort studies. *Clin Nutr*. 2014;33(1):59-63. doi:10.1016/j.clnu.2013.03.011

101. Afrin S, Giampieri F, Gasparrini M, et al. Chemopreventive and Therapeutic Effects of Edible Berries: A Focus on Colon Cancer Prevention and Treatment. *Molecules*. 2016;21(2):169. doi:10.3390/molecules21020169

102. Brown EM, Latimer C, Allsopp P, et al. In Vitro and in Vivo Models of Colorectal Cancer: Antigenotoxic Activity of Berries. *J Agric Food Chem*. 2014;62(18):3852-3866. doi:10.1021/jf4050759

103. Vauzour D. Dietary polyphenols as modulators of brain functions: Biological actions and molecular mechanisms underpinning their beneficial effects. *Oxid Med Cell Longev*. 2012;2012. doi:10.1155/2012/914273

104. Wang J, Ferruzzi MG, Ho L, et al. Brain-targeted proanthocyanidin metabolites for Alzheimer's disease treatment. *J Neurosci*. 2012;32(15):5144-5150. doi:10.1523/JNEUROSCI.6437-11.2012

105. Travica N, D'Cunha NM, Naumovski N, et al. The effect of blueberry interventions on cognitive performance and mood: A systematic review of randomized controlled trials. *Brain Behav Immun*. 2020;85:96-105. doi:10.1016/j.bbi.2019.04.001

106. Whyte AR, Rahman S, Bell L, et al. Improved metabolic function and cognitive performance in middle-aged adults following a single dose of wild blueberry. *Eur J Nutr*. 2021;60(3):1521-1536. doi:10.1007/s00394-020-02336-8

107. Dangles O, Fenger J-A. The Chemical Reactivity of Anthocyanins and Its Consequences in Food Science and Nutrition. *Molecules*. 2018;23(8):1970. doi:10.3390/molecules23081970

108. Cassidy A, O'Reilly ÉJ, Kay C, et al. Habitual intake of flavonoid subclasses and incident hypertension in adults. *Am J Clin Nutr*. 2011;93(2):338-347. doi:10.3945/ajcn.110.006783

109. Wu X, Beecher GR, Holden JM, Haytowitz DB, Gebhardt SE, Prior RL. Concentrations of anthocyanins in common foods in the United States and estimation of normal consumption. *J Agric Food Chem*. 2006;54(11):4069-4075. doi:10.1021/jf060300l

110. Patras A, Brunton NP, O'Donnell C, Tiwari BK. Effect of thermal processing on anthocyanin stability in foods; mechanisms and kinetics of degradation. *Trends Food Sci Technol*. 2010;21(1):3-11. doi:10.1016/j.tifs.2009.07.004

111. Chun OK, Chung SJ, Song WO. Estimated Dietary Flavonoid Intake and Major Food Sources of U.S. Adults. *J Nutr*. 2007;137(5):1244-1252. doi:10.1093/jn/137.5.1244

112. Zamora-Ros R, Knaze V, Luján-Barroso L, et al. Differences in dietary intakes, food sources and determinants of total flavonoids between Mediterranean and non-Mediterranean countries participating in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. *Br J Nutr*. 2013;109(8):1498-1507. doi:10.1017/S0007114512003273

113. Vogiatzoglou A, Mulligan AA, Lentjes MAH, et al. Flavonoid Intake in European Adults (18 to 64 Years). Wang D, ed. *PLoS One*. 2015;10(5):e0128132. doi:10.1371/journal.pone.0128132

114. Sebastian RS, Wilkinson Enns C, Goldman JD, et al. A New Database Facilitates Characterization of Flavonoid Intake, Sources, and Positive Associations with Diet Quality among US Adults. *J Nutr*. 2015;145(6):1239-1248. doi:10.3945/jn.115.213025

115. Zamora-Ros R, Knaze V, Rothwell J a., et al. Dietary polyphenol intake in europe: The european prospective investigation into cancer and nutrition (EPIC) study. *Eur J Nutr*. 2016;55(4):1359-1375. doi:10.1007/s00394-015-0950-x

116. Igwe EO, Charlton KE, Probst YC. Usual dietary anthocyanin intake, sources and their association with blood pressure in a representative sample of Australian adults. *J Hum Nutr Diet*. 2019;32(5):578-590. doi:10.1111/jhn.12647

117. Li G, Zhu Y, Zhang Y, Lang J, Chen Y, Ling W. Estimated daily flavonoid and stilbene intake from fruits, vegetables, and nuts and associations with lipid profiles in Chinese adults. *J Acad Nutr Diet*. 2013;113(6):786-794. doi:10.1016/j.jand.2013.01.018

118. European Food Safety Authority (EFSA). Scientific Opinion on the re-evaluation of anthocyanins (E 163) as a food additive. *EFSA J*. 2013;11(4):3145. doi:10.2903/j.efsa.2013.3145

119. Pourrat H, Bastide P, Dorier P, Tronche P. Préparation et activité thérapeutique de quelques glycosides d'anthocyanes. [Preparation and therapeutic activity of some anthocyanin glycosides.]. *Chim Ther*. 1967;2(33):8.

120. Scalbert A, Williamson G. Dietary Intake and Bioavailability of Polyphenols. *J Nutr*. 2000;130(8):2073S-2085S. doi:10.1093/jn/130.8.2073s

121. Kühnau J. The flavonoids. A class of semi-essential food components: their role in human nutrition. *World Rev Nutr Diet*. 1976;24:117-191.

122. Nurmi T, Mursu J, Heinonen M, Nurmi A, Hiltunen R, Voutilainen S. Metabolism of berry anthocyanins to phenolic acidsin humans. *J Agric Food Chem*. 2009;57(6):2274-2281. doi:10.1021/jf8035116

123. De Ferrars RM, Czank C, Zhang Q, et al. The pharmacokinetics of anthocyanins and their metabolites in humans. *Br J Pharmacol*. 2014;171(13):3268-3282. doi:10.1111/bph.12676

124. Del Rio D, Rodriguez-Mateos A, Spencer JPE, Tognolini M, Borges G, Crozier A. Dietary (Poly)phenolics in Human Health: Structures, Bioavailability, and Evidence of

Protective Effects Against Chronic Diseases. *Antioxid Redox Signal*. 2013;18(14):1818-1892. doi:10.1089/ars.2012.4581

125. Passamonti S, Vrhovsek U, Vanzo A, Mattivi F. The stomach as a site for anthocyanins absorption from food. *FEBS Lett*. 2003;544(1-3):210-213. doi:10.1016/S0014-5793(03)00504-0

126. Zou T-B, Feng D, Song G, Li H-W, Tang H-W, Ling W-H. The role of sodium-dependent glucose transporter 1 and glucose transporter 2 in the absorption of cyanidin-3-o-β-glucoside in Caco-2 cells. *Nutrients*. 2014;6(10):4165-4177. doi:10.3390/nu6104165

127. Gee JM, DuPont MS, Day a J, Plumb GW, Williamson G, Johnson IT. Intestinal transport of quercetin glycosides in rats involves both deglycosylation and interaction with the hexose transport pathway. *J Nutr*. 2000;130(July):2765-2771.

128. Williamson G, Clifford MN. Colonic metabolites of berry polyphenols: the missing link to biological activity? *Br J Nutr*. 2010;104 Suppl(2010):S48-S66. doi:10.1017/S0007114510003946

129. Lampe JW, Chang J-L. Interindividual Differences in Phytochemical Metabolism and Disposition. *Semin Cancer Biol*. 2007;17(5):347-353. doi:10.1016/j.semcancer.2007.05.003

130. Kahle K, Kraus M, Scheppach W, Ackermann M, Ridder F, Richling E. Studies on apple and blueberry fruit constituents: Do the polyphenols reach the colon after ingestion? *Mol Nutr Food Res*. 2006;50(4-5):418-423. doi:10.1002/mnfr.200500211

131. Sender R, Fuchs S, Milo R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLOS Biol*. 2016;14(8):e1002533. doi:10.1371/journal.pbio.1002533

132. Rodriguez-Mateos A, Vauzour D, Krueger CG, et al. Bioavailability, bioactivity and impact on health of dietary flavonoids and related compounds: an update. *Arch Toxicol*. 2014;88(10):1803-1853. doi:10.1007/s00204-014-1330-7

133. Melo-Filho C, Braga R, Andrade C. Advances in Methods for Predicting Phase I Metabolism of Polyphenols. *Curr Drug Metab*. 2014;15(1):120-126. doi:10.2174/1389200215666140130125339

134. Chu LL, Pandey RP, Jung N, Jung HJ, Kim EH, Sohng JK. Hydroxylation of diverse flavonoids by CYP450 BM3 variants: Biosynthesis of eriodictyol from naringenin in whole cells and its biological activities. *Microb Cell Fact*. 2016;15(1):1-15. doi:10.1186/s12934-016-0533-4

135. Bachmann K. Chapter 8 - Drug Metabolism. In: Hacker M, Messer W, Bachmann K, eds. *Pharmacology*. Academic Press; 2009:131-173. doi:https://doi.org/10.1016/B978-0-12-369521-5.00008-7

136. Doherty MM, Charman WN. The mucosa of the small intestine: how clinically relevant as an organ of drug metabolism? *Clin Pharmacokinet*. 2002;41(4):235-253. doi:10.2165/00003088-200241040-00001

137. McGhie TK, Walton MC. The bioavailability and absorption of anthocyanins: Towards a better understanding. *Mol Nutr Food Res*. 2007;51(6):702-713. doi:10.1002/mnfr.200700092

138. Crozier A, Jaganath IB, Clifford MN. Dietary phenolics: Chemistry, bioavailability and effects on health. *Nat Prod Rep*. 2009;26(8):1001-1043. doi:10.1039/b802662a

139. Pojer E, Mattivi F, Johnson D, Stockley CS. The case for anthocyanin consumption to promote human health: A review. *Compr Rev Food Sci Food Saf*. 2013;12(5):483-508.

doi:10.1111/1541-4337.12024

140. Czank C, Cassidy A, Zhang Q, et al. Human metabolism and elimination of the anthocyanin , cyanidin-3-glucoside : a 13 C-tracer study 1 – 3. *Am J Clin Nutr*. 2013;97:995-1003. doi:10.3945/ajcn.112.049247.

141. Amin HP, Czank C, Raheem S, et al. Anthocyanins and their physiologically relevant metabolites alter the expression of IL-6 and VCAM-1 in CD40L and oxidized LDL challenged vascular endothelial cells. *Mol Nutr Food Res*. 2015;59(6):1095-1106. doi:10.1002/mnfr.201400803

142. di Gesso JL, Kerr JS, Zhang Q, et al. Flavonoid metabolites reduce tumor necrosis factor-α secretion to a greater extent than their precursor compounds in human THP-1 monocytes. *Mol Nutr Food Res*. 2015;59(6):1143-1154. doi:10.1002/mnfr.201400799

143. Bharat D, Cavalcanti RRM, Petersen C, et al. Blueberry Metabolites Attenuate Lipotoxicity-Induced Endothelial Dysfunction. *Mol Nutr Food Res*. 2018;62(2):1-8. doi:10.1002/mnfr.201700601

144. Yi W, Akoh CC, Fischer J, Krewer G. Absorption of anthocyanins from blueberry extracts by Caco-2 human intestinal cell monolayers. *J Agric Food Chem*. 2006;54(15):5651-5658. doi:10.1021/jf0531959

145. Talavéra S, Felgines C, Texier O, et al. Anthocyanins Are Efficiently Absorbed from the Small Intestine in Rats. *J Nutr*. 2004;134(9):2275-2279. doi:10.1093/jn/134.9.2275

146. Prior RL, Wu X. Anthocyanins: Structural characteristics that result in unique metabolic patterns and biological activities. *Free Radic Res*. 2006;40(10):1014-1028. doi:10.1080/10715760600758522

147. Sandhu AK, Huang Y, Xiao D, Park E, Edirisinghe I, Burton-Freeman B. Pharmacokinetic Characterization and Bioavailability of Strawberry Anthocyanins Relative to Meal Intake. *J Agric Food Chem*. 2016;64(24):4891-4899. doi:10.1021/acs.jafc.6b00805

148. Mullen W, Edwards CA, Serafini M, Crozier A. Bioavailability of pelargonidin-3-O-glucoside and its metabolites in humans following the ingestion of strawberries with and without cream. *J Agric Food Chem*. 2008;56(3):713-719. doi:10.1021/jf072000p

149. Curtis PJ, Sampson M, Potter J, Dhatariya K, Kroon P a., Cassidy A. Chronic Ingestion of Flavan-3-ols and Isoflavones Improves Insulin Sensitivity and Lipoprotein Status and Attenuates Estimated 10-Year CVD Risk in Medicated Postmenopausal Women With Type 2 Diabetes: A 1-year, double-blind, randomized, controlled trial. *Diabetes Care*. 2012;35(2):226-232. doi:10.2337/dc11-1443

150. Cassidy A, Minihane A-M. The role of metabolism (and the microbiome) in defining the clinical efficacy of dietary flavonoids. *Am J Clin Nutr*. 2017;105(1):10-22. doi:10.3945/ajcn.116.136051

151. Tomás-Barberán FA, Selma M V, Espín JC. Interactions of gut microbiota with dietary polyphenols and consequences to human health. *Curr Opin Clin Nutr Metab Care*. 2016;19(6):471-476. doi:10.1097/MCO.0000000000000314

152. Setchell KDR, Brown NM, Lydeking-Olsen E. The clinical importance of the metabolite equol-a clue to the effectiveness of soy and its isoflavones. *J Nutr*. 2002;132(12):3577-3584.

153. Joy S, Siow RCM, Rowlands DJ, et al. The isoflavone equol mediates rapid vascular relaxation: Ca 2+-independent activation of endothelial nitric-oxide synthase/Hsp90 involving ERK1/2 and Akt phosphorylation in human endothelial cells. *J Biol Chem*. 2006;281(37):27335-27345. doi:10.1074/jbc.M602803200

154. Gredel S, Grad C, Rechkemmer G, Watzl B. Phytoestrogens and phytoestrogen metabolites differentially modulate immune parameters in human leukocytes. *Food Chem Toxicol*. 2008;46(12):3691-3696. doi:10.1016/j.fct.2008.09.047

155. Tomás-Barberán FA, González-Sarrías A, García-Villalba R, et al. Urolithins, the rescue of “old” metabolites to understand a “new” concept: Metabotypes as a nexus among phenolic metabolism, microbiota dysbiosis, and host health status. *Mol Nutr Food Res*. 2017;61(1). doi:10.1002/mnfr.201500901

156. González-Sarrías A, García-Villalba R, Romo-Vaquero M, et al. Clustering according to urolithin metabotype explains the interindividual variability in the improvement of cardiovascular risk biomarkers in overweight-obese individuals consuming pomegranate: A randomized clinical trial. *Mol Nutr Food Res*. 2017;61(5):1-14. doi:10.1002/mnfr.201600830

157. Williamson G, Clifford MN. Role of the small intestine, colon and microbiota in determining the metabolic fate of polyphenols. *Biochem Pharmacol*. 2017;139:24-39. doi:10.1016/j.bcp.2017.03.012

158. Mena P, Ludwig IA, Tomatis VB, et al. Inter-individual variability in the production of flavan-3-ol colonic metabolites: preliminary elucidation of urinary metabotypes. *Eur J Nutr*. 2019;58(4):1529-1543. doi:10.1007/s00394-018-1683-4

159. Curtis PJ, Van Der Velpen V, Berends L, et al. Blueberries improve biomarkers of cardiometabolic function in participants with metabolic syndrome-results from a 6-month, double-blind, randomized controlled trial. *Am J Clin Nutr*. 2019;109(6):1535-1545. doi:10.1093/ajcn/nqy380

160. Xin W, Wei W, Li X. Effect of Fish Oil Supplementation on Fasting Vascular Endothelial Function in Humans: A Meta-Analysis of Randomized Controlled Trials. *PLoS One*. 2012;7(9):e46028. https://doi.org/10.1371/journal.pone.0046028

161. Green DJ, Smith KJ. Effects of Exercise on Vascular Function, Structure, and Health in Humans. *Cold Spring Harb Perspect Med*. 2018;8(4). doi:10.1101/cshperspect.a029819

162. Carlström M, Liu M, Yang T, et al. Cross-talk between Nitrate-Nitrite-NO and NO Synthase Pathways in Control of Vascular NO Homeostasis. *Antioxidants Redox Signal*. 2015;23(4):295-306. doi:10.1089/ars.2013.5481

163. Turnbull D, Rodricks J V, Mariano GF, Chowdhury F. Caffeine and cardiovascular health. *Regul Toxicol Pharmacol*. 2017;89:165-185. doi:https://doi.org/10.1016/j.yrtph.2017.07.025

164. Chiva-Blanch G, Arranz S, Lamuela-Raventos RM, Estruch R. Effects of wine, alcohol and polyphenols on cardiovascular disease risk factors: evidences from human studies. *Alcohol Alcohol*. 2013;48(3):270-277. doi:10.1093/alcalc/agt007

165. McKeown NM, Day NE, Welch AA, et al. Use of biological markers to validate self-reported dietary intake in a random sample of the European Prospective Investigation into Cancer United Kingdom Norfolk cohort. *Am J Clin Nutr*. 2001;74(2):188-196. doi:10.1093/ajcn/74.2.188

166. Craig CL, Marshall AL, Sjöström M, et al. International physical activity questionnaire: 12-Country reliability and validity. *Med Sci Sports Exerc*. 2003;35(8):1381-1395. doi:10.1249/01.MSS.0000078924.61453.FB

167. Bonetti PO, Lerman LO, Lerman A. Endothelial dysfunction: A marker of atherosclerotic risk. *Arterioscler Thromb Vasc Biol*. 2003;23(2):168-175. doi:10.1161/01.ATV.0000051384.43104.FC

168. Celermajer DS, Sorensen KE, Gooch VM, et al. Non-invasive detection of endothelial dysfunction in children and adults at risk of atherosclerosis. *Lancet*. 1992;340(8828):1111-1115. doi:10.1016/0140-6736(92)93147-F

169. Lewington S, Clarke R, Qizilbash N, Peto R, Collins R. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet (London, England)*. 2002;360(9349):1903-1913. doi:10.1016/s0140-6736(02)11911-8

170. Fuchs FD, Whelton PK. High Blood Pressure and Cardiovascular Disease. *Hypertension*. 2020;(Cvd):285-292. doi:10.1161/HYPERTENSIONAHA.119.14240

171. Zieman SJ, Melenovsky V, Kass DA. Mechanisms, pathophysiology, and therapy of arterial stiffness. *Arterioscler Thromb Vasc Biol*. 2005;25(5):932-943. doi:10.1161/01.ATV.0000160548.78317.29

172. Safar ME, Asmar R, Benetos A, et al. Interaction between hypertension and arterial stiffness an expert reappraisal. *Hypertension*. 2018;72(4):796-805. doi:10.1161/HYPERTENSIONAHA.118.11212

173. Mitchell GF, Hwang SJ, Vasan RS, et al. Arterial stiffness and cardiovascular events: The framingham heart study. *Circulation*. 2010;121(4):505-511. doi:10.1161/CIRCULATIONAHA.109.886655

174. Townsend RR, Wilkinson IB, Schiffrin EL, et al. Recommendations for Improving and Standardizing Vascular Research on Arterial Stiffness: A Scientific Statement from the American Heart Association. *Hypertension*. 2015;66(3):698-722. doi:10.1161/HYP.0000000000000033

175. Townsend RR, Black HR, Chirinos JA, et al. Clinical Use of Pulse Wave Analysis: Proceedings From a Symposium Sponsored by North American Artery. *J Clin Hypertens*. 2015;17(7):503-513. doi:10.1111/jch.12574

176. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem*. 1972;18(6):499-502.

177. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and β-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia*. 1985;28(7):412-419. doi:10.1007/BF00280883

178. Magnusson I, Rothman DL, Katz LD, Shulman RG, Shulman GI. Increased rate of gluconeogenesis in type II diabetes mellitus. A 13C nuclear magnetic resonance study. *J Clin Invest*. 1992;90(4):1323-1327. doi:10.1172/JCI115997

179. Rassaf T, Bryan NS, Kelm M, Feelisch M. Concomitant presence of N-nitroso and S-nitroso proteins in human plasma. *Free Radic Biol Med*. 2002;33(11):1590-1596. doi:10.1016/S0891-5849(02)01183-8

180. Csonka C, Páli T, Bencsik P, Görbe A, Ferdinandy P, Csont T. Measurement of NO in biological samples. *Br J Pharmacol*. 2015;172(6):1620-1632. doi:10.1111/bph.12832

181. Lundberg JO, Weitzberg E, Gladwin MT. The nitrate-nitrite-nitric oxide pathway in physiology and therapeutics. *Nat Rev Drug Discov*. 2008;7(2):156-167. doi:10.1038/nrd2466

182. Straczkowski M, Dzienis-Straczkowska S, Stêpieñ A, Kowalska I, Szelachowska M, Kinalska I. Plasma interleukin-8 concentrations are increased in obese subjects and related to fat mass and tumor necrosis factor-α system. *J Clin Endocrinol Metab*. 2002;87(10):4602-4606. doi:10.1210/jc.2002-020135

183. De Ferrars RM, Czank C, Saha S, et al. Methods for isolating, identifying, and quantifying anthocyanin metabolites in clinical samples. *Anal Chem*. 2014;86(20):10052-10058. doi:10.1021/ac500565a

184. Verardi V, Croux C. Robust regression in Stata. *Stata J*. 2009;9(3):439-453. doi:10.1177/1536867x0900900306

185. Yellowlees A, Bursa F, Fleetwood KJ, et al. The Appropriateness of Robust Regression in Addressing Outliers in an Anthrax Vaccine Potency Test. *Bioscience*. 2015;66(1):63-72. doi:10.1093/biosci/biv159

186. Worley B, Powers R. Multivariate Analysis in Metabolomics. *Curr Metabolomics*. 2013;1(1):92-107. doi:10.2174/2213235x11301010092

187. Hong S, Song Y, Lee KH, et al. A fruit and dairy dietary pattern is associated with a reduced risk of metabolic syndrome. *Metabolism*. 2012;61(6):883-890. doi:10.1016/j.metabol.2011.10.018

188. Guasch-Ferre M, Bhupathiraju SN, Hu FB. Use of metabolomics in improving assessment of dietary intake. *Clin Chem*. 2018;64(1):82-98. doi:10.1373/clinchem.2017.272344

189. Denova-Gutiérrez E, Castañón S, Talavera JO, et al. Dietary patterns are associated with metabolic syndrome in an urban Mexican population. *J Nutr*. 2010;140(10):1855-1863. doi:10.3945/jn.110.122671

190. Windig W, Kistemaker PG, Haverkamp J, Meuzelaar HLC. Factor analysis of the influence of changes in experimental conditions in pyrolysis-mass spectrometry. *J Anal Appl Pyrolysis*. 1980;2(1):7-18. doi:10.1016/0165-2370(80)80041-6

191. van den Berg RA, Hoefsloot HCJ, Westerhuis JA, Smilde AK, van der Werf MJ. Centering, scaling, and transformations: Improving the biological information content of metabolomics data. *BMC Genomics*. 2006;7:1-15. doi:10.1186/1471-2164-7-142

192. Osborne JW, Costello AB, Kellow JT. Best Practices in Exploratory Factor Analysis. In: *Best Practices in Quantitative Methods*. SAGE Publications, Inc.; 2014:86-99. doi:10.4135/9781412995627.d8

193. Hair JF (Kennesaw SU, Black B, Babin BJ, Anderson RE (Drexel U. Chapter 3: Factor analysis. In: *Mutlivariate Data Analysis*. 7th Editio. Pearson Education; 2010:91-152.

194. MacCallum RC, Widaman KF, Zhang S, Hong S. Sample size in factor analysis. *Psychol Methods*. 1999;4(1):84-99. doi:10.1037/1082-989X.4.1.84

195. Hayton JC, Allen DG, Scarpello V. Factor Retention Decisions in Exploratory Factor Analysis: A Tutorial on Parallel Analysis. *Organ Res Methods*. 2004;7(2):191-205. doi:10.1177/1094428104263675

196. Velicer WF, Eaton CA, Fava JL. Construct Explication through Factor or Component Analysis: A Review and Evaluation of Alternative Procedures for Determining the Number of Factors or Components. In: Goffin RD, Helmes E, eds. Springer US; 2000:41-71. doi:10.1007/978-1-4615-4397-8_3

197. DiStefano C, Zhu M, Mîndrilă D. Understanding and using factor scores: Considerations for the applied researcher. *Pract Assessment, Res Eval*. 2009;14(20).

198. Barona J, Aristizabal JC, Blesso CN, Volek JS, Fernandez ML. Grape polyphenols reduce blood pressure and increase flow-mediated vasodilation in men with metabolic syndrome. *J Nutr*. 2012;142(9):1626-1632. doi:10.3945/jn.112.162743

199. Rothwell JA, Perez-Jimenez J, Neveu V, et al. Phenol-Explorer 3.0: A major update of the Phenol-Explorer database to incorporate data on the effects of food processing on

polyphenol content. *Database*. 2013;2013:1-8. doi:10.1093/database/bat070

200. Wedick NM, Pan A, Cassidy A, et al. Dietary flavonoid intakes and risk of type 2 diabetes in US men and women. *Am J Clin Nutr*. 2012;95(4):925-933. doi:10.3945/ajcn.111.028894

201. Cassidy A, Rogers G, Peterson JJ, Dwyer JT, Lin H, Jacques PF. Higher dietary anthocyanin and flavonol intakes are associated with anti-inflammatory effects in a population of US adults. *Am J Clin Nutr*. 2015;102(1):172-181. doi:10.3945/ajcn.115.108555

202. Bento-Silva A, Koistinen VM, Mena P, et al. Factors affecting intake, metabolism and health benefits of phenolic acids: do we understand individual variability? *Eur J Nutr*. 2020;59(4):1275-1293. doi:10.1007/s00394-019-01987-6

203. González-Sarrías A, Combet E, Pinto P, et al. A systematic review and meta-analysis of the effects of flavanol-containing tea, cocoa and apple products on body composition and blood lipids: Exploring the factors responsible for variability in their efficacy. *Nutrients*. 2017;9(7). doi:10.3390/nu9070746

204. Manach C, Milenkovic D, Van de Wiele T, et al. Addressing the inter-individual variation in response to consumption of plant food bioactives: Towards a better understanding of their role in healthy aging and cardiometabolic risk reduction. *Mol Nutr Food Res*. 2017;61(6):1600557. doi:10.1002/mnfr.201600557

205. Frankenfeld CL. Cardiometabolic risk and gut microbial phytoestrogen metabolite phenotypes. *Mol Nutr Food Res*. 2017;61(1):1-15. doi:10.1002/mnfr.201500900

206. Martini D, Chiavaroli L, González-Sarrías A, et al. Impact of foods and dietary supplements containing hydroxycinnamic acids on cardiometabolic biomarkers: A systematic review to explore inter-individual variability. *Nutrients*. 2019;11(8). doi:10.3390/nu11081805

207. García-Conesa MT, Chambers K, Combet E, et al. *Meta-Analysis of the Effects of Foods and Derived Products Containing Ellagitannins and Anthocyanins on Cardiometabolic Biomarkers: Analysis of Factors Influencing Variability of the Individual Responses*. Vol 19.; 2018. doi:10.3390/ijms19030694

208. Morand C, Roos B De, Garcia-conesa MT, et al. Why interindividual variation in response to consumption of plant food bioactives matters for future personalised nutrition. *Proc Nutr Soc*. 2020;(April 2019):1-11. doi:10.1017/S0029665120000014

209. Tomás-Navarro M, Vallejo F, Sentandreu E, Navarro JL, Tomás-Barberán FA. Volunteer stratification is more relevant than technological treatment in orange juice flavanone bioavailability. *J Agric Food Chem*. 2014;62(1):24-27. doi:10.1021/jf4048989

210. Hazim S, Curtis PJ, Schär MY, et al. Acute benefits of the microbial-derived isoflavone metabolite equol on arterial stiffness in men prospectively recruited according to equol producer phenotype: A double-blind randomized controlled trial. *Am J Clin Nutr*. 2016;103(3):694-702. doi:10.3945/ajcn.115.125690

211. Van Der Velpen V, Geelen A, Hollman PCH, Schouten EG, Van 't Veer P, Afman LA. Isoflavone supplement composition and equol producer status affect gene expression in adipose tissue: A double-blind, randomized, placebo-controlled crossover trial in postmenopausal women. *Am J Clin Nutr*. 2014;100(5):1269-1277. doi:10.3945/ajcn.114.088484

212. Comrey AL, Lee HB. *A First Course in Factor Analysis*. Psychology Press; 1992.

213. Gorsuch RL. Exploratory factor analysis: Its role in item analysis. *J Pers Assess*.

1997;68(3):532-560. doi:10.1207/s15327752jpa6803_5

214. Dziura JD, Post LA, Zhao Q, Fu Z, Peduzzi P. Strategies for dealing with missing data in clinical trials: From design to analysis. *Yale J Biol Med*. 2013;86(3):343-358.

215. Velicer WF. Determining the number of components from the matrix of partial correlations. *Psychometrika*. 1976;41(3):321-327. doi:10.1007/BF02293557

216. Cattell RB. The Scree Test For The Number Of Factors. *Multivariate Behav Res*. 1966;1(2):245-276. doi:10.1207/s15327906mbr0102_10

217. Costello AB, Osborne JW. Best Practices in Exploratory Factor Analysis : Four Recommendations for Getting the Most From Your Analysis. *Pract Assessment, Res Educ*. 2005;10:1-9. doi:10.1.1.110.9154

218. Zhong S, Sandhu A, Edirisinghe I, Burton-Freeman B. Characterization of Wild Blueberry Polyphenols Bioavailability and Kinetic Profile in Plasma over 24-hour Period in Human Subjects. *Mol Nutr Food Res*. Published online 2017. doi:10.1002/mnfr.201700405

219. Ludwig IA, Mena P, Calani L, et al. New insights into the bioavailability of red raspberry anthocyanins and ellagitannins. *Free Radic Biol Med*. 2015;89:758-769. doi:10.1016/j.freeradbiomed.2015.10.400

220. Andreasen MF, Kroon PA, Williamson G. Hydroxycinnamates Is Distributed along the Intestine of Mammals. Published online 2001:5679-5684.

221. Kay CD, Pereira-Caro G, Ludwig IA, Clifford MN, Crozier A. Anthocyanins and Flavanones Are More Bioavailable than Previously Perceived: A Review of Recent Evidence. *Annu Rev Food Sci Technol*. 2017;8:155-180. doi:10.1146/annurev-food-030216-025636

222. Pero R. Health Consequences of Catabolic Synthesis of Hippuric Acid in Humans. *Curr Clin Pharmacol*. 2010;5(1):67-73. doi:10.2174/157488410790410588

223. Lambert GW, Eisenhofer G, Jennings GL, Esler MD. Regional Homovanillic acid production in humans. *Life Sci*. 1993;53:63-75.

224. Akira Y, Masanori Y, Tatsuhiro Y, Toshio S, Zenzo T. Urinary levels of vanilmandelic acid and homovanillic acid determined by high-speed liquid chromatography. *Clin Chim Acta*. 1976;73(2):315-320. doi:10.1016/0009-8981(76)90178-9

225. van Dorsten FA, Grün CH, van Velzen EJJ, Jacobs DM, Draijer R, van Duynhoven JPM. The metabolic fate of red wine and grape juice polyphenols in humans assessed by metabolomics. *Mol Nutr Food Res*. 2010;54(7):897-908. doi:10.1002/mnfr.200900212

226. Rodriguez-Mateos A, del Pino-García R, George TW, Vidal-Diez A, Heiss C, Spencer JPE. Impact of processing on the bioavailability and vascular effects of blueberry (poly)phenols. *Mol Nutr Food Res*. 2014;58(10):1952-1961. doi:10.1002/mnfr.201400231

227. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16(3):297-334. doi:10.1007/BF02310555

228. Taber KS. The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Res Sci Educ*. 2018;48(6):1273-1296. doi:10.1007/s11165-016-9602-2

229. Guerra A, Folesani G, Mena P, et al. Hippuric acid in 24 h urine collections as a biomarker of fruits and vegetables intake in kidney stone formers. *Int J Food Sci Nutr*. 2014;65(8):1033-1038. doi:10.3109/09637486.2014.950210

230. Ras RT, Streppel MT, Draijer R, Zock PL. Flow-mediated dilation and cardiovascular risk prediction: A systematic review with meta-analysis. *Int J Cardiol*. 2013;168(1):344-351. doi:10.1016/j.ijcard.2012.09.047

231. Makles A. Stata tip 110: How to get the optimal k-means cluster solution. *Stata J*. 2012;12(2):347-351. https://www.stata-journal.com/article.html?article=st0262

232. Martinez-guryn K, Hubert N, Frazier K, et al. Small intestine microbiota regulate host digestive and absorptive adaptive responses to dietary lipids. *Cell Host Microbe*. 2018;23(4):458-469. doi:10.1016/j.chom.2018.03.011.Small

233. Bae JH, Bassenge E, Kim KB, et al. Postprandial hypertriglyceridemia impairs endothelial function by enhanced oxidant stress. *Atherosclerosis*. 2001;155(2):517-523. doi:10.1016/s0021-9150(00)00601-8

234. Kajikawa M, Maruhashi T, Matsumoto T, et al. Relationship between serum triglyceride levels and endothelial function in a large community-based study. *Atherosclerosis*. 2016;249:70-75. doi:10.1016/j.atherosclerosis.2016.03.035

235. Godos J, Sinatra D, Blanco I, Mulè S, La Verde M, Marranzano M. Association between dietary phenolic acids and hypertension in a mediterranean cohort. *Nutrients*. 2017;9(10). doi:10.3390/nu9101069

236. Zhu Y, Bo Y, Wang X, et al. The Effect of Anthocyanins on Blood Pressure. *Medicine (Baltimore)*. 2016;95(15). doi:10.1097/MD.0000000000003380

237. Yang LP, Ling WH, Du ZC, et al. Effects of anthocyanins on cardiometabolic health: A systematic review and meta-analysis of randomized controlled trials. *Adv Nutr*. 2017;8:684-693. doi:10.3945/an.116.014852

238. Zhang H, Park Y, Wu J, et al. Role of TNF-α in vascular dysfunction. *Clin Sci*. 2009;116(3):219-230. doi:10.1042/CS20080196

239. Rothwell JA, Madrid-Gambin F, Garcia-Aloy M, et al. Biomarkers of intake for coffee, tea, and sweetened beverages. *Genes Nutr*. 2018;13(1):1-18. doi:10.1186/s12263-018-0607-5

240. Penczynski KJ, Krupp D, Bring A, Bolzenius K, Remer T, Buyken AE. Relative validation of 24-h urinary hippuric acid excretion as a biomarker for dietary flavonoid intake from fruit and vegetables in healthy adolescents. *Eur J Nutr*. 2017;56(2):757-766. doi:10.1007/s00394-015-1121-9

241. Ross AB, Åman P, Kamal-Eldin A. Identification of cereal alkylresorcinol metabolites in human urine - Potential biomarkers of wholegrain wheat and rye intake. *J Chromatogr B Anal Technol Biomed Life Sci*. 2004;809(1):125-130. doi:10.1016/j.jchromb.2004.06.015

242. González-Barrio R, Edwards CA, Crozier A. Colonic Catabolism of Ellagitannins, Ellagic Acid, and Raspberry Anthocyanins: In Vivo and In Vitro Studies. *Drug Metab Dispos*. 2011;39:1680-1688. doi:10.1124/dmd.111.039651.

243. Zadernowski R, Naczk M, Nesterowicz J. Phenolic acid profiles in some small berries. *J Agric Food Chem*. 2005;53(6):2118-2124. doi:10.1021/jf040411p

244. Lauer T, Preik M, Rassaf T, et al. Plasma nitrite rather than nitrate reflects regional endothelial nitric oxide synthase activity but lacks intrinsic vasodilator action. *Proc Natl Acad Sci U S A*. 2001;98(22):12814-12819. doi:10.1073/pnas.221381098

245. de Ferrars RM, Cassidy A, Curtis P, Kay CD. Phenolic metabolites of anthocyanins following a dietary intervention study in post-menopausal women. *Mol Nutr Food Res*. 2014;58(3):490-502. doi:10.1002/mnfr.201300322

246. Brett GM, Hollands W, Needs PW, et al. Absorption, metabolism and excretion of flavanones from single portions of orange fruit and juice and effects of anthropometric variables and contraceptive pill use on flavanone excretion. *Br J Nutr*. 2009;101(5):664-675. doi:10.1017/S000711450803081X

247. Tomas-Barberan FA, Garcia-Villalba R, Gonzalez-Sarrias A, Selma M V., Espin JC. Ellagic acid metabolism by human gut microbiota: Consistent observation of three urolithin phenotypes in intervention trials, independent of food source, age, and health status. *J Agric Food Chem*. 2014;62(28):6535-6538. doi:10.1021/jf5024615

248. Lampe JW, Atkinson C, Hullar MAJ. Assessing exposure to lignans and their metabolites in humans. *J AOAC Int*. 2006;89(4):1174-1181.

249. Chandra P, Rathore AS, Kay KL, et al. Contribution of berry polyphenols to the human metabolome. *Molecules*. 2019;24(23). doi:10.3390/molecules24234220

250. González-Barrio R, Borges G, Mullen W, Crozier A. Bioavailability of anthocyanins and ellagitannins following consumption of raspberries by healthy humans and subjects with an ileostomy. *J Agric Food Chem*. 2010;58(7):3933-3939. doi:10.1021/jf100315d

251. Stalmach A, Edwards CA, Wightman JD, Crozier A. Gastrointestinal stability and bioavailability of (poly)phenolic compounds following ingestion of Concord grape juice by humans. *Mol Nutr Food Res*. 2012;56(3):497-509. doi:10.1002/mnfr.201100566

252. Suzuki A, Yamamoto N, Jokura H, et al. Chlorogenic acid attenuates hypertension and improves endothelial function in spontaneously hypertensive rats. *J Hypertens*. 2006;24(6). https://journals.lww.com/jhypertension/Fulltext/2006/06000/Chlorogenic_acid_attenuates_hypertension_and.15.aspx

253. Loader TB, Taylor CG, Zahradka P, Jones PJH. Chlorogenic acid from coffee beans: Evaluating the evidence for a blood pressure-regulating health claim. *Nutr Rev*. 2017;75(2):114-133. doi:10.1093/nutrit/nuw057

254. Mills CE, Flury A, Marmet C, et al. Mediation of coffee-induced improvements in human vascular function by chlorogenic acids and its metabolites: Two randomized, controlled, crossover intervention trials. *Clin Nutr*. 2017;36(6):1520-1529. doi:10.1016/j.clnu.2016.11.013

255. Ward NC, Hodgson JM, Woodman RJ, et al. Acute effects of chlorogenic acids on endothelial function and blood pressure in healthy men and women. *Food Funct*. 2016;7(5):2197-2203. doi:10.1039/c6fo00248j

256. Agudelo-Ochoa GM, Pulgarín-Zapata IC, Velásquez-Rodriguez CM, et al. Coffee consumption increases the antioxidant capacity of plasma and has no effect on the lipid profile or vascular function in healthy adults in a randomized controlled trial. *J Nutr*. 2016;146(3):524-531. doi:10.3945/jn.115.224774

257. Rechner AR, Kuhnle G, Bremner P, Hubbard GP, Moore KP, Rice-Evans CA. The metabolic fate of dietary polyphenols in humans. *Free Radic Biol Med*. 2002;33(2):220-235. doi:10.1016/S0891-5849(02)00877-8

258. Pereira-Caro G, Oliver CM, Weerakkody R, et al. Chronic administration of a microencapsulated probiotic enhances the bioavailability of orange juice flavanones in humans. *Free Radic Biol Med*. 2015;84:206-214. doi:10.1016/j.freeradbiomed.2015.03.010

259. Tilg H, Kaser A. Gut microbiome, obesity, and metabolic dysfunction. *J Clin Invest*. 2011;121(6):2126-2132. doi:10.1172/JCI58109

260. Boto-Ordóñez M, Urpi-Sarda M, Queipo-Ortuño MI, et al. High levels of Bifidobacteria

are associated with increased levels of anthocyanin microbial metabolites: A randomized clinical trial. *Food Funct*. 2014;5(8):1932-1938. doi:10.1039/c4fo00029c

261. Moreno-Indias I, Sánchez-Alcoholado L, Pérez-Martínez P, et al. Red wine polyphenols modulate fecal microbiota and reduce markers of the metabolic syndrome in obese patients. *Food Funct*. 2016;7(4):1775-1787. doi:10.1039/c5fo00886g

262. Sonnenburg JL, Bäckhed F. Diet-microbiota interactions as moderators of human metabolism. *Nature*. 2016;535(7610):56-64. doi:10.1038/nature18846

263. Rodriguez-Mateos A, Feliciano RP, Cifuentes-Gomez T, Spencer JPE. Bioavailability of wild blueberry (poly)phenols at different levels of intake. *J Berry Res*. 2016;6(2):137-148. doi:10.3233/JBR-160123

264. Rechner AR, Kuhnle G, Hu H, et al. The Metabolism of Dietary Polyphenols and the Relevance to Circulating Levels of Conjugated Metabolites. *Free Radic Res*. 2009;36(11):1229-1241. doi:10.1080/1071576021000016472

265. Selma M V., Espín JC, Tomás-Barberán FA. Interaction between phenolics and gut microbiota: Role in human health. *J Agric Food Chem*. 2009;57(15):6485-6501. doi:10.1021/jf902107d

266. Kuntz S, Kunz C, Herrmann J, et al. Anthocyanins from fruit juices improve the antioxidant status of healthy young female volunteers without affecting anti-inflammatory parameters: Results from the randomised, double-blind, placebo-controlled, cross-over ANTHONIA (ANTHOcyanins in Nutrition. *Br J Nutr*. 2014;112(6):925-936. doi:10.1017/S0007114514001482

267. Yi L, Chen CY, Jin X, et al. Differential suppression of intracellular reactive oxygen species-mediated signaling pathway in vascular endothelial cells by several subclasses of flavonoids. *Biochimie*. 2012;94(9):2035-2044. doi:10.1016/j.biochi.2012.05.027

268. Stanikunaite R, Khan SI, Trappe JM, Ross SA. Cyclooxygenase-2 inhibitory and antioxidant compounds from the truffle Elaphomyces granulatus. *Phyther Res*. 2009;23(4):575-578. doi:10.1002/ptr.2698

269. Hein TW, Singh U, Vasquez-Vivar J, Devaraj S, Kuo L, Jialal I. Human C-reactive protein induces endothelial dysfunction and uncoupling of eNOS in vivo. *Atherosclerosis*. 2009;206(1):61-68. doi:10.1016/j.atherosclerosis.2009.02.002

270. Kelley DS, Adkins Y, Reddy A, Woodhouse LR, Mackey BE, Erickson KL. Sweet bing cherries lower circulating concentrations of markers for chronic inflammatory diseases in healthy humans1-4. *J Nutr*. 2013;143(3):340-344. doi:10.3945/jn.112.171371

271. Queipo-Ortuño MI, Boto-Ordóñez M, Murri M, et al. Influence of red wine polyphenols and ethanol on the gut microbiota ecology and biochemical biomarkers. *Am J Clin Nutr*. 2012;95(6):1323-1334. doi:10.3945/ajcn.111.027847

272. Vugic L, Colson N, Nikbakht E, et al. Anthocyanin supplementation inhibits secretion of pro-inflammatory cytokines in overweight and obese individuals. *J Funct Foods*. 2020;64(October 2019):103596. doi:10.1016/j.jff.2019.103596

273. Kerimi A, Kraut NU, da Encarnacao JA, Williamson G. The gut microbiome drives inter- and intra-individual differences in metabolism of bioactive small molecules. *Sci Rep*. 2020;10(1):1-12. doi:10.1038/s41598-020-76558-5

274. Burke GL, Bertoni AG, Shea S, et al. The impact of obesity on cardiovascular disease risk factors and subclinical vascular disease: The multi-ethnic study of atherosclerosis. *Arch Intern Med*. 2008;168(9):928-935. doi:10.1001/archinte.168.9.928

275. Saude EJ, Sykes BD. Urine stability for metabolomic studies: Effects of preparation and storage. *Metabolomics*. 2007;3(1):19-27. doi:10.1007/s11306-006-0042-2

276. Thomas CE, Sexton W, Benson K, Sutphen R. Urine Collection and Processing for Protein Biomarker Discovery and Quantification. *Biomarkers*. 2011;19(4):953-959. doi:10.1158/1055-9965.EPI-10-0069.Urine

277. European Medicines Agency. Guideline on bioanalytical method validation. Published online 2011. doi:10.4155/bio.12.44

278. Wells G, Prest H, Russ IV CW. *Signal, Noise, and Detection Limits in Mass Spectrometry*.; 2011.

279. Welch A, MacGregor A, Jennings A, Fairweather-Tait S, Spector T, Cassidy A. Habitual flavonoid intakes are positively associated with bone mineral density in women. *J Bone Miner Res*. 2012;27(9):1872-1878. doi:10.1002/jbmr.1649

280. Conroy RM. What hypotheses do “nonparametric” two-group tests actually test? *Stata J*. 2012;12(2):182-190. doi:10.1177/1536867x1201200202

281. Gonthier MP, Rios LY, Verny MA, Rémésy C, Scalbert A. Novel liquid chromatography-electrospray ionization mass spectrometry method for the quantification in human urine of microbial aromatic acid metabolites derived from dietary polyphenols. *J Chromatogr B Anal Technol Biomed Life Sci*. 2003;789(2):247-255. doi:10.1016/S1570-0232(03)00073-4

282. Holmes E, Loo RL, Stamler J, et al. Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature*. 2008;453(7193):396-400. doi:10.1038/nature06882

283. Elliott P, Posma JM, Chan Q, et al. Urinary metabolic signatures of human adiposity. *Sci Transl Med*. 2015;7(285):1-17. doi:10.1126/scitranslmed.aaa5680

284. Pallister T, Jackson MA, Martin TC, et al. Hippurate as a metabolomic marker of gut microbiome diversity: Modulation by diet and relationship to metabolic syndrome. *Sci Rep*. 2017;7(1):1-9. doi:10.1038/s41598-017-13722-4

285. Mullen W, Edwards CA, Serafini M, Crozier A. Bioavailability of Pelargonidin-3- O -glucoside and Its Metabolites in Humans Following the Ingestion of Strawberries with and without Cream. *J Agric Food Chem*. 2008;56(3):713-719. doi:10.1021/jf072000p

286. Reddy VC, Vidya Sagar G V, Sreeramulu D, Venu L, Raghunath M. Addition of Milk Does Not Alter the Antioxidant Activity of Black Tea. *Ann Nutr Metab*. 2005;49(3):189-195. doi:10.1159/000087071

287. Van Het Hof KH, Kivits GAA, Weststrate JA, Tijburg LBM. Bioavailability of catechins from tea: The effect of milk. *Eur J Clin Nutr*. 1998;52(5):356-359. doi:10.1038/sj.ejcn.1600568

288. Renouf M, Marmet C, Guy P, et al. Nondairy creamer, but not milk, delays the appearance of coffee phenolic acid equivalents in human plasma. *J Nutr*. 2010;140(2):259-263. doi:10.3945/jn.109.113027

289. Zhang H, Jiang L, Guo H, et al. The inhibitory effect of milk on the absorption of dietary phenolic acids and the change in human plasma antioxidant capacity through a mechanism involving both milk proteins and fats. *Mol Nutr Food Res*. 2013;57(7):1228-1236. doi:10.1002/mnfr.201200803

290. de Roos B, Aura AM, Bronze M, et al. Targeting the delivery of dietary plant bioactives to those who would benefit most: from science to practical applications. *Eur J Nutr*. 2019;58(s2):65-73. doi:10.1007/s00394-019-02075-5

291. Trøseid M, Andersen GØ, Broch K, Hov JR. The gut microbiome in coronary artery disease and heart failure: Current knowledge and future directions. *EBioMedicine*. 2020;52:102649. doi:10.1016/j.ebiom.2020.102649

292. Setchell KDR, Cole SJ. Method of Defining Equol-Producer Status and Its Frequency among Vegetarians. *J Nutr*. 2006;136(8):2188-2193. http://jn.nutrition.org/cgi/content/long/136/8/2188

293. Schork NJ, Goetz LH. Single-Subject Studies in Translational Nutrition Research. *Annu Rev Nutr*. 2017;37(1):395-422. doi:10.1146/annurev-nutr-071816-064717

294. Ballermann BJ, Dardik A, Eng E, Liu A. Shear stress and the endothelium. *Kidney Int*. 1998;54(67):S100-S108. doi:10.1046/j.1523-1755.1998.06720.x

295. National Library of Medicine (US). PubMed.

296. Brat P, Mennen L, Georgé S, et al. Determination of the Polyphenol Content of Fruits and Vegetables. Establishment of a Database and Estimation of the Polyphenol Intake in the French Diet. *Acta Hortic*. 2007;(744):61-70. doi:10.17660/actahortic.2007.744.5

297. Neveu V, Perez-Jiménez J, Vos F, et al. Phenol-Explorer: an online comprehensive database on polyphenol contents in foods. *Database*. 2010;2010:doi: 10.1093/database/bap024. doi:10.1093/database/bap024

298. Kerio LC, Wachira FN, Wanyoko JK, Rotich MK. Characterization of anthocyanins in Kenyan teas: Extraction and identification. *Food Chem*. 2012;131(1):31-38. doi:https://doi.org/10.1016/j.foodchem.2011.08.005

299. He F, Liang NN, Mu L, et al. Anthocyanins and their variation in red wines I. Monomeric anthocyanins and their color expression. *Molecules*. 2012;17(2):1571-1601. doi:10.3390/molecules17021571

300. Mazza G, Velioglu YS. Anthocyanins and other phenolic compounds in fruits of red-flesh apples. *Food Chem*. 1992;43(2):113-117. doi:https://doi.org/10.1016/0308-8146(92)90223-O