Metabolomic Profiles and Prostate Cancer Risk

Thesis submitted for the degree of

Doctor of Philosophy

Hilary Term 2025

Zoe Grenville

Green Templeton College

University of Oxford

Supervised by

Professor Ruth Travis, Dr. Karl Smith-Byrne, Dr. Julie Schmidt

Cancer Epidemiology Unit

Oxford Population Health

Abstract

Metabolomic Profiles and Prostate Cancer Risk

By Zoe Grenville, University of Oxford, DPhil thesis, Hilary Term 2025

Pre-diagnostic blood metabolites may help identify novel risk factors for prostate cancer incidence, and these metabolite profiles may be modifiable through lifestyle and dietary factors. For this thesis, I investigated circulating metabolites in relation to (a) potential determinants and (b) prostate cancer risk, using EPIC and the UK Biobank for two of the largest prospective studies to date of metabolites and prostate cancer risk.

In EPIC, plasma metabolites were measured by liquid chromatography-mass spectrometry. In 3,042 men from EPIC free of cancer at blood draw, in a two-stage cross-sectional discovery (2,524 men) and validation (518 men) study, I estimated the associations of 24 dietary and lifestyle variables with three metabolite patterns (and levels of the individual metabolites loading on each pattern). Intakes of alcohol and fish were positively associated with Pattern 1 (64 phosphatidylcholines and two hydroxysphingomyelins). BMI was positively associated with Pattern 2 (two acylcarnitines, glutamate, ornithine, and taurine), which appeared to be driven by a strong positive BMI-glutamate association. Finally, both BMI and fatty fish were inversely associated with Pattern 3 (8 lyso phosphatidylcholines). These results indicated robust associations of fish, alcohol, and BMI with several individual metabolites, as well as three metabolite patterns previously found to be inversely associated with risk of aggressive prostate cancer.

In an extended EPIC nested case-control study of 4,387 matched case-control pairs, I investigated the associations of 148 individual metabolites and the three previously defined metabolite patterns with prostate cancer risk, overall and by time since blood collection, tumour characteristics, and risk of prostate cancer death. After correction for multiple testing, six phosphatidylcholines were inversely associated with risk of advanced prostate cancer within ten years of blood collection (this association was weaker with longer follow-up time). Patterns 1 and 2 were also inversely associated with advanced prostate cancer; these associations were strongest for men diagnosed within 10 years of recruitment (OR_{1SD} 0.80, 95% CI 0.66–0.96 and 0.76, 0.59–0.97, respectively). Pattern 3 was associated with prostate cancer death (0.82, 0.68–0.98). These results suggest that plasma metabolite profiles change up to a decade before detection of advanced-stage disease.

Finally, in a prospective cohort analysis in UK Biobank (n=126,206 men), I assessed 249 metabolites measured using nuclear magnetic resonance imaging in relation to risk of overall prostate cancer (n=6,575 cases) and death from prostate cancer (n=596). Men with a higher percentage of docosahexaenoic acid (DHA) to total fatty acids had a higher risk of prostate cancer (HR per 1 SD 1.05, 95%CI 1.03-1.08). In addition, a higher percentage of phospholipids to total lipids in small LDL was associated with higher risk of prostate cancer (1.06, 1.02-1.09), with a greater magnitude of association in the first 3 years since blood collection (1.12, 1.04–1.20). Men with higher circulating citrate concentrations had an increased risk of prostate cancer death (1.18, 1.08-1.28). Finally, lower concentrations of glycoprotein acetyls (1.25, 1.14-1.4) and higher concentrations of albumin (0.86, 0.79-0.93) were associated with lower risk of prostate cancer death, with a greater magnitude of association closer to blood collection, likely indicating changes in blood metabolite profiles in response to an existing tumour. In cross-sectional analyses, higher fish intake was strongly associated with higher plasma DHA concentrations.

In conclusion, alcohol, fish intake, and BMI affect men's blood metabolic profiles. Several metabolites and metabolite patterns were inversely associated with advanced stage prostate cancer and death, possibly reflecting that there are changes in men's blood metabolic profiles in response to cancer up to ten years before diagnosis, while DHA was associated with higher overall prostate cancer risk. Ultimately, I highlight modifiable correlates of the metabolic landscape, offering further context into pathways that may be associated with preclinical prostate cancer. Furthermore, the findings of a possible aetiological role of DHA in prostate cancer risk contributes to previous observational research indicating an association between DHA (and other long-chain omega-3 polyunsaturated acids) and prostate cancer risk. Future work should involve pooled analyses of several cohorts (with longer follow-up time to investigate the temporality of these associations), investigating potential mechanisms including lipid metabolism and inflammation, and assessing causality using approaches such as Mendelian randomisation.

Acknowledgements

First and foremost, I would like to express my gratitude for my DPhil supervisors, Professor Ruth Travis, Dr. Karl Smith-Byrne, and Dr. Julie Schmidt, as well as Professor Tim Key. While there are no words that can fully convey how grateful I am for their supervision these past four years, I would like to thank them for their unwavering support, mentorship, and guidance throughout the course of my DPhil. As a result of my supervisors' encouragement and motivation, I had the opportunity to present at several conferences, publish my first primary author paper, and engage with a global network of research collaborators. Furthermore, their insightful advice and overall warmth fostered a positive, rewarding environment in which I could grow both academically and personally, and I sincerely appreciate their belief in me.

I also wish to show my appreciation for the lifelong friends I have made in Oxford throughout the course of the DPhil, both in Oxford Population Health and across the university. Whether it was helping with a last-minute coding question, offering each other a listening ear, having a yoga buddy, or sharing in the unique experience that is life at Oxford University, your support and friendship truly made Oxford feel like home. Though I cannot fit everyone on one page, I would especially like to thank Sofia Mouchti, Dixa Thakrar, Dorottya Zsiboracs, Asmita Poddar, Gabrielle Smith, Larissa Mark, Richin Kabra, Nick Shafer, Imogen Malpas, Andrew Rosenblatt, Charlotte Clarke, Maira Khan, Tonima Trisa, Izabella Lawson, Kiran Manku, and Ellena Damini. I owe a special thanks to Julia Reinhardt, my oldest friend in Oxford, for keeping me company throughout every day of thesis writing, and for always bringing positivity and laughter in the midst of our Branca work sessions.

Endless gratitude goes to my entire family for their patience, support, and belief in me, even during the most challenging times. I wish to extend my deepest gratitude to my mom, who I can always depend on for unconditional love and wisdom, and whom I can always count on to answer my calls at all hours, all the way in San Francisco. I would also like to acknowledge my dad, who always knows how to help me see the positive side of any situation, gives the best advice, and helped read over my thesis. An honourable mention goes out to my sister, Nadia, who provided invaluable nutrition input throughout the course of this thesis. I wish to also thank my brothers Andre and Dean, and well as my grandparents, for their constant support.

I feel extremely fortunate to have had incredible mentors before starting Oxford, whose support and guidance contributed to me ultimately pursuing my DPhil. In particular, I would like to thank Frances Bright, Dr. Jason Patent, Dr. Clive Pullinger, and Dr. John Kane for inspiring me to further pursue my education and believe in myself.

Finally, I would like to thank Cancer Research UK for supporting my DPhil research, as well as the research colleagues involved with the EPIC and UK Biobank studies, and the EPIC and UK Biobank participants, for helping make this thesis possible.

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Full Term
AA	Amino acid
ATBC	Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study
BCAA	Branched-chain amino acids (Leucine, Isoleucine, Valine)
Biocrates	Biocrates AbsoluteIDQ™ p180 kit (targeted metabolomics platform)
BMI	Body mass index
CI	Confidence Interval
CPS-II	Cancer Prevention Study-II Nutrition Cohort
DHA	Docosahexaenoic acid
DPA	Docosapentaenoic acid
EHNBPCCG	Endogenous Hormones, Nutritional Biomarkers and Prostate Cancer Collaborative Group
ESI	Electro-spray ionization
EPIC	European Prospective Investigation into Cancer and Nutrition
FA	Fatty acid
FDR	False discovery rate

FFP	Fish and fish products
FFQ	Food frequency questionnaires
FIA	Flow injection analysis
g/day	Grams per day (unit of dietary intake)
GWAS	Genome-wide association study
HDL	High-density lipoprotein cholesterol
HR	Hazard ratio
IARC	International Agency for Research on Cancer
ICD-10	International Classification of Diseases, 10^{th} Revision
IDL	Intermediate-density lipoprotein cholesterol
IGF-1	Insulin-like growth factor-1
LC-MS	Liquid chromatography-mass spectrometry
LDL	Low-density lipoprotein cholesterol
LL	Lower limit (of confidence interval)
Lyso PC	Lysophosphatidylcholine
MET	Metabolic equivalent of task

MR	Mendelian randomisation
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MUFA	Monounsaturated fatty acids
NHS	National Health Service
nmol/L	Nanomoles per liter (unit of concentration)
NMR	Nuclear magnetic resonance
NSHDS	Northern Sweden Health and Disease Study
OR	Odds ratio
Padj	Adjusted p-value (statistical significance after correction)
PC aa	Diacyl-phosphatidylcholines
PC ae	Acyl-alkyl-phosphatidylcholines
PCPT	Prostate Cancer Prevention Trial
PL	Phospholipid
PLCO	Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial
PRACTICAL	Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome

PRS	Polygenic risk score
PSA	Prostate-specific antigen
PUFA	Polyunsaturated fatty acids
SD	Standard deviation
SELECT	Selenium and Vitamin E Cancer Prevention Trial
SM	Sphingomyelins
SM(OH)	Hydroxysphingomyelins
SNP	Single-nucleotide polymorphism
SU.VI.MAX	Supplementation en Vitamines et Mineraux Antioxydants Study
TNM	Tumour, Node, Metastasis (cancer staging system)
UHPLC-HRMS	Ultra high-performance liquid chromatography–high-resolution mass spectrometry
UL	Upper limit (of confidence interval)
VLDL	Very low-density lipoprotein cholesterol

Table of Contents

1. Introduction

1.1 Rationale

Prostate cancer is the second-most commonly diagnosed cancer in men, with over 1.4 million cases worldwide in 2022 [1]. To date, age, family history, ethnicity, and genetics are well-established, non-modifiable prostate cancer risk factors [2]. Besides compelling evidence for free-testosterone and insulin-like growth factor 1 (IGF-1), few potentially modifiable risk factors have been established [3].

Metabolomics is a rapidly evolving field in which large numbers of metabolites (small molecules in body fluids and tissues), are measured in order to establish a metabolic fingerprint for different diseases [4,5]. These metabolite profiles can be determined by internal (i.e. genetic and epigenetic, physiological) or external (i.e. dietary and environmental) factors [6]. Several observational studies have aimed to identify novel risk factors and pathways in prostate cancer aetiology by measuring metabolite levels in in pre-diagnostic blood samples [4,6–15]. While many of these previous studies have included investigation of associations between circulating metabolites and more aggressive prostate cancer tumour subtypes, these studies have generally lacked statistical power due to limited sample size, including in analyses stratified by follow-up time and tumour subtype [4,7–10,12].

Metabolites have a key role in homeostatic processes, both in their regulation and as a byproduct, and in supporting the body's biochemical processes. As byproducts of cellular metabolism, metabolites are both involved in or result from essential functions including energy production, cellular signalling and regulation, and the biosynthesis of proteins, nucleic acids, and lipids crucial for cell structure and function. Metabolic profiles have been associated with prostate cancer development and progression; Metabolic profiles have been linked to prostate cancer development; for instance, disruptions in fatty acid metabolism (which plays a key role in energy production, membrane synthesis, and post-translational modification of signalling molecules) have been implicated in the early stages of malignant transformation [16–20]. In addition to evidence that changes in fatty acid metabolism may contribute to early oncogenic processes, they may also arise as an effect of tumorigenesis, indicating a complex, potentially bidirectional relationship [19]. Prostate cancer cells significantly rely on fatty acid absorption and metabolism, especially for long-chain fatty acids, and have increased uptake of fatty acids. Consequently, abnormal fatty acid metabolism has frequently been found in prostate cancer patients. In addition, prostate cancer tissues demonstrate higher IGF-1 and fatty acid transport protein, both which have a role in fatty acid metabolism [19]. Ultimately, the measurement and profiling of metabolites may provide insight into the aetiology of prostate cancer [4,11].

1.2 Research Aims

The aims of this DPhil are to characterize the correlates of blood metabolic profiles in men, and to investigate their associations with prostate cancer risk, using different metabolomics platforms and a sample size large enough to provide robust estimates of risk. This will involve analyses of data from two prospective cohorts (the European Prospective Investigation into Cancer and Nutrition (EPIC) and the UK Biobank using two different metabolomics platforms.

This thesis comprises three main projects, detailed below:

1. The first project of this thesis aims to determine dietary and lifestyle correlates of three metabolite patterns (and the circulating individual metabolites that comprise the patterns), which had already been derived in a previous EPIC nested case-control study using the statistical method treelet transform and been identified as being associated with aggressive prostate cancer risk [4].
2. The second project investigates the associations of the three metabolite patterns and circulating individual metabolites with aggressive prostate cancer tumour subtypes and death from prostate cancer using a larger sample size and longer follow-up in the EPIC cohort.
3. The third project aims to identify the associations between circulating metabolite concentrations and prostate cancer risk and death from prostate cancer in the UK Biobank, in order to expand on the previous project's results using a different cohort and metabolomics platform.

Figure 1.1 Aims of the thesis

This chart represents the three main aims of the thesis: (1) to investigate potentially modifiable correlates of metabolites; (2) to investigate the association between metabolites and prostate cancer risk in the EPIC cohort; and (3) to investigate the association between metabolites and prostate cancer risk in the UK Biobank.

1.3 Role of the author in the preparation of this thesis

Julie Schmidt, Ruth Travis, and Tim Key conceived the original idea to study circulating metabolites and prostate cancer risk, with research design input throughout the thesis from Karl Smith-Byrne. This work originated in order to follow up on the findings from Julie Schmidt's previous work on metabolites and prostate cancer, and with the guidance of my supervisors (Julie Schmidt, Ruth Travis, and Karl Smith-Byrne), I have been involved in all aspects of the research. I have led the literature review, data analysis and interpretation, manuscript preparation, and writing. Funding for this thesis was secured by Tim Key and Ruth Travis from Cancer Research UK.

I have written all the chapters, including reviewing all the available literature, and reporting and interpreting my original analyses. However, beyond the guidance of my supervisors, co-authors read and provided input for the two published manuscripts that were based on the results in Chapters 3 and 4 [21,22].

All metabolomics assays for EPIC were conducted by laboratory personnel at the World Health Organization's International Agency for Research on Cancer (IARC), Lyon, France. All metabolomics assays for UK Biobank were performed by laboratory personnel at Nightingale Health, Helsinki, Finland. For Chapters 3 and 5, the dataset was cleaned by former CEU statistician Urwah Noor before I began analyses, while Urwah and I cleaned the dataset (which included new incoming data) for Chapter 4 together. Afterwards, I prepared all the analytic datasets.

I conducted all statistical analyses for Chapters 3, 4, and 5 in Stata Statistical Package versions 17, 18, and 19 (StataCorp, Texas, USA) [23–25], and used R versions 4.1.0 to 4.3.2 (R Core Team, Vienna, Austria) [26] for preparation of figures and plots. For Chapters 3 and 4, Urwah advised with some technical questions, while for Chapter 5 CEU statisticians Georgina Fensom and Mahboubeh Parsaeian advised with some additional technical questions. For Chapter 4, I modified and tailored treelet transform code originally written by Georgina Fensom to fit my data and analytical specifications, while for Chapter 5 I similarly modified UK Biobank code previously written by Urwah Noor.

I designed all figures and tables, with coding assistance for forest and/or volcano plots in R provided by Karl Smith-Byrne (Chapter 3) and Mahboubeh Parsaeian (Chapters 4 and 5).

1.4 Publications arising from this thesis

The majority of the work presented in Chapters 3 and 4 has been published in *Nutrients* (2022) and the *International Journal of Cancer* (2024), and I am the first author of both manuscripts. Chapter 3 is a cross-sectional study of diet/lifestyle factors and metabolites/metabolite patterns previously associated with prostate cancer, using EPIC data. Chapter 4 is a nested case-control study of three metabolite patterns and individual metabolites with risk of more aggressive prostate cancer tumour subtypes.

2. An overview of prostate cancer epidemiology and metabolomics

2.1 Prostate cancer epidemiology

2.1.1 Geographical trends

In 2022, prostate cancer comprised 7.3% of total cancers diagnosed worldwide [11], with over 1.4 million new cases. It was the most commonly diagnosed cancer in 118 countries. In 52 countries, prostate cancer was the leading cause of cancer death, and approximately 396,000 men died from the disease [27].

There is geographic variation in incidence and mortality rates [27]. Most diagnoses are in countries with a higher human development index (HDI), with the exception of the Caribbean; the top regions with the highest age-standardised incidence rates were (in descending order) Northern Europe, Australia-New Zealand, Caribbean, and Northern America. The lowest age-standardised rates were found in Asian and African regions. The highest age-standardised death rates are found in some countries with a lower HDI, including (in descending order) southern Africa, Caribbean, middle Africa, and western Africa, while the lowest age-standardised death rates were in parts of Asia [27].

2.1.2 Temporal trends

Between 1973 and 1997, there was a global increase in prostate cancer incidence. During this period, annual incidence rates in more-developed regions (often aligned with countries classified as having a high Human Development Index [HDI]) were considerably higher than those in less-developed regions, with North America reporting the highest rates. This likely reflects greater PSA screening use in more developed countries, perhaps leading to higher detection rates, including diagnoses of indolent disease [27]. In fact, evidence suggests that prostate cancer is present in most older men: autopsy studies in several countries (including western Europe, South Africa, and Japan) have identified prostate cancer in approximately 80 percent of men who died in their seventies from other causes [28]. In recent years, however, there has been a general decrease in prostate cancer diagnoses in the US, Canada, and Australia, compared to an increase in Europe. This variation is likely due to earlier and more widespread adoption of widespread PSA testing in the former countries during the 1980s and 1990s. In contrast, European countries have begun implementing PSA testing more broadly in more recent decades, which may explain the continued increase in diagnoses [29]. Especially in the US, PSA screening for early prostate cancer detection has remained a subject of controversy. Two large randomised-controlled trials, the European Randomised Study of Screening for Prostate Cancer (ERSPC) and the Prostate, Lung, Colorectal and Ovarian screening trial (PLCO) cohort, both investigated the efficacy of PSA screening. While the ERSPC trial found that PSA testing decreased prostate cancer mortality over 11 years of follow-up, the PLCO cohort reported no difference. In fact, both trials also highlighted the adverse effects of PSA screening, including false-positives, unnecessary biopsies, and detrimental effects from treatment. In light of this evidence, the U.S. Preventive Services Task Force (USPSTF) updated its PSA screening recommendations. In 2008, it recommended against PSA screening for men 75 years and older, and in 2012, updated this recommendation to include all men, deciding that the risks of routine screening outweighed the benefits. This was largely in response to concerns that men who were unlikely to benefit from early detection, including those with limited life expectancy or at risk of overtreatment of low-risk disease, were being subjected to unnecessary screening [30].

In general, there has been a decrease in age-standardised mortality rates in high-income countries in the last thirty years, particularly in Northern America, Oceania, and Northern and Western Europe [27]. The decrease in mortality is likely due in part to improved treatment options, and possibly supported by earlier detection through PSA screening, though the latter remains subject to ongoing debate. In countries in Central and Eastern Europe, Asia, and Africa, age-standardised incidence and mortality rates increased, which may be attributed in part to increased uptake of PSA testing in previously underdiagnosed population, and persistently limited access to effective treatment compared to high-HDI countries [27].

2.1.3 Established risk factors

2.1.3.1 Age

Age is a non-modifiable, well-established prostate cancer risk factor. According to 2022 cancer statistics, men's risk of developing prostate cancer increases from 1.8 percent in men aged 60-69 years old, to nine percent in men aged 70 years and older [2]. Globally, 55 percent of deaths occur in men aged 65 years or older [29].

2.1.3.2 Family history

Family history has been established as another non-modifiable prostate cancer risk factor [31]. In a 2025 systematic review (which included observational and modelling studies), the odds of detecting clinically significant prostate cancer in men with a family history ranged from 1.3- to 8.4-fold across studies, with risk increasing further in individuals with multiple affected first-degree relatives [31]. Family history of prostate cancer may not only indicate genetic factors, but could also be a result of a shared environment of socioeconomic status, deprivation, diet, stress, and other factors, including health behaviours, awareness of the disease and likelihood of receiving testing [31].

2.1.3.3 Genetics

Genetics are also an established prostate cancer risk factor. Low-penetrance variants, or those in which not everyone with a particular genotype will present with clinical symptoms by a certain age [32], account for the majority of prostate cancer cases. Genome-wide association studies (GWAS) have identified common risk loci by analysing differences in the frequency of single-nucleotide polymorphisms (SNPs) between cases and controls across the genome. The most common variants are SNPs that are present in more than one percent of the population; while individually these SNPs only confer a small increase in prostate cancer risk, their cumulative effect can raise risk to a clinically significant level when combined [33]. For example, four SNPs (rs111906923, rs11568818, rs2735839, rs10993994) that are located within or near genes implicated in increased cancer risk may be associated with risk of aggressive prostate cancer and metastatic disease [34]. The susceptibility region on chromosome 8q24 has been estimated to contribute to 25 percent of total familial risk, more than any other known locus, and has largely been identified in men with African ancestry [34–36]. Overall, GWAS have identified 451 known prostate cancer-associated SNPs, including 269 SNPs mainly found in Europeans, and 187 novel SNPs from a recent multiethnic GWAS [33,37]. Their cumulative effect can be used to generate a polygenic risk score (PRS), which predicts an individual's lifetime risk of prostate cancer depending on their genotype, helping stratify individuals into categories of genetic risk [33].

Though rarer (found in less than 2 percent of the population), high-penetrance pathogenic variants, which include autosomal-dominant DNA-repair gene variants or mismatch repair (MMR) variants, also associated with increased prostate cancer risk and aggressiveness. For example, variants in *BRCA1 and BRACA2*, both tumour suppressor genes, are linked to elevated prostate cancer risk compared to the general population [33]. In fact, high risk variants in *BRACA2* have been associated with earlier age at diagnosis, 2.5-fold to 8.6-fold increased risk of being diagnosed by age 65 in carriers compared to non-carriers, and more aggressive prostate cancer phenotypes that have been associated with positive lymph nodes and distant metastasis [33,38]. Similarly, *BRCA1* variants confer a 1.8- to 3.5-fold increased relative risk of overall prostate cancer compared to the general population [33]. High risk *BRCA1 and*

BRACA2 genetic variants are more prevalent in individuals of Ashkenazi Jewish descent [39] than in the general White population.

Pathogenic germline variants in *ATM* and *CHEK2*, both tumour suppressor genes involved in apoptosis and DNA damage repair, respectively, have also been associated with increased risk, including for more aggressive or metastatic disease. The Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL) consortium reported an odds ratio (OR) of 4.4 for prostate cancer risk for variants in the *ATM* gene [33].

In addition, *HOXB13* variants have been associated with prostate cancer risk and aggressiveness. The *HOXB13* gene is associated with prostate cancer development, encoding the transcription factor homeobox B13 (HOXB13), an androgen receptor repressor and transcriptional regulator that modulates signals from androgen receptors [33]. A systematic review and kin-cohort study found that *HOXB13 G84E* carriers had a 3-6-fold increased risk of prostate cancer, which increased further for those who additionally had a family history of prostate cancer [40]. Similarly, in a Scandinavian population of approximately 5,000 cases, it was reported that those who carried a pathogenic germline missense variant of the gene had a 33 percent risk of developing prostate cancer, in comparison to 12 percent risk in non-carriers [38].

Finally, variants in *PALB2* (which encodes a binding protein involved in involved in DNA homologous recombination repair), *MSH2,* and *MSH6* (both MMR genes that repair errors in DNA replication) have been associated with more aggressive disease. Carriers of pathogenic *PALB2* variants have an approximately 3.5-fold increased risk of developing prostate cancer, while pathogenic variants in *MSH2* and *MSH6* have been associated with a 2-fold to 10-fold increase in overall prostate cancer relative risk [33].

2.1.3.4 Ancestry

There is variation in prostate cancer incidence amongst different ancestry groups. Black men in the United States have approximately twice the risk of being diagnosed with and dying from prostate cancer as White men [41]. Similarly, in the UK, Black British men have two-to-three times higher lifetime risk of being diagnosed with prostate cancer compared to White men, and twice the risk of prostate cancer death [42,43]. Incidence proportions reported in the UK (based on prostate cancer incidence and mortality data provided by Public Health England) found that the lifetime risk of being diagnosed with prostate cancer is 13.3% for White men, 29.3% for Black men, and 7.9% for Asian men [43]. The lifetime risk of prostate cancer death was 4.2% for White men, 8.7% for Black men, and 2.3% for Asian men [43]. A recent retrospective cohort study in England also assessed the incidence of prostate cancer within one year following a raised PSA test result in men from different ancestry groups, and reported that Black men with a raised PSA are about 1.25 times more likely to be diagnosed with prostate cancer than White men, and White men are approximately 1.48 times more likely to be diagnosed than Asian men [42].

Studies have shown that on average, African-American men have higher PSA levels than White men [29]. A large multi-ancestry GWAS published in 2021 identified 86 new prostate cancer risk variants, and reported that men of African ancestry have approximately a 2.18-fold higher genetic risk of prostate cancer compared to men of European ancestry [37]. This elevated risk appears to be primarily due to a higher frequency of high-risk alleles (such as those at loci 8q24 and KLK3) in men of African ancestry, rather than stronger per-allele risk effects [37]. In addition, the study developed and validated a genetic risk score (GRS) based on 269 variants [37]. Although the odds ratio (OR) for prostate cancer in the top GRS decile (compared to men with average genetic risk) was lower in African ancestry men (OR 3.74) than in European ancestry men (OR 5.06), the higher overall GRS distribution among African ancestry men indicates that a much larger proportion of these individuals fall into high genetic risk categories [37]. These findings are supported by other studies highlighting key variants, such as SNPs in the 8q24 region, which have been strongly associated with earlier age at diagnosis in African men [37], west African ancestry, increased risk of prostate cancer in African-American and Puerto Rican men, and positive prostate biopsy in African-American men [29,44]. In addition, a SNP in the *HOXB13* gene called X285K, which is largely unique to men of west African ancestry, has been found to be strongly associated with more aggressive prostate cancer [45].

In summary, these studies suggest that evaluating genetic ancestry, along with other risk factors, could offer a more precise way to assess individual prostate cancer risk, rather than using broad ethnic groups as proxies [45].

2.1.3.5 Free Testosterone

Supported by recent evidence [46], free testosterone is an established prostate cancer risk factor. Testosterone is a steroid hormone that is mainly synthesised by the testes, and is the most abundant androgen in men [47]. Free testosterone refers to testosterone that is not bound to blood proteins, and is therefore biologically active. In an analysis in the Endogenous Hormones, Nutritional Biomarkers and Prostate Cancer Collaborative Group (EHNBPCCG), a pooled individual participant case-control dataset of prospective studies of prostate cancer risk (including up to 25 studies, 14,944 cases and 36,752 controls, and including 1,870 aggressive prostate cancers), circulating free testosterone was positively associated with overall (OR per 1 standard deviation [SD]) 1.03) and early-onset (OR per 1 SD 1.08) prostate cancer risk [46]. Mendelian randomisation (MR) findings have further supported an association with higher disease risk through genetic evidence (and are less susceptible to biases such as confounding, reverse causation, or detection bias). Using genetic instruments that were identified through the UK Biobank (up to 194,453 men) and outcome data from PRACTICAL (including up to 79,148 cases and 61,106 controls, which included 15,167 aggressive prostate cancers), the authors reported that genetically predicted free testosterone was positively associated with overall (OR per 1 SD 1.20), early-onset (OR per 1 SD 1.37), and aggressive (OR per 1 SD 1.23) prostate cancer [46].

2.1.3.6 IGF-1

Men with higher circulating IGF-1 are at increased risk of prostate cancer [3]. Though the exact biological role of IGF-1 in prostate cancer development is unclear, IGF-1 is involved in promoting cell proliferation, and regulating differentiation and apoptosis [48]. In the most robust prospective analyses to date using data from the EHNBPCCG, IGF-1 was associated with overall (OR per 1 SD 1.09), aggressive (OR per 1 SD 1.09), and possibly early-onset disease (OR per 1 SD 1.11) [3]. MR analyses using instruments identified via UK Biobank, and outcome data obtained from PRACTICAL, reported a similar association between higher IGF-1 and higher risk of overall (OR per 1 SD 1.07), aggressive (OR per 1 SD 1.10), and possible early-onset disease (OR per 1 SD 1.13), complementing the observational findings with genetic evidence [3]. This study represents the largest available collection of both observational and genetic data on IGF-1 and prostate cancer risk, and the pooled analyses represent nearly all of the available global data [3].

2.1.4 Body Mass Index

Prospective studies have demonstrated an inverse association between body mass index (BMI) and risk of overall prostate cancer [49], but a positive association with risk of more advanced disease and prostate cancer mortality [50,51]. Evidence from a recent MR analysis also supports a protective association between higher BMI and overall prostate cancer risk [49]. However, the extent to which BMI biologically influences disease initiation and progression, or whether detection bias contributes to the observed associations with less aggressive tumours (inverse association) and more aggressive tumour subtypes (positive association), remains unclear.

For example, men with obesity are less likely to undergo a prostate biopsy, primarily because higher BMI is associated with lower PSA concentrations, and also because obesity makes thorough digital rectal examination more challenging. This can lead to later detection of tumours, and worse prognosis. Additionally, because men with obesity often have enlarged prostates, the sensitivity of biopsies is further reduced, also contributing to delayed cancer diagnosis [51].

2.2 Metabolomics

The aim of *omics* approaches is largely to identify and quantify the biomolecules present in an organism (e.g. in tissue, cell, fluid, etc.), generating large datasets that can be analysed via bioinformatics and biostatistics. Metabolomics, which can be broadly defined as the measurement of metabolites (small molecules) in a biological specimen, is an emerging field. Meanwhile, the metabolome refers to the comprehensive set of metabolites expressed by an organism. Enabling the profiling of much larger numbers of metabolites than those covered in current standard clinical laboratory panels, metabolomics can provide promising insight into metabolic pathways. When developing analytical procedures, researchers can employ a combination of different metabolomics platforms to analyse metabolite subsets with shared functional groups or similar structures. Thus, the metabolome can be profiled through integrated methodologies [52,53].

2.2.1 Multiplex Metabolomics Platforms

2.2.1.1 Mass Spectrometry

Mass spectrometry (MS) is an analytical technique used to quantity metabolites (present in biological samples such as plasma, serum, urine, tissues, and cell extracts) by identifying the molecular mass to charge ratio of compounds to determine their structure and abundance [54]. Liquid-chromatography-mass spectrometry (LC-MS) is the most commonly used method in MS; the technique involves combining high-pressure liquid chromatography with mass spectrometry, separating molecules depending on different physical and chemical properties (including affinity towards other molecules, charge, polarity, and molecular size) [55]. Advantageous for its high sensitivity and specificity compared to other analytical methods, MS studies the gas phase of ionized molecules [54,56]. This approach concentrates on the quantification of selected metabolites, including those involved in a metabolic pathway of interest. MS is particularly useful for targeted metabolomics, in which the metabolites being investigated are already known.

2.2.1.1.1 Biocrates

The Biocrates Absolute*IDQ*™ p180 kit is a commercially available high-throughput targeted metabolomics assay, which uses tandem mass spectrometry (MS) to quantify metabolites. It detects 21 amino acids, 21 biogenic amines, hexose, 40 acylcarnitines, 14 lysophosphatidylcholines, 76 phosphatidylcholines, and 15 sphingolipids (insert Biocrates website source).

Several prospective cohorts have used the Biocrates Absolute*IDQ*™ p180 kit to quantify metabolites, including the European Prospective Investigation into Cancer and Nutrition (EPIC) and the Northern Sweden Health and Disease Study (NSHDS). Two of the three results chapters in this thesis will use data from the EPIC cohort, while the NSHDS will be referenced in discussion chapters as a comparable cohort.

According to the manufacturer's protocol, human plasma samples are first vortexed and centrifuged to ensure proper mixing and separation. Using a 96-well plate with filter inserts, 10 µL of each prepared plasma sample is pipetted onto the filters, along with a zero sample (phosphate-buffered saline), seven concentrations of calibration standards, and internal standards that include quality control samples at low, medium, and high concentrations. After the filters are dried in nitrogen stream, derivatisation solution (5% phenylisothiocyanate reagent, PITC; 50 µL required) is added and the filters are dried again. Metabolites and internal standards are then extracted using 5 mM ammonium acetate in methanol, and afterwards the solution is centrifuged through the filter membrane.

Amino acids and biogenic amines are analysed using liquid chromatography (LC)-electrospray ionization (ESI) coupled to tandem mass-spectrometry (MS/MS), and are validated using external standards in seven different concentrations. The acylcarnitines, glycerophospholipids, sphingolipids, and hexose are measured by flow injection analysis (FIA)-ESI-MS/MS, and are quantified using a one-point internal standard calibration (compared to seven-point for the amino acids and biogenic amines). Several acylcarnitines, lipids, and FIA-detected metabolites are considered semi-quantitative because standards were not commercially available.

Assessment of the reproducibility of the kit on human plasma and serum samples following the manufacturer's procedure across six participating laboratories found it to be highly reproducible [57].

2.2.1.1.1.1 Information on Metabolites

Throughout this thesis, fatty acid side chains in acylcarnitines, glycerophospholipids and sphingolipids are referred to "Cx:y", (x and y respectively refer to the total number of carbon atoms (x) and double bonds (y) in each molecule. This abbreviation reflects the specific fatty acid side chain present for acylcarnitines. All measured glycerophospholipids in the Biocrates chapters in this thesis are phosphatidylcholines, which are further sub-categorised based on the type and number of fatty acid side chains: Lysophosphatidylcholines (lyso PC a) refer to phosphatidylcholines with a single acyl chain; diacyl-phosphatidylcholine (PC aa) indicates diacyl (two acyl chains); and acyl-alkyl phosphatidylcholine (PC ae) designates molecules with one acyl and one alkyl side chain. The sphingolipids assessed are all sphingomyelins (SM), which may contain a hydroxy (OH) group, and are also classified by their fatty acid side chains. Finally, hexose refers to the combined measurement of several six-carbon monosaccharides, such as glucose, fructose, and galactose [58,59].

2.2.1.1.2 Metabolon

Metabolon is a second platform that uses MS to analyse metabolites, quantifying them using ultra high performance liquid chromatography/mass spectrometry (UHPLC/MS) and gas chromatography/mass spectrometry (GC/MS). Metabolites are categorised according to eight chemical classes: amino acids (including amino acid derivatives), carbohydrates, cofactors and vitamins, energy metabolites, lipids, nucleotides, peptides, and xenobiotics [9,12]. The Alpha-tocopherol, Beta-carotene Cancer Prevention Study (ATBC), PLCO (see Table 2), and Cancer Prevention Study-II Nutrition Cohort (CPS-II) cohorts used the Metabolon platform to measure a total of approximately 650 metabolites [8,9,14,60].

2.2.1.2 NMR Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is a second method used to identify metabolites from a serum or plasma sample. In brief, NMR spectroscopy exposes nuclei in a magnetic field to a radiofrequency pulse, causing their protons to temporarily transition to a higher energy state. When they return to their normal state, this produces a distinctive radiowave. Different metabolite have unique energy release patterns, which are demonstrated as peaks in a chromatogram (while the area of the peaks shows the concentration of each metabolite). In addition to being highly reproducible and quantitative, NMR is ideal for identifying unknown metabolites (untargeted metabolomics), is low in cost, and requires minimal sample preparation. [61,62].

2.2.1.2.1 Nightingale Health

Nightingale Health Plc market a metabolomics platform that uses NMR to identify metabolite biomarkers of disease. Nightingale Health partnered with the UK Biobank, a prospective cohort of 500,000 individuals who have voluntarily shared their health information with scientists in order to advance health research. The Nightingale Health-UK Biobank initiative is one of the largest metabolic profiling studies to date.

In 2021, the first tranche of data was released, which included detailed metabolomic assay results for approximately 120,000 UK Biobank participants. The second tranche of analysed blood samples for an additional 200,000 participants was released in 2023, with the ultimate goal of releasing assay data for all UK Biobank participants.

Classes of metabolites measured by Nightingale Health (n = 249) include: lipids, fatty acid composition metabolites, low-molecular weight metabolites (amino acids (n = 10), ketone bodies and glycolysis metabolites), and 14 lipoprotein subclasses (triglycerides, phospholipids, total cholesterol, cholesterol esters, and free cholesterol, and total lipid concentration within each subclass). In addition to absolute concentrations, the platform also provides derived features, including metabolite ratios and percentages of total components [63,64].

2.2.1.3 How the techniques compare

LC-MS is a widely used metabolomics approach, ideal for a targeted approach where metabolites are initially identified and quantitated. In addition to providing valuable insights regarding the dynamics and fluxes of metabolites, this quantitative method allows for identifying metabolite classes with better classification accuracy. Meanwhile, NMR's high resolution and reproducibility allow for the identification of a large number of metabolites and accurate determination of their concentrations. NMR is ideal for identifying unknown metabolites via an untargeted approach. NMR and LC-MS studies are complementary, in tandem identifying larger numbers and more classes of metabolites, thus covering more of the metabolome (the complete set of metabolites found in a biological sample) [61,62,65,66]

2.3 Prospective studies of circulating metabolites and prostate cancer

There have been several prospective studies of circulating metabolites and prostate cancer, using both MS and NMR (see Table 2.1). The MS studies reviewed here mainly used the Biocrates and Metabolon platforms. This section will review the key findings.

2.3.1 Acylcarnitines

Acylcarnitines have been associated with more aggressive prostate cancer tumour subtypes in prior prospective studies, particularly C18:2 and C18:0 [4,9,10]. Higher circulating concentrations of acylcarnitine C18:2 were associated with a lower risk of more aggressive disease in both EPIC and the NSHDS [4,10]. C18:0 was also implicated in prostate cancer risk, though in a different direction; in the PLCO cohort, stearoylcarnitine (C18:0) was positively associated with overall and aggressive prostate cancer [9].

2.3.2 Lysophosphatidylcholines

Lysophosphatidylcholines (lyso PC a) have been associated with aggressive, advanced, and lethal prostate cancer across several cohorts. The direction of association reported was mainly inverse, with the exception of the NSHDS [4,10,14]. In EPIC, lysophosphatidylcholines C16:0, C16:1, C18:0, and C18:1 were inversely associated with advanced prostate cancer, and in the ATBC cohort, higher concentrations of lysolipid 1-linoleoyl-glycerophosphatidylcholine (18:2) (also known as lysoPC a C18:2) were associated with lower risk of lethal prostate cancer [4,14]. Contrastingly, in the NSHDS, lysophosphatidylcholines C17:0 and C18:0 were positively associated with overall and aggressive disease [10].

2.3.3 Lipid metabolites

Higher levels of phosphatidylcholines and other lipid metabolites have also been implicated in reduced risk of overall and more advanced tumour subtypes for several cohorts, with the exception of one cohort, which reported a positive association [4,8,10,67]. In EPIC, two diacyl-phosphatidylcholines (PC aas) associated inversely with aggressive disease, and 21 PC diacyl-phosphatidylcholines and nine acyl-alkyl-phosphatidylcholines (PC aes) associated inversely with advanced prostate cancer [4]. Similarly, in the ATBC cohort, higher levels of six glycerophospholipids were associated with lower risk of tumour stage T3 prostate cancer [8], which is included in the advanced cancer category in EPIC [4]. However, in the NSHDS, three acyl-alkyl phosphatidylcholines were associated with greater risk of aggressive disease [10].

2.3.4 Sphingomyelins

For studies that found an association between sphingomyelins and sphingolipid-related metabolites, the association has mainly been inverse, with the exception of one cohort. The

EPIC and NSHDS cohorts both reported associations, which were inconsistent [4,13]. In EPIC, sphingomyelins C16:1 and C24:0 were linked to advanced prostate cancer, with the former associated with higher risk and the latter with lower risk [4]. Meanwhile, in the NSHDS, one sphingomyelin was associated with greater risk of overall disease in the untargeted study [13]. In line with the EPIC findings, however, the CPS-II cohort reported an inverse association between sphingomyelin d18:1/18:0 (also known as SM C36:1) and risk of prostate cancer death [60]. Similarly, the supplementation en Vitamines et Mineraux Antioxydants Study (SU.VI.MAX) cohort reported an inverse association between two sphingolipid bases, sphinganine and sphingosine, with overall prostate cancer risk (using MS to detect metabolites via an in-house standard) [68].

2.3.5 Amino acids

Several cohorts have found an association between certain amino acids and overall or more aggressive prostate cancer. However, the particular amino acids, and direction of risk, vary within and across studies [9,10,14,15]. In the ATBC cohort, non-proteinogenic amino acids such as thioproline, and thioproline combined with cysteine-related amino acids were inversely associated with risk of prostate cancer death, while the dipeptide leucylglycine and three gamma-glutamyl amino acids were positively associated with risk of prostate cancer death [14]. The NSHDS also reported an association between glycine and greater risk of overall prostate cancer [10]. In the PLCO cohort, higher concentrations of Pyroglutamine, gamma-glutamylphenylalanine, phenylpyruvate, and *N*-acetylcitrulline were associated with lower risk of overall and aggressive prostate cancer [9]. Finally, in the SU.VI.MAX cohort, the amino acids valine, tyrosine, histidine, methylhistidine, and phenylalanine were associated with greater overall prostate cancer risk [15].

2.3.6 Glycolysis metabolites

Two cohorts reported an inverse association between glycolysis-linked metabolites and prostate cancer [10,67]. An analysis in the UK Biobank reported a strong inverse association between glucose and citrate and overall prostate cancer [67], and similarly, the NSHDS found an inverse association between the glycolysis metabolite pyruvate and overall prostate cancer in men aged 40-50 years [10]. The ATBC cohort also reported an inverse association between citrate and more aggressive prostate cancer [12]

2.3.7 Fatty acids

Previous work has also reported an association between fatty acids and prostate cancer risk, especially a positive association of long-chain omega-3 fatty acids with overall risk. In an individual participant meta-analysis, including 5,098 cases from seven studies, the saturated fatty acid stearic acid was inversely associated with overall prostate cancer, while higher levels of the omega-3 long chain polyunsaturated fatty acids eicosapentaenoic acid and docosapentaenoic acid were each associated with a 14% and 16% increased risk of prostate cancer, respectively (though there was heterogeneity in the latter two associations between studies) [69]. Two prospective studies included in the analysis reported a positive role of omega-3 long chain polyunsaturated fatty acids, including docosahexaenoic acid, with overall prostate cancer risk [70,71].

2.3.8 Dimension-reduction methods in prospective studies

Analysing each metabolite individually can present statistical challenges, especially due to correlations among metabolites. Dimension reduction methods (e.g. principal component analysis (PCA) and treelet transform) can address these challenges, by summarising complex metabolite data into a smaller number of components or patterns that capture the main sources of variation within the data [72]. By grouping correlated metabolites into patterns, dimension reduction methods can improve statistical power and reduce noise. In particular, treelet transform offers an advantage over traditional approaches like PCA by naturally producing sparse, interpretable components and revealing the underlying structure of the data through its combination of hierarchical clustering and dimension reduction [73].

In EPIC, three metabolite patterns derived using treelet transform were identified to associate with prostate cancer risk (details in Chapters 3 and 4): a metabolite pattern characterized by 64 phosphatidylcholines (including PC aas and PC aes) and three hydroxysphingomyelins, as well as a metabolite pattern of acylcarnitines C18:1 and C18:2, glutamate, ornithine, and taurine, were both found to be associated inversely with advanced and aggressive prostate cancer risk. Furthermore, a metabolite pattern of eight lysophosphatidylcholines was observed to be associated inversely with risk of advanced prostate cancer and prostate cancer death [4]. Treelet transform was previously applied in EPIC-Potsdam untargeted metabolomics and lipidomics data, where a pattern characterised by positive loadings on a sphingomyelin, three phosphatidylcholines and other glycerophospholipids (a metabolite class that includes phosphatidylcholines) was identified [74]. Outside of EPIC, the PLCO cohort also applied dimension-reduction methods to a full set of 695 metabolites, deriving ten metabolite patterns that were found to not be associated with overall or aggressive prostate cancer [9].

2.3.9 Recent meta-analysis

Recently, a systematic review and meta-analysis of twelve pre-diagnostic circulating metabolomic studies and prostate cancer risk identified several metabolites associated with overall, high-risk, and lethal prostate cancer, including the previous work in EPIC, using MS (including Metabolon and Biocrates) and NMR platforms [75]. In the study, the authors suggest that the metabolite-prostate cancer associations identified may be useful for risk prediction [75]. Three metabolites, including the amino acid-derived metabolites serotonin and tiglylcarnitine and the lipid sphinganine, were associated with reduced overall prostate cancer risk after multiple testing. Higher concentrations of nine phosphatidylcholines and the carnitine C18:2 were associated with reduced risk of high risk prostate cancer. Higher concentrations of three lipids (androstenediol (3beta,17beta) disulfate (2), 3hydroxybutyrate, and glycerol), the amino acids glutamate, taurine, C-glycosyltryptophan, and N-acetylserine, and three peptides were associated with higher risk of death from prostate cancer. Meanwhile, higher concentrations of the glycerophospholipid 1-linoleoyl-GPC, the amino acid cysteine, and non-proteinogenic amino acid derivative 4hydroxyphenolpyruvate were found to associate with lower risk of prostate cancer death. Finally, no metabolites associated with low- to intermediate-risk prostate cancer after correction for multiple testing. Ultimately, lipids were the metabolites most commonly associated with high and very high risk disease. There were few common metabolite associations across outcomes, suggesting that there may be distinct metabolic profiles associated with overall prostate cancer, high risk prostate cancer and risk of prostate cancer death [75].

However, comparing results across different cohorts and metabolomics platforms is challenging due to variations in analytical methods, populations, and metabolites. Furthermore, several of the included studies in the meta-analysis were modest in size, and more evidence is needed to confirm and interpret the magnitude of metabolite-prostate cancer risk associations. To address this research need, for this thesis I have conducted analyses in two distinct cohorts, using extended datasets to provide new prospective metabolomics and prostate cancer data. A separate systematic review was not conducted as part of this thesis, because a comprehensive meta-analysis has already been published, and because methodological heterogeneity across studies (different platforms, populations, and metabolites assessed) makes direct comparisons complex.

Table 2.1 Studies in other cohorts investigating circulating metabolites and prostate cancer risk

Platform	Cohort	Study	Study Design	Sample size
Biocrates (MS)	EPIC	**Schmidt et al. (2020) [4]**	Case-control	3,057 matched sets
		*Breeuer et al. (2022) [76]	Case-control	533 advanced matched sets
				1,301 localized matched sets
		*Schmidt et al. (2017) [6]	Case-control	1,077 matched sets
		*Kuhn et al. (2016) [11]	Case-cohort	310 incident cancer cases
	NSHDS	Östman et al. (2022) [13] (Untargeted metabolomics)	Case-control	752 matched sets
		Röhnisch et al. (2020) [10] (Targeted metabolomics)	Case-control	777 matched sets
Metabolon (MS)	ATBC	Huang et al. (2019) [14]	Case-control	523 matched sets
		Huang et al. (2017) [8]	Case-control	T_2 (n = 71), T_3 (n = 51), and T_4 (n = 15) matched to 200 controls
		*Mondul et al. (2015) [12]	Case-control	200 matched sets
		*Mondul et al. (2014) [7]	Case-control	74 matched sets
	PLCO	Huang et al. (2016) [9]	Case-control	380 matched sets
	CPS-II	Wang et al. (2021) [60]	Case-cohort	241 incident cancer cases
UHPLC-HRMS (MS, metabolites detected via in-house standard)	SU.VI.MAX	Lin et al. (2021) [68]	Case-control	146 cases matched to 272 controls
Nightingale Health (NMR)	UK Biobank	Hu et al. (2024) [67]	Prospective cohort	83,290 participants (1,436 cases)
BRUKER AVANCE III (NMR)	SU.VI.MAX	Lécuyer et al. (2021) [15]	Case-control	171 matched sets

Table 2.1 Studies in other cohorts investigating circulating metabolites and prostate cancer risk. Studies with a * indicate that the data included is a subset of data in the bolded study above, and thus will not be referred to here forth. EPIC, European Prospective Study into Cancer and Nutrition; NSHDS, Northern Sweden Health and Disease Study; ATBC, the Alpha-tocopherol, Beta-carotene Cancer Prevention Study; PLCO, the Prostate, Lung, Colorectal, and Ovarian Screening Trial; CPS-II, the Cancer Prevention Study-II Nutrition Cohort; The Supplementation en Vitamines et Mineraux Antioxydants Study. MS, mass spectrometry; UHPLC, ultra high

Table 2.1 Studies in other cohorts investigating circulating metabolites and prostate cancer risk
performance liquid chromatography; HRMS, high-resolution mass spectrometry; NMR, nuclear magnetic resonance; T_2, T_3, and T_4; stage of tumour based on tumour node metastasis staging.

3. A cross-sectional study of correlates of prostate-cancer associated metabolites in men

3.1 Introduction

Blood metabolite concentrations are affected by both internal and external factors, including modifiable factors, such as diet and body mass index (BMI) [77,78]. Previously in an EPIC cross-sectional study of breast cancer-associated metabolites in women, higher total and saturated fat intakes were associated with higher concentrations of four phosphatidylcholines, while alcohol consumption was associated with lower concentrations of one phosphatidylcholine [79]. In the UK Fenland Study cohort, which also used the Biocrates platform to quantify metabolites, several components of the Mediterranean diet were associated with variations in plasma acylcarnitines, amino acids, and phospholipids in a cross-sectional study [80].

In the prior case-control study nested within the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort, three patterns were identified and assessed in relation to risk for prostate cancer [4]; a metabolite pattern of 64 diacyl- and acyl-alkyl phosphatidylcholines and three hydroxysphingomyelins, as well as a metabolite pattern of two acylcarntines, glutamate, ornithine, and taurine, were both found to be inversely associated with risk of advanced and aggressive prostate cancer. Furthermore, a metabolite pattern of eight lysophosphatidylcholines was observed to be inversely associated with risk of advanced prostate cancer and prostate cancer death [4]. Thus, a better understanding of how these prostate cancer-related metabolite patterns can potentially be modifiable through dietary and lifestyle factors might offer insights into possible avenues for prostate cancer prevention.

3.2 Aims

In this cross-sectional study nested in the EPIC cohort, I aim to investigate the associations of dietary variables and BMI with metabolite patterns previously found to be inversely associated with more aggressive prostate cancer tumour subtypes. I will also investigate whether these exposure-metabolite pattern associations are driven by specific metabolites loading on each pattern.

3.3 Methods

3.3.1 Study population

EPIC is a multi-centre prospective cohort study, which recruited approximately 500,000 Europeans, including 153,457 men, between 1992 and 2000. The current analyses include men mainly aged between the ages of 35 and 70 years at recruitment from 19 centres in seven countries (Denmark, Germany, Italy, Netherlands, Spain, Sweden and United Kingdom) [81]. Detailed information on dietary, lifestyle, and anthropometric data was gathered at recruitment, previously described in *Riboli et al* [81]. 139,600 men provided blood samples [81]. EPIC participants signed an informed consent agreement, then were mailed two questionnaires, on diet and lifestyle, respectively, both which were completed at home by most participants. Participants were then invited to a study centre for an examination, which included collection of the completed questionnaires, obtainment of blood samples, and measurement of anthropometry [81].

In all EPIC centres (except the Oxford cohort), height, weight, and waist and hip circumference were measured on all subjects using similar protocols. In Oxford, weight, height, and waist and hip circumference were mainly self-reported in the questionnaires [81].

Centres in Germany, Italy, The Netherlands, Spain, and the UK collected questionnaire information on type of physical activity at work, physical exercise to keep fit, and vigorous physical activity, as well as time spent on specific activities including walking, cycling, gardening, housework, and number of stairs climbed per day. Meanwhile, non-dietary questionnaire questions in Denmark and Sweden differed slightly from the other EPIC centres, though Denmark had complete information, and Sweden had the majority of the information available [81].

All EPIC centres had information on smoking status (current, past, never), as well as smoking intensity (amount of cigarettes smoked). In addition, all centres used in this study (except those in The Netherlands) have information on current and past cigar and pipe smoking.

For all EPIC centres, information on current alcohol consumption was available from the dietary questionnaires. Centres in Germany, Italy, The Netherlands, Spain, and the UK had information on past amount of wine, beer/cider, fortified wine and spirit/liquor consumed. Sweden and Denmark had information on current alcohol consumption only. Non-dietary questionnaire questions also included level of education [81].

All participants in the EPIC study provided written informed consent, and the study was approved by the ethics committees of the International Agency for Research on Cancer (IARC) and all participating centres [81].

Men were eligible for the current study if they (1) were free of cancer (except non-melanoma skin cancer) at baseline; (2) had a known date of blood collection; (3) had been included as control participants in one of four case-control studies on metabolite concentrations and cancer risk nested within the EPIC cohort (on prostate [4], colorectal [82], kidney [83], and liver cancer [84]) with available blood concentrations of all of the metabolites included in the metabolite patterns that were previously found to be associated with more aggressive prostate cancer subtypes; and (4) had blood samples that were included in an analytical batch that had at least 10 samples, to ensure proper normalization of metabolite concentrations. Thus, data for 3,198 men were available for this study.

3.3.2 Laboratory measurements

For participants from Germany, Italy, the Netherlands, Spain and the UK, biological samples are stored at IARC in plastic straws at -196°C (details published elsewhere) [81]. In Sweden and Denmark, blood samples are stored in tubes in local repositories; in Sweden, the samples are kept in freezers at -80°C, and in Denmark in nitrogen vapour at -150°C [81].

All samples were previously assayed at IARC in Lyon, France using the Absolute*IDQ* p180 Kit (Biocrates Life Sciences AG, Innsbruck, Austria), and following the procedure recommended by the vendor. To quantify metabolites, liquid chromatography mass spectrometry (LC-MS) was applied. All samples were assayed using one LC instrument (Agilent 1290, Santa Clara, CA) coupled with one of two different triple quadrupole MS instruments (Triple Quad 4500, AB Sciex, Framingham, MA for prostate and colorectal cancer [4,82] and Q-Trap 5500, AB Sciex, MA for liver and kidney cancer [84,85]). Of note, within each cancer site-specific study a single pair of LC-MS instruments was used for all samples (Appendix 3.2) [86]. 118 common metabolites were measured across all studies [86].

3.3.3 Diet, BMI, and covariate data

Detailed information on dietary, lifestyle, and anthropometric data was gathered at recruitment, previously described in *Riboli et al* [81]. At baseline, lifestyle, socio-economic, and medical factors were assessed via a questionnaire. In order to determine usual dietary intakes, centre- or country-specific validated dietary questionnaires covering the previous 12 months were used [87]. The dietary variables (continuous, consumption in g/day) investigated in this study were intakes of total dairy (sum of milk, cheese, and yogurt), milk, cheese, yogurt, eggs, red meat, poultry, processed meat, total fish products (refers to fish and shellfish combined), total fish (subset of total fish products), fatty fish (subcategory of total fish), lean fish (subcategory of total fish), fats and oils (sum of butter, margarine, and vegetable oils), butter, margarine, vegetable oils, total vegetables (sum of leafy, root, and fruiting vegetables), leafy vegetables, root vegetables, fruiting vegetables, total fruit, cereals and cereal products, and alcohol. In order to reflect average daily consumption, increments were chosen for each dietary variable to represent typical intakes in an average European male population (Table 3.1) [88–93].

Table 3.1 Increments for each dietary exposure variable

Dietary variable	Increment (grams per day)
Dairy	200 g
Milk	200 g
Cheese	30 g
Yogurt	30 g
Eggs	7 g
Total fish products	30 g
Total fish	30 g
Lean fish	10 g
Fatty fish	10 g
Red meat	40 g
Poultry	20 g
Processed meat	40 g
Fats and oils	10 g
Butter	5 g
Margarine	5 g
Vegetable oils	5 g
Fruits	100 g
Vegetables	100 g
Leafy vegetables	25 g
Root vegetables	25 g
Fruiting Vegetables	100 g
Cereals and cereal products	200 g
Alcohol	10 g

BMI (continuous, kg/m^2) was also examined as a possible correlate of the metabolite patterns, calculated from weight and height (measured, except self-reported in some participants in the EPIC-Oxford cohort) [82].

3.3.4 Statistical analysis

3.3.4.1 Participant characteristics

Participants' characteristics at baseline were summarized using frequencies for categorical variables and mean (standard deviation) for continuous variables. To further contextualise a key finding, average total fish intake by country (mean and standard deviation) is also shown. In order to evaluate whether fasting status could bias comparisons between the discovery and validation sets, adjusted mean metabolite pattern scores were compared across fasting categories in each dataset (Table 3.7). Adjusted means and 95% confidence intervals were obtained from linear regression models with metabolite pattern scores as the outcome and fasting status as the exposure, adjusted for age, time of day of blood collection, level of education, physical activity, energy intake, alcohol intake, and BMI. Likelihood ratio tests were used to evaluate heterogeneity across fasting categories within each dataset (overall, discovery, and validation).

3.3.4.2 Normalization of metabolite concentrations

A statistical pipeline has been developed for the EPIC metabolomics data [86] and was applied in this analysis to the raw metabolite concentrations. In brief, log-transformed concentrations of the metabolites of interest were normalized to remove effects of analytical batch and study, which were estimated as random effects in mixed-effects linear models correcting for possible heteroscedasticity, while centre was a fixed effect to account for heterogeneity by centre [86]. Corrected metabolite concentrations investigated in this work correspond to residuals from the model. This pipeline was shown to be efficient in removing unwanted variability and improving the comparability of measurements acquired across the different cancer-specific studies [86].

3.3.4.3 Metabolite patterns

Patterns in metabolite profiles were previously identified using treelet transform in an EPIC nested case-control study of prostate cancer [4]. In summary, treelet transform is a linear dimension-reduction method aiming at summarizing the metabolite variables into fewer latent variables that best capture the observed variation in the overall set of metabolites [73,94]. *Schmidt et al.* identified three treelet components (henceforth referred to in the text as metabolite patterns, which together explained 31.4% of the total variance in metabolite concentrations), all of which were found to have an inverse association with advanced and/or aggressive prostate cancer risk. The first treelet component (Pattern 1) had positive loadings on all diacyl-phosphatidylcholines (PC aa; n=31) and acyl-alkyl-phosphatidylcholines (PC ae; n=33), as well as three hydroxysphingomyelins (SM(OH)): C14:1, C16:1, and C22:2. The second treelet component (Pattern 2) had positive loadings on acylcarnitines C18:1 and C18:2, and the amino acids glutamate, ornithine, and taurine. Finally, the third treelet component (Pattern 3) had positive loadings on eight lysophosphatidylcholines (lyso PC a): C16:0, C16:1, C17:0, C18:0, C18:1, C18:2, C20:3, and C20:4 (Table 3.2) [4]. Full details of each individual metabolite loading on each metabolite pattern can be found in Appendix 3.1. Each metabolite pattern was scaled to units of one standard deviation (SD) as done in the previous study [4].

Table 3.2 Metabolite patterns and their loadings

Metabolite Pattern	Contributing metabolites all with positive loadings	Percent explained variance (%)
1	All diacyl and acyl-alkyl phosphatidylcholines; (SM (OH) C14:1, SM (OH) C16:1, and SM (OH) C22:2)	21.5
2	Acylcarnitines C18:1 and C18:2, glutamate, ornithine, and taurine	5.2
3	Lyso PCs C16:0, C16:1, C17:0, lC18:0, C18:1, C18:2, C20:3, and C20:4	4.7

Table 3.2 Metabolite patterns and their loadings. SM(OH), hydroxysphingomyelin; lyso PC, lysophosphatidylcholine.

3.3.4.4 Correlates of metabolites

After excluding participants with missing values for time at blood collection (78), fasting status (65), energy intake (2), BMI (23), and level of education (31), the current cross-sectional analysis included data from 3,042 participants. These data were subsequently split into a discovery set (N=2,524; 83% of the population) and a validation set (N=518; 17.0% of the population). Specifically, the discovery set included all controls from the EPIC prostate cancer nested case-control study (used in the identification of metabolite patterns), while the validation set comprised the male controls from the other three EPIC cancer sub-studies (kidney [83], liver [84], and colorectal [82] cancer). Table 3.3 shows the distribution of controls from each study, used to comprise the discovery and validation sets, tabulated by country.

Table 3.3 Distribution of control participants by study and country in the discovery and validation sets

Cancer study (Set)	Italy	Spain	UK	Holland	Germany	Sweden	Denmark	Total
Prostate (Discovery)	446 (16.9%)	572 (21.7%)	675 (25.6%)	168 (6.4%)	779 (29.5%)	-	-	2,640
Colorectal (Validation)	75 (48.1%)	35 (22.4%)	24 (15.4%)	4 (2.6%)	18 (11.5%)	-	-	156
Kidney (Validation)	51 (18.4%)	72 (26.0%)	47 (17.0%)	12 (4.3%)	95 (34.3%)	-	-	277
Liver (Validation)	14 (11.2%)	17 (13.6%)	21 (16.8%)	1 (0.8%)	32 (25.6%)	20 (16.0%)	20 (16.0%)	125
Total	586 (18.3%)	696 (21.8%)	767 (24.0%)	185 (5.8%)	924 (28.9%)	20 (0.6%)	20 (0.6%)	3,198

Table 3.3. Distribution and percentage contribution of control participants by study and country in the discovery and validation sets. Values are number of control participants, with percentages indicating the proportion contributed by each country within that study/set. Percentages in the total row indicate the proportion across all controls combined (n = 3,198).

A discovery-validation set design was chosen to both reduce the in-sample bias from the samples used to determine patterns, and to afford an external validation for any associations that appeared statistically significant in initial analyses.

First, analyses were run in the discovery set. For each of the three metabolite patterns and each dietary or lifestyle variable, a linear regression model was run with the metabolite pattern as the dependent variable. Models were adjusted for age at blood collection (continuous), time of day of blood collection (continuous), fasting status at blood collection (< 3 hours since last meal, 3–6 hours, > 6 hours, and missing), baseline education level (primary/no schooling, secondary, professional/technical, university/higher, not specified, and missing), physical activity (Cambridge index [95]: inactive, moderately inactive, moderately active, active, and missing), smoking status (never, former, current, and missing), energy (continuous, kcal/day) and alcohol intakes (continuous, g/day), and BMI (continuous, kg/m²). Models that examined alcohol intake and BMI as main exposures were not adjusted for alcohol intake and BMI, respectively.

In the discovery set, to account for multiple testing, I used a Benjamini-Hochberg false discovery rate (FDR) by metabolite pattern at a 5% threshold to define statistical significance [96]. Each statistically significant association in the discovery set was re-assessed in the validation set, using the same variables and adjusted models. Results from the analyses in the validation set were not corrected for multiple testing. Associations between exposures and metabolites that passed the FDR threshold in the discovery set, and the significance threshold in the validation set ($P < 0.05$), were considered robust (Appendix 3.3).

3.3.4.5 Individual metabolite analysis

A supplementary analysis was conducted of dietary exposures and BMI with the individual metabolites that contributed to metabolite patterns with which they were robustly associated (Appendix 3.4). Models were adjusted as described above for the main analysis. Individual metabolite values were log-transformed. Linear regressions were run in the overall dataset (n between 2,136 and 3,042, depending on exposure). To account for multiple testing, dietary and lifestyle correlates of metabolites that passed the FDR of 0.05 were determined to be statistically significant.

3.4 Results

3.4.1 Participant characteristics

Main characteristics of the participants, overall and in the discovery and validation sets, are shown in Table 3.4. 46.4% of men in the discovery set and 31.7% of men in the validation set were not considered fasting at blood collection (< 3 hours since last meal), while 32.0% and 45.7% of men in the discovery and validation sets, respectively, were fasting (> 6 hours since last meal). Otherwise, participant characteristics were relatively similar in the discovery and validation sets.

Table 3.4 Main characteristics of men included in the analysis, overall and separately in discovery and validation sets

	Overall (n=3,198)	**Discovery (n=2,640)**	**Validation (n=558)**
Age at blood collection (years)	57.2 (7.2)	57.5 (7.1)	56.0 (7.8)
Fasting status at blood collection (time since last meal) (n (%))			
<3h	1402 (43.8)	1225 (46.4)	177 (31.7)
3-6h	631 (19.7)	526 (19.9)	105 (18.8)
>6h	1100 (34.4)	845 (32.0)	255 (45.7)
Missing	65 (2.0)	44 (1.7)	21 (3.8)
Socio-economic and lifestyle factors (n (%))			
Education level			
Primary/no schooling	1216 (38.0)	992 (37.6)	224 (40.1)
Secondary	347 (10.9)	289 (11.0)	58 (10.4)
Technical/professional	744 (23.3)	612 (23.2)	132 (23.7)
University or higher	761 (23.8)	633 (24.0)	128 (22.9)
Not specified	99 (3.1)	88 (3.3)	11 (2.0)
Missing	31 (0.9)	26 (0.9)	5 (0.9)
Physical activity (Cambridge Index)			
Inactive	722 (22.6)	582 (22.1)	140 (25.1)
Moderately inactive	1048 (32.8)	869 (32.9)	179 (32.1)
Moderately active	731 (22.9)	609 (23.1)	122 (21.9)
Active	637 (19.9)	523 (19.8)	114 (20.4)
Missing	60 (1.9)	57 (2.2)	3 (0.5)
Smoking status			
Never	1025 (32.1)	843 (31.9)	182 (32.6)
Former	1374 (43.0)	1129 (42.8)	245 (43.9)
Current	765 (23.9)	640 (24.2)	125 (22.4)
Unknown	34 (1.1)	28 (1.1)	6 (1.1)
Alcohol consumption at recruitment			

Table 3.4 Main characteristics of men included in the analysis, overall and separately in discovery and validation sets			
Non-drinker (<0.1 g/day)	286 (8.9)	235 (8.9)	51 (9.1)
>0.1-3 g/day	432 (13.5)	360 (13.6)	72 (12.9)
>3-12 g/day	730 (22.8)	605 (22.9)	125 (22.4)
>12-24 g/day	644 (20.1)	539 (20.4)	105 (18.8)
>24 g/day	1106 (34.6)	901 (34.1)	205 (36.7)
Anthropometric variables (mean (SD))			
Height (cm)	172.7 (7.0)	172.7 (7.1)	173.0 (6.7)
BMI (kg/m²)	26.9 (3.4)	26.9 (3.4)	26.9 (3.3)
Dietary variables (g/day) (mean (SD))			
Total energy (kcal/day)	2390 (649)	2375 (650)	2440(641)
Dairy	303 (229)	302 (227)	306 (237)
Milk	198 (205)	199 (204)	195 (212)
Cheese	34.5 (35.2)	33.6 (34.1)	38.7 (39.6)
Yogurt	38.9 (70.4)	37.5 (67.2)	45.8 (83.5)
Egg	18.6 (17.9)	18.4 (18.1)	19.5 (16.7)
Fish and fish products	40.9 (41.8)	40.9 (41.8)	41.0 (41.6)
Fish	35.1 (38.3)	35.2 (38.0)	34.8 (39.4)
Lean fish	24.9 (31.8)	25.1 (31.6)	24.2 (33.0)
Fatty fish	12.8 (18.2)	12.8 (18.4)	13.0 (17.5)
Red meat	49.6 (36.6)	49.0 (36.2)	52.5 (38.2)
Processed meat	45.9 (42.7)	45.9 (43.4)	45.9 (38.8)
Poultry	21.9 (21.2)	21.9 (21.0)	21.8 (22.4)
Fats and oils	32.6 (17.4)	32.3 (17.3)	33.9 (17.6)
Butter	5.24 (10.5)	5.45(10.6)	4.26 (9.68)
Margarine	9.74 (14.7)	9.69 (14.5)	9.93 (15.8)
Vegetable oil	16.5 (17.7)	16.1 (17.5)	18.4 (18.5)
Vegetables	190 (129)	191 (130)	186 (128)
Leafy vegetables	30.4 (49.0)	30.0 (49.1)	32.2 (48.6)
Fruiting vegetables	67.6 (56.3)	67.0 (56.1)	70.2 (57.1)
Root vegetables	19.6 (24.2)	20.1 (24.6)	17.5 (22.0)
Fruit	236 (206)	233 (204)	251 (214)
Cereal	257 (139)	253 (134)	273 (161)
Scores for metabolite patterns			
Pattern 1 (geometric mean (SD))	10.2 (1.3)	10.2 (1.3)	10.2 (1.2)
Pattern 2 (geometric mean (SD))	1.98 (0.44)	1.98 (0.44)	1.98 (0.45)
Pattern 3 (geometric mean (SD))	6.13 (0.61)	6.13 (0.61)	6.13 (0.61)

Table 3.4 Main characteristics of men included in the analysis, overall and separately in discovery and validation sets. BMI, body mass index; SD, standard deviation.

The frequency of men from each country, in the overall dataset, and discovery and validation groups, are depicted in Table 3.5. Numbers were broadly consistent between groups, though Italy was more represented in the validation set (25.1% versus 16.9% in the discovery), while the UK was more represented in the discovery set (25.6% versus 16.5% in the validation).

Table 3.5. Total number (n) of men from each country and percentage (%) of total						
	Overall		**Discovery**		**Validation**	
Country	**n**	**%**	**n**	**%**	**n**	**%**
Italy	586	18.3	446	16.9	140	25.1
Spain	696	21.8	572	21.7	124	22.2
UK	767	24.0	675	25.6	92	16.5
Netherlands	185	5.8	168	6.40	17	3.10
Germany	924	28.9	779	29.5	145	26.0
Sweden	20	0.63			20	3.58
Denmark	20	0.63			20	3.58
Total	3198	100	2640	100	558	100

The average fish intake of men from each country represented in this study is depicted in Table 3.6, including in the overall dataset, and in the discovery and validation sets. Spain had the highest average fish consumption overall (84 g/day), while the Netherlands had the lowest (11.4 g/day).

Table 3.6 Mean and SD of total fish intake (g/day) of men from each country						
	Overall		**Discovery**		**Validation**	
Country	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Italy	30.0	22.1	30.1	22.3	29.5	21.5
Spain	84.0	53.5	85.0	52.4	79.2	58.7
UK	35.6	27.5	35.3	26.9	37.8	31.9
Netherlands	11.4	12.5	11.6	12.8	9.75	9.33
Germany	25.2	30.9	25.6	31.9	23.2	24.6
Sweden	43.3	24.8			43.3	24.8
Denmark	53.0	28.9			53.0	28.9

Table 3.6 Mean and SD of total fish intake (g/day) of men from each country. SD, standard deviation; g/day, grams per day.

The association of fasting status with each metabolite pattern across the overall, discovery, and validation datasets is shown in Table 3.7, to assess whether fasting status differed between the discovery and validation sets. Adjusted means were very similar across fasting categories within both the discovery and validation sets, and likelihood ratio tests provided no evidence of heterogeneity. These results indicate that fasting status is unlikely to explain differences between the discovery and validation sets.

Table 3.7 Adjusted means of metabolite pattern scores by fasting status (hours) in the overall, discovery, and validation sets														
		Overall (n=3,198)				**Discovery (n=2,640)**				**Validation (n=558)**			
Pattern	**Fasting status (hours)**	**Adj mean**	**95% CI LL**	**95% CI UL**	**P-het**	**Adj mean**	**95% CI LL**	**95% CI UL**	**P-het**	**Adj mean**	**95% CI LL**	**95% CI UL**	**P-het**
1					0.79				0.71				0.93
	0-3	11.8	11.8	11.9		11.8	11.8	11.9		11.8	11.6	12.0	
	3-6	11.8	11.7	11.9		11.8	11.7	12.0		11.8	11.5	12.0	
	>6	11.8	11.7	11.9		11.8	11.7	11.9		11.8	11.7	12.0	
2					0.30				0.43				0.45
	0-3	6.10	6.07	6.14		6.10	6.07	6.14		6.11	6.02	6.21	
	3-6	6.12	6.07	6.17		6.12	6.06	6.17		6.10	5.96	6.23	
	>6	6.13	6.08	6.17		6.13	6.08	6.18		6.12	6.03	6.20	
3					0.67				0.67				0.98
	0-3	6.00	5.97	6.03		6.00	5.96	6.03		6.01	5.92	6.09	
	3-6	6.01	5.96	6.06		6.00	5.94	6.05		6.11	5.99	6.24	
	>6	5.91	5.87	5.95		5.92	5.87	5.97		5.88	5.80	5.96	
Adjusted means and 95% confidence intervals of metabolite pattern scores by fasting status (hours) in the overall, discovery, and validation sets. Models were adjusted for age, time of day of blood collection, level of education, physical activity, energy intake, alcohol intake, and body mass index. Adj mean, adjusted mean; 95% CI LL, 95% confidence interval lower limit; 95% CI UL, 95% confidence interval upper limit. P-het values are from likelihood ratio tests of heterogeneity across fasting categories within each dataset.													

3.4.2 Correlates of metabolite patterns

Figure 3.1 depicts the betas and 95% confidence intervals for associations between the metabolite patterns and selected potential correlates in the discovery and validation sets. Appendix 3.3 shows the full results for betas, p-values, and Padj values (p-values after adjusting for multiple testing in the discovery set) for the exposure-metabolite pattern associations in the discovery and validation sets. Associations with individual metabolites are shown in Appendix 3.4.

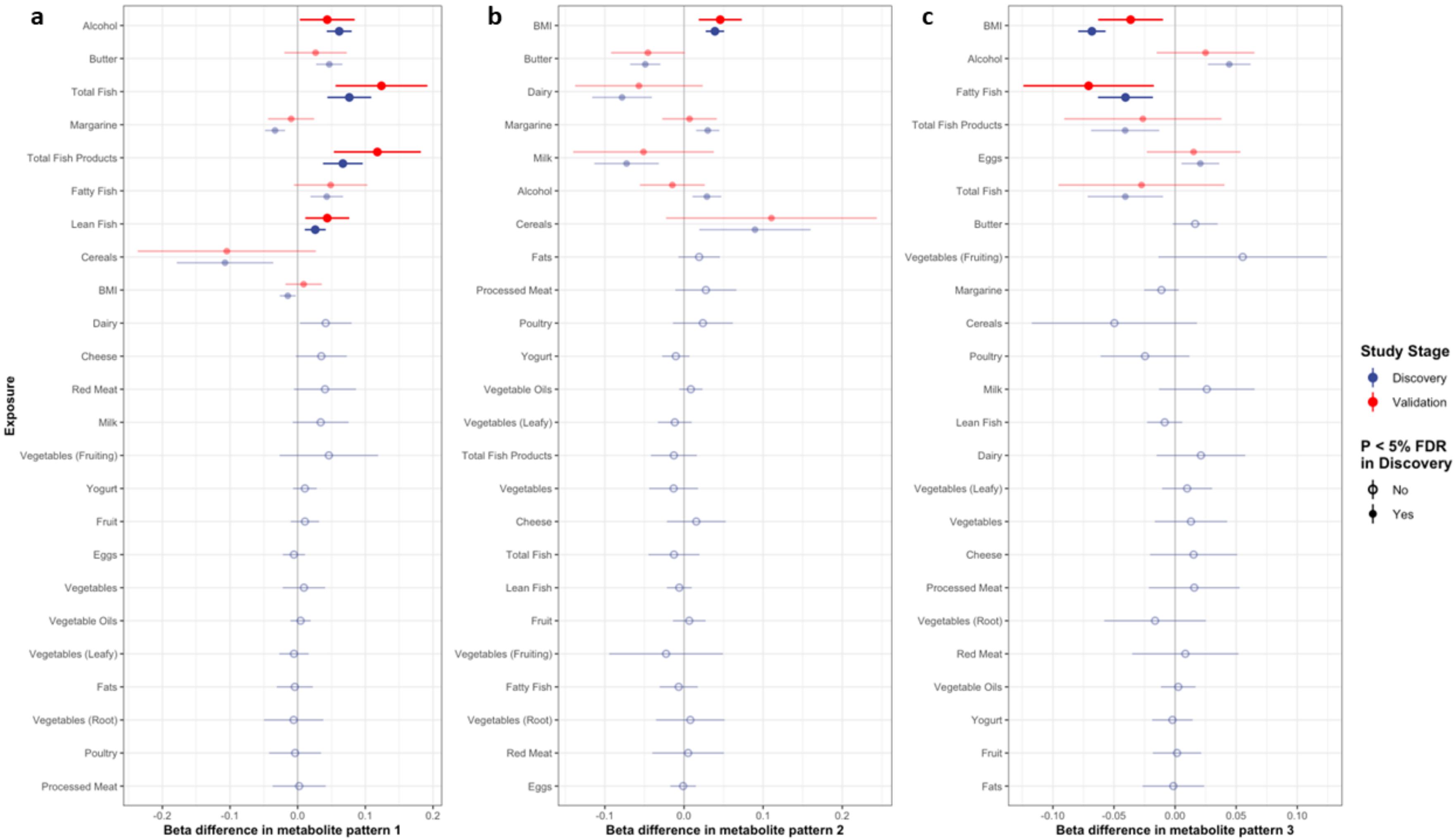


Figure 3.1 Association of dietary and lifestyle factors with metabolite patterns. Estimates presented in bold are for dietary and lifestyle factors with significant Padj associations with at least one metabolite pattern in the discovery dataset (blue) and $P < 0.05$ in the validation dataset (red). Estimates with associations that do not pass the false discovery rate threshold are presented with hollow circles. FDR, false discovery rate.

Of the nine exposures that passed the significance threshold after adjusting for multiple testing in the discovery set, intakes of total fish products, total fish, lean fish, and alcohol all remained statistically significantly ($P < 0.05$) positively associated with pattern 1 in the validation set (Figure 3.1a). In the analysis of individual metabolites contributing to Pattern 1, the associations with the lowest P-values included alcohol with PC aa C32:1, C34:1, and C36:4, and total fish products, total fish, and lean fish with PC aa C42:2 (Figure 3.2; see Appendix 3.4 for full results).

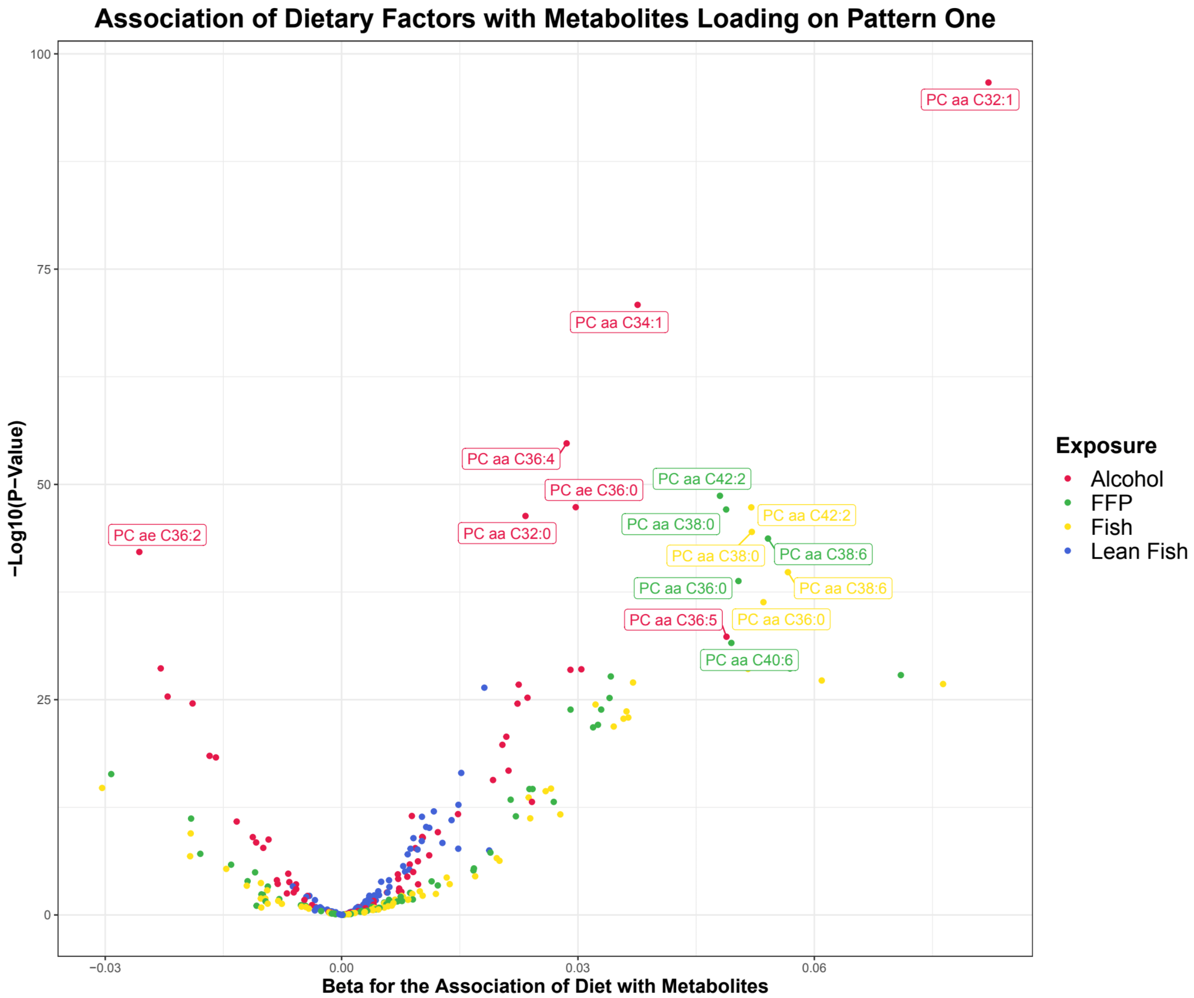


Figure 3.2. Volcano plot depicting the beta and -log10(p-value) of the individual metabolites loading on Pattern 1 associated with an increase in intake of alcohol, total fish products, and its subtypes fish and lean fish. FFP, fish and fish products.

For Pattern 2, seven exposures passed the significance threshold after multiple testing in the discovery set, though only BMI remained positively and statistically significantly associated with the metabolite pattern in the validation set (Figure 3.1b). This relationship appeared to be strongly driven by a strong, positive BMI-glutamate association (Figure 3.3; Appendix 3.4).

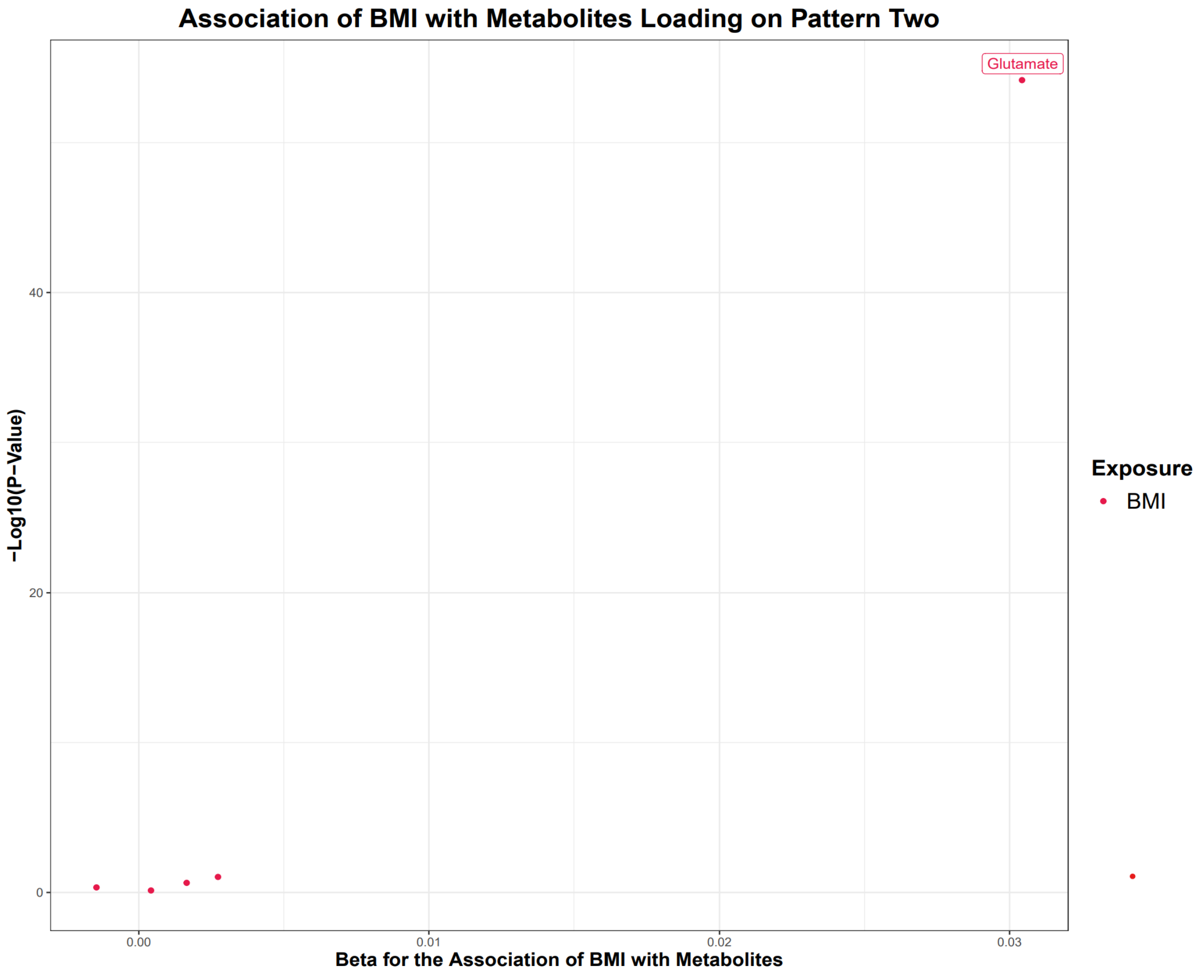


Figure 3.3 Volcano plot depicting the beta and -log10(p-value) of the individual metabolites loading on Pattern 2 associated with an increase in BMI. BMI, body mass index.

Of the six exposures that were significant after multiple testing in the discovery set, fatty fish intake and BMI remained statistically significantly inversely associated with Pattern 3 in the validation set (Figure 3.1c). In the analysis of individual metabolites, BMI was strongly and inversely associated with all eight lyso PCs loading on the metabolite pattern, while fatty fish was significantly inversely associated with lysophosphatidylcholines C16:1, C18:0, C18:1, C20:3, and C20:4 (Figure 3.4; Appendix 3.4).

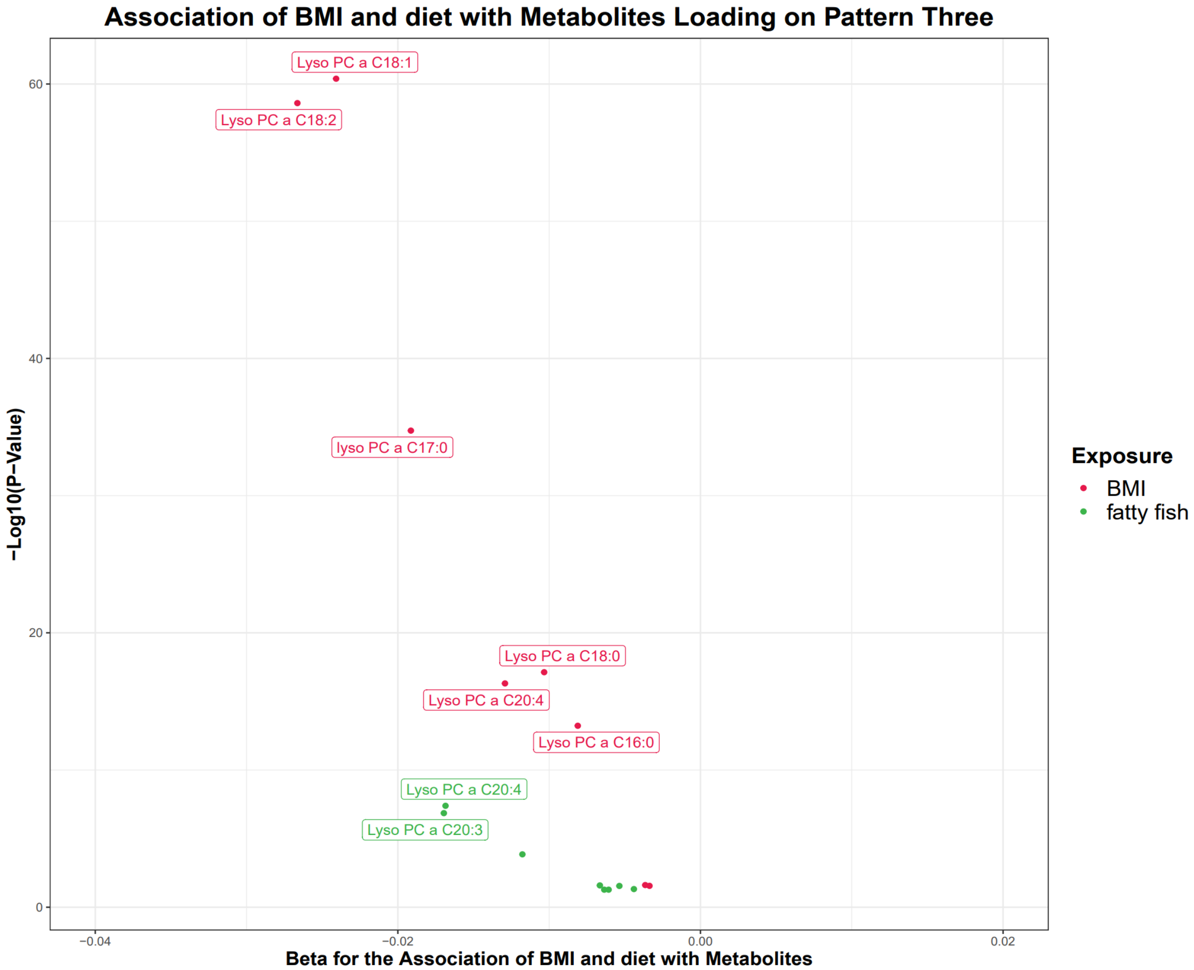


Figure 3.4 Volcano plot depicting the beta and -log10(p-value) of the individual metabolites loading on Pattern 3 associated with an increase in BMI and fatty fish. BMI, body mass index.

3.5 Discussion

This large cross-sectional study identified several dietary and lifestyle correlates of metabolite patterns that were inversely associated with more aggressive prostate cancer subtypes. The intakes of alcohol, total fish products, and its subsets fish and lean fish, were all positively associated with a metabolite pattern with higher concentrations of 31 PC aas, 33 PC aes, and three hydroxysphingomyelins. BMI was positively associated with the second metabolite pattern of two acylcarnitines, glutamate, ornithine, and taurine, and an individual metabolite analysis showed that this was driven by a specific association with glutamate. Finally, BMI and fatty fish intake were inversely associated with scores of a third metabolite pattern of eight lysophosphatidylcholines. No additional associations of dietary variables or BMI with metabolite patterns were validated.

Comparing these results to previous studies is complex; this analysis primarily investigated metabolite patterns rather than individual metabolites. Broadly in line with my results, however, positive associations of some of the metabolites loading on metabolite Pattern 1 (phosphatidylcholines and hydroxysphingomyelins) with alcohol intake have previously been reported in other analyses [79,97–99].

Though there are limited prior studies of associations between fish intake and metabolites, a positive association between fish intake and certain phosphatidylcholines has also been previously reported in an analysis in the TwinsUK cohort [97], and in an intervention study at the University of Otago [100]. This may be partially attributed to fish being a dietary source of choline, which is a requirement for hepatic phosphatidylcholine biosynthesis [101].

The positive association between BMI and the metabolites loading on metabolite pattern 2 (driven largely by glutamate) was consistent with findings in other studies, including the Framingham Heart Study, and the Malmö Diet and Cancer Study [102]. Additionally, a Mendelian randomization analysis suggested that the positive effect of BMI on circulating glutamate may be causal [83]. Previous studies have also demonstrated that glutamate is positively linked to visceral obesity [82,83,102,103].

This study found inverse associations of fatty fish intake and BMI with Pattern 3, which comprised eight lysophosphatidylcholines. To date, there are still limited data available regarding the effects of diet on lysophosphatidylcholine concentrations [104]. However, an eight-week sequential therapy clinical trial for adults with diabetes mellitus found a reduction in levels of lyso PC a C16:1 after consistent fish oil supplementation [104], which may support the current study's findings that fatty fish is inversely associated with a metabolite pattern of eight lyso PCs, including lyso PC a C16:1.

For BMI and metabolite pattern 3, similar to my study, the aforementioned Mendelian randomization analysis on BMI and metabolites also found an inverse effect of BMI on blood levels of lyso PCs C18:1 and C18:2 [83], both of which were included in the pattern.

The current analyses have identified dietary (fish and alcohol) and anthropometric (BMI) correlates of three metabolite patterns that were previously found to be inversely associated with more aggressive prostate cancer subtypes [4]. The implications of these associations are not yet clear; fish [105–113] and alcohol intakes [114–123] are not established risk factors for prostate cancer, while any associations of BMI with prostate cancer risk, of which positive associations with advanced disease and death have previously been reported [124], might be due to differences in timing of detection in men with obesity compared to men with a normal BMI [57,79,125–128]. Furthermore, research is ongoing to determine whether the metabolite pattern-prostate cancer associations previously reported are likely to be causal.

3.5.1 Exposures and prostate cancer risk in previous work

Previous results studies on the association between fish and prostate cancer have been mixed. In a Swedish population-based prospective cohort study, an increased proportion of dietary fish intake was associated with decreased risk of prostate cancer and death from prostate cancer [105]; this is broadly consistent with the results of a United States prospective cohort study, which found a decreased risk of prostate cancer, specifically metastatic, with greater overall fish and marine fatty acid consumption [106]. In contrast, a Japanese prospective cohort study found that higher fish consumption led to elevated risk of overall prostate cancer, [107] while a Californian prospective study also reported a positive association between fish intake and prostate cancer, though this relationship was attenuated when adjusted for fruit and vegetable consumption [108]. Meanwhile, several studies have reported no association between fish intake and risk of prostate cancer [109–113]. Similarly, the association between alcohol intake and prostate cancer risk has been inconclusive [114]. Limited studies have reported a moderate positive association between alcohol intake and overall prostate cancer incidence [121–123]. Within EPIC, no association was observed between alcohol consumption (both current and lifetime) and prostate cancer risk [115]. This was consistent with several epidemiological studies in other cohorts [116–120], though one study did report a weak association between increased alcohol intake and reduced risk of aggressive prostate cancer and prostate cancer death [119].

In a recent EPIC study, higher BMI was associated with increased risk of high grade prostate cancer and prostate cancer death [129]. Other prospective cohort studies that have investigated an association between BMI and clinically relevant prostate cancer have largely found that higher BMI was associated with increased risk of aggressive prostate cancer [125], nonmetastatic high-grade prostate cancer and risk of metastatic or fatal prostate cancer [126], and in a meta-analysis of prospective studies, death from prostate cancer [127]. However, one study found that higher BMI (at age 21) was associated with reduced risk of advanced and lethal prostate cancer [128].

3.6 Strengths and limitations

A major strength of this study is its large sample size owing to the pooling of metabolomics data from four sub-studies within EPIC. Furthermore, using the metabolite patterns as outcome variables accounted for correlations between metabolites [4]. In addition, having access to a wide variety of collected exposure data allowed for the investigation of a range of potential dietary variables and BMI, and adjustment for potential confounding factors. Finally, the discovery/validation approach likely reduced the in-sample bias due to participant overlap between those used to derive treelets and those used to validate the diet- and BMI-metabolite pattern associations.

One limitation of this study is the cross-sectional design, which prevents drawing any definitive conclusions about the temporality and causality of the reported associations. Potential heterogeneity in metabolite concentrations, such as between the four studies nested within the EPIC cohort, was addressed by applying a dedicated pipeline to the data prior to statistical analyses [86], and the analytical protocol used has demonstrated high reproducibility between instruments [57]. To assess dietary intakes, food frequency questionnaires were used in most EPIC centres, which can result in some measurement error due to misreporting of food consumption, recall bias, or errors related to the food composition tables used. Despite this, numerous pilot and cross-sectional studies have supported the reproducibility and validity of the food frequency questionnaire method [79,130]. This study was conducted in a European population, and while there is limited information on racial and ethnic diversity of the participants, it is expected that the participants are primarily of European ancestry. This limited diversity may hamper the generalizability of the current findings to non-European populations. Future research should evaluate associations in different ethnic and racial groups to provide a more generalisable understanding of determinants of the circulating metabolome.

3.7 Conclusion

I report in this large, cross-sectional study in European men that BMI and intakes of total fish products, fish subtypes, and alcohol are associated with metabolite patterns that have been previously linked to a lower risk of aggressive prostate cancer tumour subtypes. In particular, fish and alcohol intake are positively associated with several phosphatidylcholines (including PC aas and PC aes), while BMI strongly associates with the amino acid glutamate. The nature and possible causality of these associations warrants further investigation.

4. An EPIC nested case-control study of 4,387 matched pairs investigating the association of metabolites and aggressive prostate cancer risk

4.1 Introduction

Several studies have investigated circulating metabolites and their associations with more aggressive prostate cancer tumour subtypes [4,7,9,10,12,14]. In the most recent nested case-control study of 3,057 matched pairs nested within EPIC, three metabolite patterns were identified and assessed in relation to prostate cancer risk [4] (outlined in Chapter 3), including Patterns 1 (64 diacyl- and acyl-alkyl phosphatidylcholines and three hydroxysphingomyelins) and 2 (two acylcarntines, glutamate, ornithine, and taurine), which were both found to be associated inversely with advanced and aggressive prostate cancer risk. Furthermore, Pattern 3 (eight lysophosphatidylcholines) was observed to be associated inversely with risk of advanced prostate cancer and prostate cancer death [4]. In line with these results, other cohorts have reported inverse associations of glycerophospholipids [8] and acylcarnitine C18:2 [10] with aggressive prostate cancer risk [10]. While these previous studies have provided important insights, their smaller sample sizes (<580 advanced cases) may limit their statistical power to assess metabolite-prostate cancer associations by tumour characteristic [4,7,9,10,12,14]. Thus, prospective studies with larger sample sizes in more aggressive subgroups could be useful to more robustly assess the association between metabolites and prostate cancer risk.

I also had an *a priori* interest in the amino acids based on previous evidence that amino acids, such as branched-chain and other essential amino acids, and serine, may have an integral role in cancer cell proliferation [12,131–133], and could be a biomarker in distinguishing more aggressive prostate cancer tumour subtypes [133,134].

4.2 Aims

For the current analysis, the previous EPIC dataset was expanded considerably from 3,057 sets [4] to 4,387 matched sets, with extended median follow-up from 9.7 years to 10.8 years, and with 2,477 cases diagnosed more than 10 years after recruitment, compared to 1,391 previously. With this larger dataset, I investigated associations of 148 metabolites and three previously determined metabolite patterns [4] with overall, aggressive, advanced, high grade, and death from prostate cancer using the largest prospective sample size to date, stratified at 10 years of follow-up time.

4.3 Methods

4.3.1 Study Population

The EPIC study is a multi-centre prospective cohort study of >520,000 individuals including 153,400 men, aged mainly between 35 and 70, from 19 centres in eight countries (Denmark, Germany, Greece (not used in this study), Italy, Netherlands, Spain, Sweden and United Kingdom) who were recruited between 1992 and 2000 [81]. Detailed information on diet and lifestyle was collected at recruitment, and 139,600 men provided a blood sample. Men were eligible for the current study if they had blood stored at the central biobank at the International Agency of Research on Cancer, Lyon, France (IARC; centres in Germany, Italy, Netherlands, Spain and United Kingdom), or if recruited in Denmark, samples were stored locally only. Further eligibility criteria were that the date of blood collection was known, and no cancer (except non-melanoma skin cancer) had been diagnosed at the time of blood collection. As a result, data for 4,387 cases and 4,387 matched controls were available for this study.

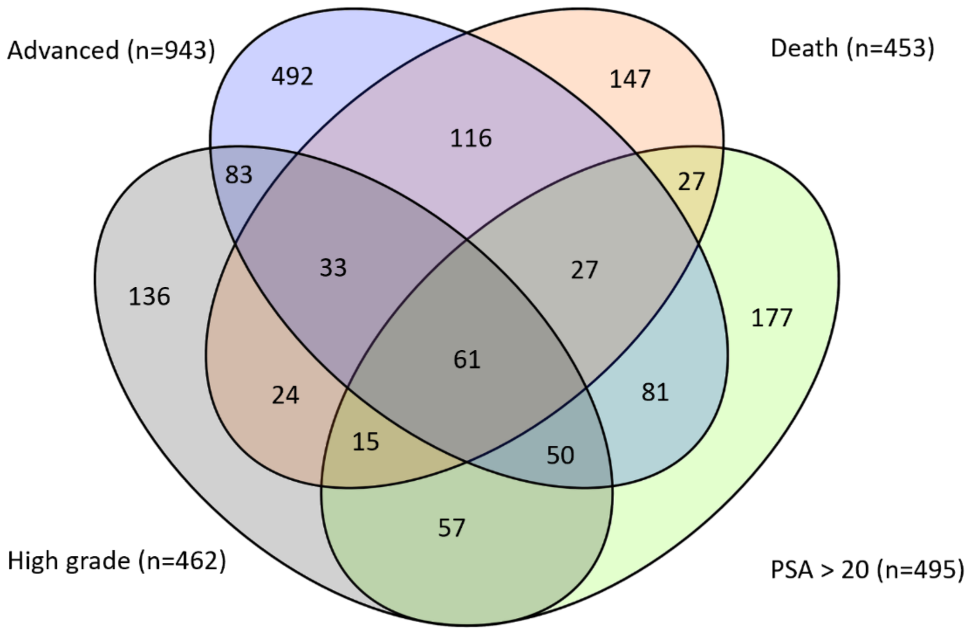
4.3.2 Follow up for cases and controls

In Denmark, Italy, the Netherlands, Spain, and the UK, information on cancer cases, tumour subtypes, and vital status was identified through population cancer registries. In Germany, a combination of methods, such as cancer and pathology registries, and health insurance records, and active follow-up of study subjects were used [81].

Men diagnosed with prostate cancer (defined as code C61 in the 10th revision of the International Statistical Classification of Diseases and Related Health Problems [ICD-10]) after blood collection and before the end of follow-up were categorized as cases [4]. Each case was matched to one control participant, selected randomly from male cohort participants who were alive and free of cancer (except non-melanoma skin cancer) at the time the case was diagnosed. An incidence density sampling procedure was used, so that a control could become a case at a later date, or be a control for multiple cases. Thus, the odds ratio provides an unbiased estimate of the incidence rate ratio, which would have been obtained from the full cohort [135]. Matching criteria included study centre, length of follow-up and age (±6 months), time of day (±1 hour) and fasting status (< 3 hours, 3–6, > 6) at blood collection.

Prostate cancer subtypes were categorised based on the tumour-node-metastasis (TNM) system and histological grade as follows. Advanced (T3–4 and/or N1–3 and/or M1, or coded as advanced, n = 943), high grade (Gleason score 8+ or coded as undifferentiated tumours, n = 462), aggressive (advanced, and/or high grade, and/or preoperative PSA > 20ng/ml, and/or death from prostate cancer as the underlying cause of death (n = 1,495). Overall, 453 men died from prostate cancer as the underlying cause during follow-up. Figure 4.1 depicts the overlap in cases among each tumour subtype (which are included in the broader aggressive definition).

Figure 4.1 Case overlap of tumour subtypes included in the broader aggressive definition (n=1,495)



4.3.3 Blood collection and laboratory analysis/assay

A standardised protocol was followed for blood collection and processing, and fasting was not required (details published elsewhere) [81]. In brief, for participants from Germany, Italy, the Netherlands, Spain and the UK, samples were stored at IARC in plastic straws at −196 °C. In Denmark, blood samples were stored in tubes in local repositories, and kept in nitrogen vapor at −150 °C. Samples were all assayed at IARC in Lyon, France using the AbsoluteIDQ® p180 Kit (Biocrates Life Sciences AG, Innsbruck, Austria), and following the procedure recommended by Biocrates. To quantify metabolites, liquid chromatography mass spectrometry (LC-MS) was applied. All samples were assayed using one LC instrument (Agilent 1290, Santa Clara, CA, USA) coupled with a MS instrument (Triple Quad 4500, AB Sciex, Framingham, MA, USA). Samples from matched case-control sets were assayed in the same analytical batch, along with quality control samples from pooled plasma. Laboratory personnel were blinded to sample category, which includes case, study control, or quality control. A total of 148 metabolites were quantified.

A majority of metabolites (119/148) were measured in all the participants. These were 8 acylcarnitines, 21 amino acids, five biogenic amines, 72 phosphatidylcholines (lysophosphatidylcholines [lyso PC a, n = 8], diacyl phosphatidylcholines [PC aa, n = 31] and acyl-alkyl phosphatidylcholines [PC ae, n = 33]), hexose and 12 sphingomyelins (denoted hydroxysphingomyelins [SM(OH), n = 5] and sphingomyelins [SM, n = 7]). In samples from Denmark, additional data were available for 17 metabolites that had either been excluded in the previous dataset due to not passing quality control thresholds (six acylcarnitines, one amino acid, two biogenic amines, one lyso PC, two PC aas, and two PC aes) or were not previously available (One lyso PC a and two sphingomyelins). However, I report them here for the Danish data as they passed quality control in the Danish dataset.

4.3.4 Statistical analysis

Participants' characteristics at baseline were summarised using frequencies for categorical variables and mean (standard deviation) for continuous variables (Table 4.1). Participant characteristics were also similarly summarised in Table 4.2, categorised by cases with complete stage and grade data, incomplete stage and grade data, and incomplete stage or grade data. To evaluate predictors of missing tumour stage and grade information, multinomial logistic regression was used. The outcome had three categories: complete information (reference), missing both stage and grade, and missing either stage or grade. Exposures of interest included demographic, lifestyle, and recruitment characteristics. Models were adjusted for study centre to account for variation across recruitment countries. Results are presented as odds ratios with 95% confidence intervals, comparing each incomplete category with the reference group of complete information.

4.3.4.1 Individual metabolites

Metabolite measurements were logarithmically transformed for all analyses.

To estimate the risk of prostate cancer per one standard deviation increase in log metabolite concentration, conditional logistic regression was used conditioned on the matching factors. The analysis was further adjusted for exact age at blood collection (continuous), baseline values of body mass index (quartiles; unknown [0.5%]), smoking (never, past, current and unknown [1%]), alcohol intake (<10, 10–19, 20–39 and ≥40 g of alcohol/day; unknown [0.1%]), attained education level (primary, secondary, degree level and unknown [3.3%]) and marital status (married or cohabiting, not married or cohabiting, and unknown [47%]).

The analyses were performed for the full data set (n case = 4,387), and were also run by tumour subtype (advanced, aggressive, high grade). To assess the potential for reverse causation in the associations of metabolites with risk of more aggressive prostate cancer stages, analyses for advanced, aggressive, and high grade prostate cancer, as well as death from prostate cancer, were further stratified by time to diagnosis (≤10/>10 years); this was done to assess the potential for reverse causality in our analysis of the association of metabolites and metabolite patterns with risk of prostate cancer.

For analyses of individual metabolites, I accounted for multiple testing using the false discovery rate (FDR) as defined by Benjamini-Hochberg [96], with a threshold of 0.05. All tests of statistical significance were two-sided. P-values after FDR are referred to as Padj (0.05 was considered as the conventional level for statistical significance).

4.3.4.2 Metabolite Patterns

To compare with the previous EPIC study [4], I projected the three treelet components previously identified in 3,057 matched sets into the new data from Denmark (1,330 matched sets). The treelet components were identified using treelet transform (as described by Gorst-Rasmussen et al.) [4,73,94]. Full details of the individual metabolites contributing to each metabolite pattern can be found in Appendix 3.1. Conditional logistic regression using the treelet components as exposure variables was conducted as described for individual metabolites but without correction for multiple testing (P-value < 0.05).

All analyses were conducted in Stata Statistical Software Package, version 17 (StataCorp, College Station, TX).

4.4 Results

Table 4.1 shows the main characteristics of the participants. On average, participants were 58 years old (SD=6.5 years) at blood collection. There were no marked differences between cases and controls for selected characteristics.

Table 4.1 Characteristics of 4,387 prostate cancer patients and 4,387 controls from EPIC

Characteristic	Cases (4,387)	Controls (4,387)
Age at blood collection, years mean (SD)	57.8 (6.50)	57.7 (6.50)
Height, cm mean (SD)	173.7 (7.15)	173.8 (7.26)
Body Mass Index, kg/m^2 mean (SD)	26.7 (3.38)	26.8 (3.51)
Smoking, *n* (%)		
Never	1433 (33.0)	1334 (30.7)
Former	1782 (41.1)	1813 (41.7)
Current	1123 (25.9)	1198 (27.6)
Alcohol consumption, *n* (%)		
<10 g/day	1582(36.2)	1580(36.0)
10–19 g/day	950(21.7)	958(21.8)
20–40 g/day	961(22.0)	991(22.6)
≥40 g/day	883(20.2)	857(19.5)
Physical activity, *n* (%)		
Inactive	852 (19.7)	887 (20.5)
Moderately inactive	1350 (31.3)	1362 (31.5)
Moderately active	1028 (23.8)	1030 (23.8)
Active	1090 (25.2)	1045 (24.2)
Marital status, *n* (%)		
Married or cohabiting	2051 (88.5)	2063 (88.9)
Not married or cohabiting	267 (11.5)	257 (11.1)
Educational attainment, *n* (%)		
Primary or equivalent	1637 (38.6)	1615 (38.0)

Table 4.1 Characteristics of 4,387 prostate cancer patients and 4,387 controls from EPIC		
Secondary	1489 (35.2)	1529 (36.0)
Degree	1110 (26.2)	1107 (26.0)
Cases only		
Year of diagnosis, median(range)	2006 (1994-2012)	
Time to diagnosis, *n* (%)		
< 5 years	493 (11.5)	
5-10 years	1316 (30.7)	
10-15 years	1904 (44.4)	
≥ 15 years	573 (13.4)	
Data source, *n* (%)		
Schmidt et al. (2020)	3,057 (69.7%)	3,057 (69.7%)
New Danish data	1,330 (30.3%)	1,330 (30.3%)
SD, standard deviation; g/day, grams per day		

Table 4.2 shows the main characteristics of participants with complete stage and grade data, stage and grade missing, and stage or grade missing. Among 4,387 prostate cancer cases, 2,041 (47%) had complete information on both stage and grade, 850 (19%) were missing both, and 1,496 (34%) were missing either stage or grade (Table 4.2). Cases with incomplete data were generally similar to those with complete data in terms of demographic and lifestyle factors, while the most pronounced differences were in year of diagnosis and time since blood collection. Cases with incomplete stage and/or grade were more often diagnosed in later calendar years and had longer follow-up times, reflecting variation in registry completeness across countries. This appeared to be largely driven by the Danish data, which had the largest number of incomplete cases, but extended follow-up.

Table 4.2 Characteristics of cases with complete stage and grade data compared to cases with incomplete stage and/or grade data

Characteristic	Complete stage and grade data (2,041)	Stage and grade missing (850)	Stage or grade missing (1,496)
Age at blood collection, years mean(SD)	57.6 (6.35)	58.2 (6.48)	57.7 (6.71)
Height, cm mean(SD)	173.0 (7.14)	174.4 (7.11)	174.1 (7.13)
Body Mass Index, kg/m^2 mean(SD)	27.2 (3.35)	26.3 (3.25)	26.4 (3.42)
Smoking, *n* (%)			
Never	691 (33.9)	268 (31.5)	474 (31.7)
Former	805 (39.4)	365 (42.9)	612 (40.9)
Current	529 (25.9)	207 (24.4)	387 (25.9)

Table 4.2 Characteristics of cases with complete stage and grade data compared to cases with incomplete stage and/or grade data			
Missing/unknown	54 (2.6)	27 (3.2)	70 (4.7)
Alcohol consumption, *n* (%)			
<10 g/day	743 (36.4)	274 (32.2)	565 (37.8)
10–19 g/day	431 (21.1)	192 (22.6)	327 (21.9)
20–40 g/day	428 (21.0)	214 (25.2)	319 (21.3)
≥40 g/day	433 (21.2)	168 (19.8)	282 (18.9)
Missing/unknown	54 (2.6)	27 (3.2)	70 (4.7)
Physical activity, *n* (%)			
Inactive	399 (19.5)	143 (16.8)	310 (20.7)
Moderately inactive	630 (30.9)	258 (30.4)	462 (30.9)
Moderately active	517 (25.3)	192 (22.6)	319 (21.3)
Active	468 (22.9)	251 (29.5)	371 (24.8)
Missing/unknown	54 (2.6)	27 (3.2)	70 (4.7)
Marital status, *n* (%)			
Married or cohabiting	137 (6.7)	44 (5.2)	86 (5.7)
Not married or cohabiting	1090 (53.4)	284 (33.4)	677 (45.3)
Missing/unknown	54 (2.6)	27 (3.2)	70 (4.7)
Educational attainment, *n* (%)			
Primary or equivalent	843 (41.3)	296 (34.8)	498 (33.3)
Secondary	639 (31.3)	302 (35.5)	548 (36.6)
Degree	505 (24.7)	225 (26.5)	380 (25.4)
Missing/unknown	54 (2.6)	27 (3.2)	70 (4.7)
Year of diagnosis, median (range)	2005 (1995-2011)	2009 (1994-2012)	2007 (1994-2012)
Time to diagnosis, *n* (%)			
< 5 years	289 (14.2)	43 (5.06)	161 (10.8)
5-10 years	824 (40.4)	124 (14.6)	368 (24.6)
10-15 years	769 (37.7)	375 (44.1)	760 (50.8)
≥ 15 years	159 (7.80)	222 (26.1)	192 (12.8)
Country			
Italy	168 (8.2)	96 (11.3)	218 (14.6)
Spain	461 (22.6)	47 (5.5)	141 (9.4)
UK	310 (15.2)	133 (15.6)	340 (22.7)
Holland	88 (4.3)	13 (1.5)	86 (5.7)
Greece	38 (1.9)	27 (3.2)	36 (2.4)
Germany	652 (31.9)	73 (8.6)	130 (8.7)
Denmark	324 (15.9)	461 (54.2)	545 (36.4)
SD, standard deviation; g/day, grams per day			

To further evaluate predictors of missing stage and grade, multinomial logistic regression was performed, adjusted for study centre, with complete data as the reference category (Table 4.3). Consistent with the descriptive results, later year of diagnosis (p-value=9.9E-36) and longer time to diagnosis (p-value=8.3E-35) were strongly associated with incomplete stage/grade information. Associations with age, BMI, height, smoking, alcohol intake, physical activity, level of education, and marital status were weaker and largely not statistically significant.

Table 4.3 Odds ratios and 95% confidence intervals from multinomial logistic regression of predictors of incomplete cancer stage and grade compared with complete stage and grade (reference), adjusted for study centre								
	Complete stage and grade data (n=2,041)	Stage and grade missing (n=850)			Stage or grade missing (n=1,496)			
Characteristic		**OR**	**95% CI LL**	**95% CI UL**	**OR**	**95% CI LL**	**95% CI UL**	**Overall p-value for association**
Age at blood collection	Ref.	1.02	1.00	1.03	1.00	0.99	1.01	0.035
Height	Ref.	0.99	0.97	1.00	1.00	0.99	1.02	0.086
BMI	Ref.	0.97	0.95	1.00	0.98	0.96	1.00	0.086
Smoking status	Ref.	0.97	0.86	1.08	1.01	0.92	1.11	0.71
Alcohol intake (g/day)	Ref.	1.01	0.93	1.09	1.01	0.94	1.07	0.97
Physical activity	Ref.	0.96	0.88	1.05	0.95	0.88	1.01	0.29
Marital status	Ref.	0.95	0.86	1.05	1.06	0.98	1.14	0.042
Level of education	Ref.	0.98	0.92	1.05	0.99	0.94	1.04	0.81
Year of cancer diagnosis	Ref.	1.23	1.19	1.27	1.08	1.06	1.11	9.9E-36
Time to diagnosis	Ref.	1.22	1.18	1.26	1.08	1.06	1.10	8.3E-35
Year of entry	Ref.	0.99	0.91	1.08	1.00	0.94	1.08	0.92

Table 4.3 Odds ratios and 95% confidence intervals from multinomial logistic regression of predictors of incomplete cancer stage and/or grade information compared with complete stage and grade (reference), adjusted for study centre. OR, odds ratio; CI LL, 95% confidence interval lower limit; CI UL, 95% confidence interval upper limit; BMI, body mass index; g/day, grams per day

4.4.1 Individual metabolites

There were no associations of the individual metabolites, including the amino acids, with overall, aggressive, high grade prostate cancer, prostate cancer death, or advanced prostate cancer before stratification by follow-up time, after adjustment for multiple testing (Appendices 4.1 to 4.3, 4.5 to 4.10).

For men with advanced prostate cancer diagnosed at or within 10 years of blood collection, the association with PC aas C42:4 (OR_{1SD} = 0.64, 95% CI 0.50—0.82, P-value = 0.0003, Padj = 0.02), C40:2 (OR_{1SD} = 0.68, 95% CI 0.56—0.85, P-value = 0.0005, Padj = 0.02), C40:3 (OR_{1SD} = 0.72, 95% CI 0.59—0.88, P-value = 0.001, Padj = 0.04), and PC aes C38:2 (OR_{1SD} = 0.70, 95% CI 0.57—0.85, P-value = 0.0003, Padj = 0.02), C40:3 (OR_{1SD} = 0.67, 95% CI 0.53—0.86, P-value = 0.001, Padj = 0.04), and C42:3 (OR_{1SD} = 0.74, 95% CI 0.60—0.89, P-value = 0.002, Padj = 0.05) passed correction for multiple testing (FDR < 0.05) (see Figure 4.2, Appendix 4.4).

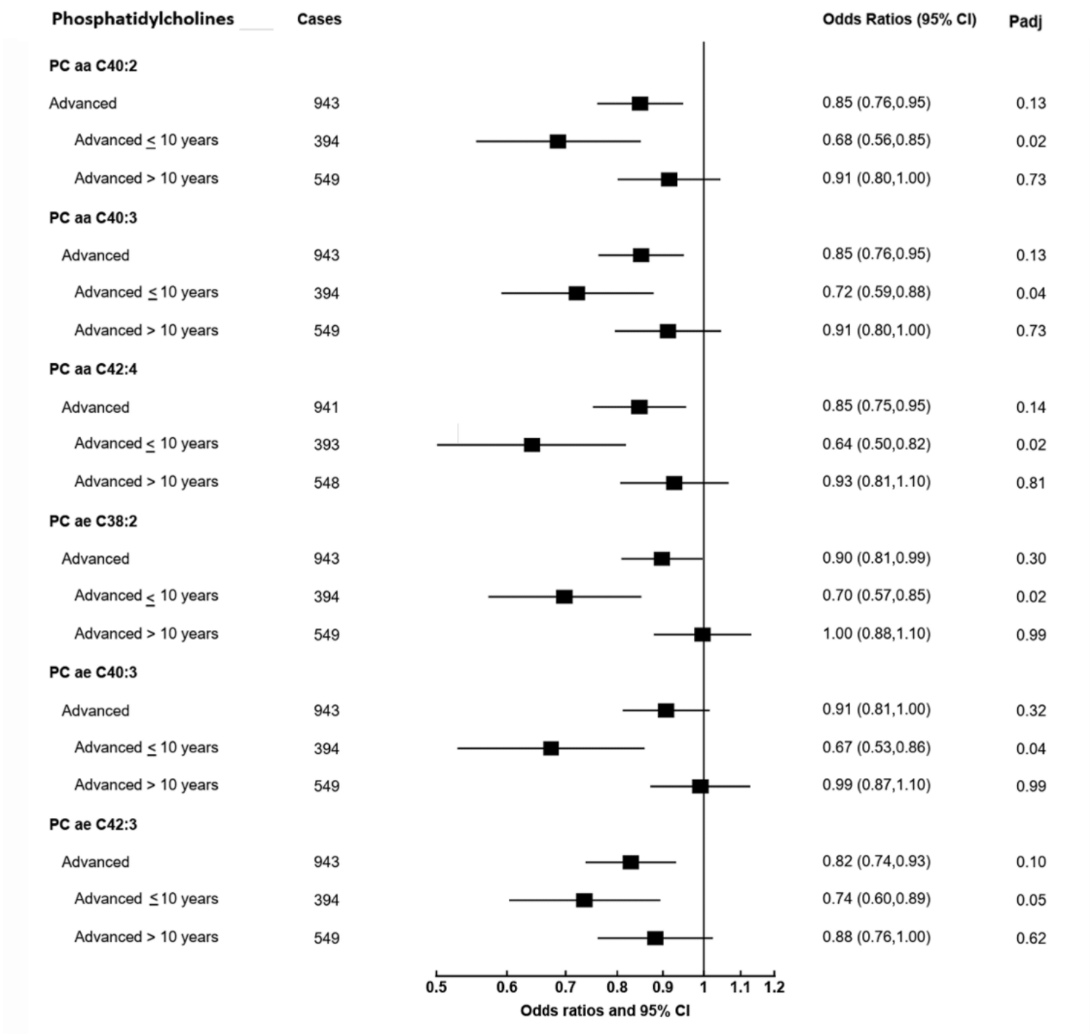


Figure 4.2 Odds ratios of advanced prostate cancer by concentration of six phosphatidylcholines. Odds ratios of advanced prostate cancer associated with a one standard deviation increase in concentration of six PCs. Stage and grade of prostate cancer were categorized using the tumour-node-metastasis (TNM) system and Gleason score, respectively; advanced ($T_{3–4}$ and/or $N_{1–3}$ and/or M_1, or coded as advanced. CI, confidence interval; OR_{1SD}, odds ratio for a one standard deviation increase in metabolite concentration; PC aa, diacyl-phosphatidylcholine; PC ae, acyl-alkyl phosphatidylcholine.

Of the amino acids, eight (serine, threonine, aspartic acid, histidine, glutamine, proline, arginine, and asparagine) were conventionally significantly (P-value < 0.05) associated with at least one prostate cancer outcome (Appendices 4.1 to 4.10). The strongest association was observed for serine, which associated positively with aggressive prostate cancer risk

(OR_{1SD} = 1.10, 95% CI 1.01—1.19, P-value = 0.02, Padj = 0.49), including for diagnoses more than 10 years after blood collection (OR_{1SD} = 1.20, 95% CI 1.07–1.35, P-value = 0.002, Padj = 0.31).

4.4.2 Metabolite Patterns and Prostate Cancer Risk

Figure 4.3 and Appendices 4.11 to 4.23 show associations of metabolite patterns and risk of prostate cancer. There were no associations found for prostate cancer overall, nor with aggressive or high-grade prostate cancer.

Metabolite Pattern 1 (characterised by higher concentrations of diacyl- and acyl-alkyl phosphatidylcholines and three hydroxysphingomyelins) was associated inversely with advanced stage prostate cancer risk (odds ratio for a one standard deviation increase in treelet component score OR_{1SD} = 0.88, 95% CI 0.79–0.99, P-value = 0.03). When stratifying by follow-up time, the association only remained for men diagnosed at or within 10 years of blood collection (OR_{1SD} = 0.80, 95% CI 0.66–0.96, P-value = 0.02), while for men diagnosed later, the odds ratio was 0.95 (95% CI 0.82-1.10, P-value = 0.47).

Pattern 2, which comprised higher acylcarnitines C18:1 and C18:2, glutamate, ornithine and taurine, was also associated inversely with advanced stage prostate cancer risk (OR_{1SD} = 0.80, 95% CI 0.68–0.94, P-value = 0.008). Similar to Pattern 1, the association was stronger in men with 10 or fewer years of follow-up (OR_{1SD} = 0.76, 95% CI 0.59–0.97, P-value = 0.03), than for cases diagnosed later (OR_{1SD} = 0.84, 95% CI 0.67-1.06, P-value = 0.15).

Men who scored higher on Pattern 3, which is characterized by eight lysophosphatidylcholines, had a lower risk of prostate cancer death (OR_{1SD} = 0.82, 95% CI 0.68–0.98, P-value = 0.03). In those diagnosed at or within 10 years, the odds ratio was 0.77 (0.61, 0.99, P-value = 0.04), and in those diagnosed after longer follow-up, the odds ratio was 0.87 (0.65, 1.18, P-value = 0.39).

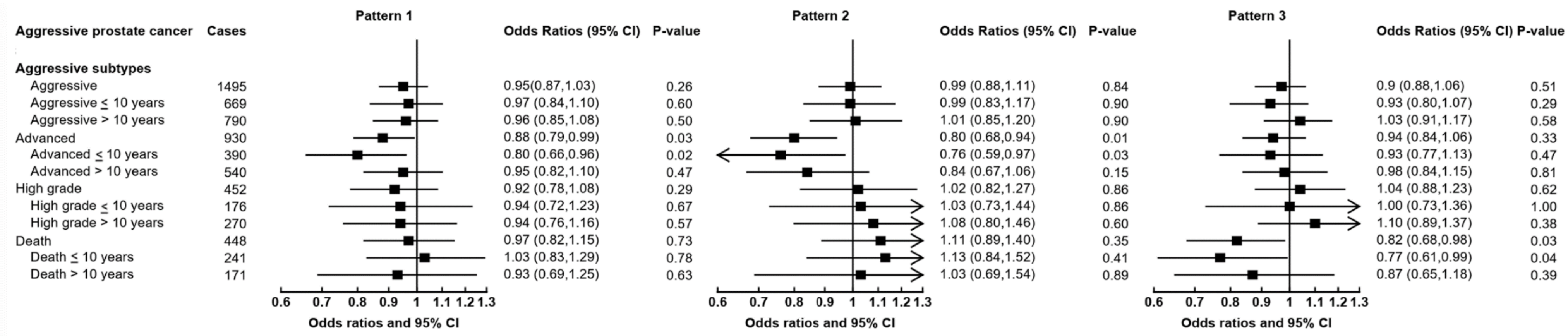


Figure 4.3 Odds ratio of prostate cancer subgroups associated with a one standard deviation increase in metabolite pattern scores. Odds ratio of prostate cancer subtypes associated with a one standard deviation increase in metabolite pattern scores. Stage and grade of prostate cancer were categorized using the tumour-node-metastasis (TNM) system and Gleason score, respectively; advanced ($T_{3–4}$ and/or $N_{1–3}$ and/or M_1, or coded as advanced), high grade (Gleason 8+ or coded as undifferentiated tumours), and aggressive (advanced, high grade, and death combined). Death from prostate cancer during follow-up was defined as prostate cancer listed as the underlying cause of death on the death certificate. CI, confidence interval; OR_{1SD}, odds ratio for a one standard deviation increase in metabolite pattern score.

4.5 Discussion

In this large prospective case-control study of plasma metabolites and prostate cancer risk, there were associations observed between six phosphatidylcholines (three PC aas and three PC aes) and lower risk of advanced prostate cancer, in men diagnosed at or within ten years of blood collection. For the amino acids, serine was associated with an increased risk of aggressive prostate cancer after ten years of follow-up, though this association did not remain after correction for multiple testing. Finally, using three metabolite patterns that were previously identified [4], we observed a lower risk of advanced stage prostate cancer in men with metabolite profiles characterized by higher concentrations of phosphatidylcholines and hydroxysphingomyelins (Pattern 1) and acylcarnitines C18:1 and C18:2, glutamate, ornithine and taurine (Pattern 2), diagnosed at or within ten years of blood collection. Furthermore, higher levels of Pattern 3 (8 lysophosphatidylcholines) were associated with lower risk of death from prostate cancer.

Previously in EPIC, metabolite patterns and prostate cancer risk were investigated, using a subset of the current data (~70%) [4]. In this earlier analysis, it was reported that Patterns 1 and 2 were inversely associated with advanced prostate cancer overall, however in the current extended analysis this association does not persist in men diagnosed more than ten years after blood collection. PC aas C40:2, C40:3, and C42:4, which were associated with men diagnosed with advanced stage disease at or within ten years of blood collection in our individual metabolites analysis, were among several diacyl-phosphatidylcholines that were associated with lower risk of advanced prostate cancer previously in EPIC [4]. However, in this previous study associations were similar in cases by duration of follow-up. Similarly, higher levels of several acyl-alkyl phosphatidylcholines were associated with lower risk of advanced prostate cancer in the previous study, among them PC aes C38:2 and C42:3, which were associated with advanced disease in men with up to 10 years of follow-up in our current individual metabolites analyses [4].

Finally, in line with my current findings, the previous study found Pattern 3 to be associated with lower risk of death from prostate cancer [4]. In my current analyses by follow-up time, however, Pattern 3 was associated with death only in men diagnosed within 10 years of blood collection, though there were a relatively small number of deaths limiting the power of the stratified analyses.

The weaker association in the subgroups with longer follow-up, despite the now larger sample size and greater power to detect an association, suggests that the associations observed may be due to reverse causation, with the presence of pre-clinical cancer altering metabolite levels [14].

Comparing these results to other prospective cohort studies is complex, due to differing analytical platforms and different classifications of aggressive prostate cancer tumour subtypes. Though there are no other studies that applied dimension reduction methods to the full set of metabolites available in this EPIC study, a principal component analysis conducted in the ATBC cohort reported that the first principal components of metabolites in pathways of uracil-containing pyrimidine, dipeptide, glycine/serine/threonine, gamma-glutamyl amino acid, aminosugar, polyunsaturated fatty acid (n3 and n6), and endocannabinoid metabolism were associated positively with risk of prostate cancer death [14]. In a gene set analysis, the ATBC study reported strong inverse associations between lipid and energy metabolite chemical classes and aggressive cancer [12]. Meanwhile, the PLCO cohort did not find any of their 10 calculated principal components to have an association with aggressive prostate cancer [9]. Regarding individual metabolite analyses, in the ATBC study higher glycerophospholipid concentrations were associated with lower risk of tumour stage T3 prostate cancer [14]. In contrast, a nested case-control in the NSHDS found that higher concentrations of three phosphatidylcholines (PC aes C38:3, C38:4, and C40:2) were associated with increased aggressive prostate cancer risk, which included tumour stage T3 [10]. There have been few studies on the metabolites loading on Pattern 2 and their associations with risk of advanced prostate cancer.

The current analyses found that men with higher scores of Pattern 3 (eight lysophosphatidylcholines) had a lower risk of death from prostate cancer. In line with this, in an ATBC study, higher levels of lysolipid 1-linoleoyl-glycerophosphatidylcholine (18:2) (which is comparable to lysophosphatidylcholine C18:2 included in Pattern 3) were also associated with reduced risk of prostate cancer death [14]. In contrast, in the NSHDS, three lysophosphatidylcholines (C17:0, C20:3, and C20:4) were associated positively with aggressive prostate cancer risk [10].

My results by time to diagnosis suggest that the associations with advanced disease are likely due to the influence of prostate cancer on the circulating metabolome, rather than the circulating metabolome affecting risk of prostate cancer development. This is in line with findings of marked changes in the metabolite profile of men with prostate cancer (as reviewed in Kelly et al. (2016) and in Lima et al. (2016)), including in circulating levels of glycerophospholipids and amino acids [136,137].

Many of the measured metabolites have been implicated in post-diagnostic prostate cancer studies. For example, in a review of metabolomics biomarkers of prostate cancer, the author noted that there was evidence of glutamate and taurine in distinguishing malignant from benign prostatic tissue in a magnetic resonance spectroscopy (NMR) study [136]; a more recent NMR study also found increased glutamate levels in tumours with progressively higher Gleason scores [138]. Men diagnosed with prostate cancer have also been found to have altered levels of acylcarnitines, lysophosphatidylcholines, and phosphatidylcholines [137,139].

Following up on an *a priori* interest in the amino acids, none of the amino acids were significantly associated with prostate cancer subtypes after correction for multiple testing. However, serine was positively associated at conventional significance with aggressive disease in men with more than ten years of follow-up; chance cannot be excluded as a possible explanation, and replication and a larger sample size would be needed to further investigate this. Several laboratory studies have suggested that serine may have an integral role in cancer cell proliferation [133], and may be a biomarker in distinguishing more aggressive prostate cancer tumour subtypes [134]. Serine is strongly interconnected with glycine metabolism, which has been implicated in cancer cell growth and as a plausible predictor of metastatic prostate cancer [140–142]. Furthermore, a case-control analysis that pooled five different cancers (prostate included) observed that plasma-free amino acids, including serine, were present in higher concentrations in those who had cancer [143].

4.6 Strengths and Limitations

Due to the large cohort size, this analysis provided greater statistical power than prior prospective studies, including the previous EPIC analysis, especially for analyses by prostate cancer subtype and time to diagnosis. For instance, I was particularly interested in advanced prostate cancer; for men with advanced disease there were 350 additional cases in the current study, including 62 new cases for men diagnosed within ten years of blood collection, and 288 new cases for men with advanced disease after ten years of blood collection [4]. Furthermore, using both individual metabolites and metabolite patterns as exposure variables allowed for the identification of specific metabolite risk associations, as well as investigating the associations with groups of correlated metabolites which might implicate potential metabolic pathways.

Limitations include different sample handling procedures between study centres, as well as the use of non-fasting blood samples, both which may impact measured metabolite levels [144,145]. In order to mitigate these effects, cases and controls were matched on study centre and fasting status [146]. Furthermore, only one blood sample was available per participant, which may lead to attenuation of risk estimates if a single measurement does not represent long-term exposure [147,148]. Finally, the missing stage/grade information in cases with ler follow-up (largely driven by the Danish data) should be taken into consideration in future pooled/collaborative analyses.

Our work on an extended dataset emphasizes the importance of a large sample size for stable risk estimates. These results may have implications for other nested case-control studies of metabolites and cancer types.

4.7 Conclusion

These findings suggest that the results of the previous EPIC study between circulating metabolites and prostate cancer risk may be explained by the effects of pre-clinical disease. I identified several associations between metabolites and risk of advanced prostate cancer among men diagnosed in the first ten years after blood collection. These results may indicate possible reverse causation, and that the metabolite profile of blood starts to change in response to prostate cancer up to a decade before the cancer is detected at an advanced stage. To further understand these observations, it is important to investigate whether comparable metabolic signals for prostate cancer risk can be replicated in an independent cohort with extended follow-up and a different metabolomics platform. Validating these findings in another population with long-term data could provide further evidence as to whether the observed metabolite associations are indicative of reverse causality, or have a more stable role in prostate cancer incidence or mortality.

5. A prospective study of circulating metabolites and prostate cancer risk in UK Biobank

5.1 Introduction

Several prospective studies have suggested that higher circulating concentrations of certain lipids and other energy-containing metabolites may be associated with prostate cancer [4,8,10,14,67]. Most of these studies were conducted using MS assays, including the nested case-control analysis described in Chapter 4 of this thesis. A strength of the MS analytical technique is that it can detect metabolites in low concentrations that NMR may not be able to reliably quantify [149]. Using an NMR metabolomics platform can be complementary to MS studies, as it allows for the identification of additional classes of metabolites. In comparison to the Biocrates MS platform utilised to measure metabolites in Chapter 4, the Nightingale Health Plc NMR-based metabolomics platform (utilised to measure metabolite profiles in UK Biobank participants) provides detailed quantification of a wide range of biomarkers related to lipid and lipoprotein metabolism, as well as amino acids, glycolysis-related metabolites, and markers of inflammation such as albumin and glycoprotein acetyls [63]. The platform includes the size and concentration of various lipoprotein particles (including high-density lipoprotein cholesterol [HDL], low-density lipoprotein cholesterol [LDL] and very low-density lipoprotein cholesterol [VLDL]), and the absolute and relative concentrations of abundant plasma fatty acids (including omega-3, saturated, and monounsaturated) [63]. As a results, NMR-based analyses allow for the investigation of other parts of the metabolome in relation to disease associations, and potentially mechanistically relevant metabolic signatures [63]. One previous NMR-based prospective cohort study reported that men with higher levels of the amino acids valine, tyrosine, histidine, methylhistidine, phenylalanine, and albumin lysyl had a higher risk of developing prostate cancer during follow-up [15], and in the UK Biobank, found lower levels of glucose, citrate, and lipoproteins (HDL and other cholesterols [carrier molecules of phosphatidylcholines]) in prostate cancer cases compared to participants who did not develop prostate cancer [67].

5.2 Aim

The aim of this chapter is to assess the association between 249 circulating metabolites and overall prostate cancer risk (6,575 total incident cases of prostate cancer, compared to 1,436 in the most recent UK Biobank analysis [67]) and risk of prostate cancer death (596 total deaths) in the UK Biobank, using an NMR-based metabolomics platform, and conducting detailed analyses by time from blood draw to diagnosis to assess reverse causality. Based on previous NMR [15] and MS [4,150,151] findings, I had an *a priori* interest in the associations of the 10 available amino acids with disease risk, which will be discussed in this chapter. Finally, to further contextualise key findings, additional cross-sectional analyses were conducted to explore potential pathways, including by investigating dietary fish intake as an exposure, its association with PSA testing and prostate cancer risk, and the association of the omega-3 fatty acid docosahexaenoic acid (DHA) with IGF-1 (which has previously been associated with both fish intake and prostate cancer risk [152]).

5.3 Methods

5.3.1 Study population

The UK Biobank is a large and detailed prospective cohort study, comprising over 500,000 participants aged approximately 40-69 years at recruitment, including 228,863 men. Participants were first identified using National Health Service (NHS) records, and were recruited between 2006 and 2010 in 22 assessment centres throughout the UK. Sociodemographic, lifestyle, health, medical history, and dietary data were collected at recruitment via questionnaires, while collection of blood, urine, and saliva samples also took place at recruitment. Follow-up information was obtained using data-linkage to health records.

All participants provided written informed consent. Full details of the study protocol are available on the UK Biobank website. The NHS North West Multicentre Research Ethics Committee (21/NW/0157) approved the UK Biobank study.

5.3.2 Exposure and outcome assessment

Nightingale Health Plc' s metabolic biomarker platform uses high-throughput NMR to quantify metabolic measures per plasma sample, following EN ISO 13485 certified procedures. Data were available for 249 metabolites (including 168 absolute concentrations and 81 percentage and ratio measures). Metabolites included routine lipids, lipoprotein subclass profiling with lipid concentrations within 14 subclasses, fatty acid composition, and low-molecular weight metabolites such as amino acids, ketone bodies and glycolysis metabolites. The 14 lipoprotein subclasses include triglycerides, phospholipids, total cholesterol, cholesterol esters, and free cholesterol, and total lipid concentration within each subclass [63]. To date, there are data on metabolites for 274,187 individuals, including 126,206 men [64,67,153].

Cancer registration and death data were obtained through record linkage to national registries (NHS Digital for England and Wales using participants' NHS numbers, and NHS Central Register in Scotland using the Community Health Index). Using the 10th revision of the World Health Organization's International Statistical Classification of Diseases (ICD-10) codes, participants were classified as having an event if they had an incident diagnosis of prostate cancer/prostate cancer case first recorded in death certificates (C61).

5.3.3 Follow-up

The person-years of follow-up were calculated from baseline assessment until the first registration of malignant cancer, date of death, loss to or end of follow-up (31 December 2020 for England and 31 December 2016 for Wales, 30 November 2021 for Scotland).

5.3.4 Statistical analysis

Participants' characteristics at baseline were summarised using frequencies for categorical variables and mean (standard deviation) for continuous variables.

Metabolite measurements were logarithmically transformed for all analyses. Prostate cancer and death from prostate cancer were treated as the respective endpoints in Cox proportional hazards regression analyses, with age used as the underlying time variable, to determine the association with prostate cancer risk and death from prostate cancer risk per 1 standard deviation increase in log metabolite concentration, ratio, or percentage. Models were stratified by age group at recruitment (<45, 45 to <50, 50 to <55, 55 to <60, 60 to <65 and ≥65 years), geographical region of baseline assessment centre (ten regions: London, North-West, North-East, Yorkshire and Humber, West Midlands, East Midlands, South-East, South-West, Wales, Scotland), and Townsend deprivation score (quintiles). In addition, models were adjusted for fasting status (< 3 hours, 3-6 hours, > 6 hours), body mass index (<25, 25-25.99, 30-34.99, ≥ 35, missing/unknown), ethnicity (White, Asian or Asian British, Black or Black British, mixed race or other, and missing/unknown), family history of prostate cancer (yes, no, and missing/unknown), height (continuous), qualification obtained (O Level/GCSE/CSE, A-levels or equivalent, professional qualification/national vocational qualification/higher national diploma/higher national certificate/university degree or other professional qualification, missing/unknown), smoking status (never, former, current, missing/unknown), total physical activity ([MET hours per week] <10, 10 to <20, 20 to <40, 40 to <60, ≥ 60), and alcohol intake ([grams/day] non-drinker, >0 to <1 g/day, 1 to <10 g/day, 10 to < 20 g/day , ≥20, missing/unknown). As a sensitivity analysis, the full model was also run for all 249 metabolites with an additional adjustment for (a) Type 2 diabetes (yes/no) and (b) ever had PSA test at recruitment (yes, no, and missing/unknown).

To guide covariate selection (particularly relevant for the glucose–prostate cancer analyses), a directed acyclic graph (DAG; Figure 5.1) was specified. In this DAG, diabetes was considered a potential confounder, as it influences circulating glucose concentrations and has been inversely associated with prostate cancer risk in epidemiological studies [154–156]. Free testosterone is shown in the DAG as a potential pathway variable through which diabetes may influence prostate cancer, since diabetes has been linked to reduced free testosterone concentrations [157], and lower free testosterone levels may in turn influence prostate cancer risk [158]. Based on this framework, primary Cox models estimated the total effect of glucose; a sensitivity model additionally adjusted for diabetes to evaluate potential confounding.

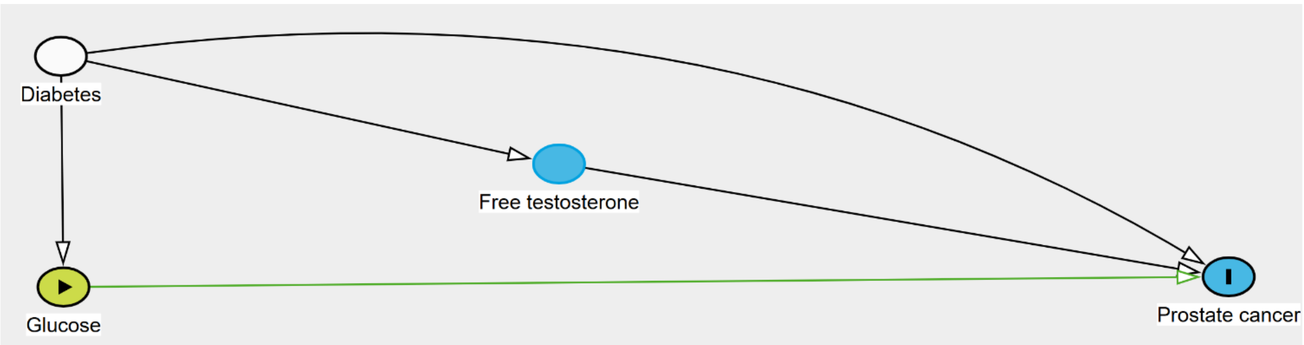


Figure 5.1 Directed acyclic graph for the glucose–prostate cancer analysis. Diabetes is depicted as a potential confounder, influencing both glucose and prostate cancer risk. Diabetes is depicted as a potential confounder, influencing both glucose and prostate cancer risk. Free testosterone is shown as a potential pathway variable through which diabetes may influence prostate cancer. The primary model estimated the total effect of glucose, while a sensitivity model additionally adjusted for diabetes to account for confounding.

Analyses were further stratified by time to diagnosis (overall prostate cancer risk: <3 years, 3-7 years, >7 years (median follow-up 12.5 years); risk of prostate cancer death: ≤ 5 years, > 5 years (median follow-up 5.04 years); this was done to the potential for reverse causality in the associations of metabolites with prostate cancer risk and mortality.

Multiple testing was accounted for using the FDR as defined by Benjamini–Hochberg, with a threshold of 0.05 [96]. All tests of statistical significance were two-sided. P-values after correction for FDR are referred to as Padj (0.05 was considered as the conventional level for statistical significance).

Because metabolite levels fluctuate within individuals over time and single baseline measurements are subject to measurement error, hazard ratios based only on baseline values may underestimate the true associations with prostate cancer risk. [102,159]. To address this, regression dilution bias was corrected using repeat metabolite measurements available from a subsample of 7,438 men who provided a second blood sample, collected on average 4.3 years after recruitment. Regression dilution ratios (RDR) were then applied to the log hazard ratios from Cox models to obtain corrected estimates. This method was applied to the three metabolites most strongly associated with prostate cancer death (glycoprotein acetyls, albumin, and citrate).

For the cross-sectional analyses, linear regression was used to determine the associations of DHA and percentage DHA to total fatty acids with IGF-1, and associations of fish intake with mean DHA and mean percentage DHA to total fatty acids. All models were adjusted for age, fasting status, body mass index, height, alcohol intake, ethnicity, level of education, smoking status, and weekly physical activity. Multivariable logistic regression was used to estimate odds ratios and 95% confidence intervals for the association between oily and non-oily fish intake and ever having had a PSA test at baseline. Models were adjusted for the same covariates as above, and additionally family history of prostate cancer. To determine the association of oily fish and non oily fish intakes with prostate cancer risk, prostate cancer was treated as the endpoint in Cox proportional hazards regression analyses, with age used as the underlying time variable, and the same stratifications and adjustments as the main prospective analyses.

To further assess whether glucose and free testosterone may act in related pathways, a cross-sectional analysis of glucose in relation to free testosterone was conducted. In this analysis, linear regression models were used to estimate geometric mean free testosterone concentrations across fifths of glucose, adjusted for age, fasting status, body mass index, height, alcohol intake, ethnicity, level of education, smoking status, and physical activity.

5.4 Results

5.4.1 Participant Characteristics

Participant characteristics are shown in in Table 5.1. While the baseline characteristics were generally comparable between all participants and those who developed prostate cancer, some characteristics differed. Notably, the mean age at blood collection was 56.8 years for the entire population, and 61.5 years for men who developed prostate cancer. A higher percentage of men who later developed prostate cancer reported having ever had a PSA test at recruitment (43.4% versus 28.7% for all participants) and had a family history of prostate cancer (10.2% versus 6.8%). There was also some difference by Townsend Deprivation Index; the most affluent group made up 20.6% of the overall population, but 24.1% of those who developed prostate cancer were from the most affluent group. For both the overall cohort and men who developed prostate cancer, most participants were of White ethnicity (approximately 95%).

Table 5.1 Characteristics of 126,206 men free from prostate cancer and 6,575 men who developed prostate cancer in the UK Biobank

Characteristic	All participants (126,206)	Participants who developed prostate cancer (6,575)
Age at blood collection, mean years (SD)	56.8 (8.19)	61.5 (5.78)
Height, mean cm (SD)	175.6 (6.82)	175.1 (6.60)
Age at prostate cancer diagnosis, mean years (SD)	N/A	67.0 (7.02)
Age at blood collection, n (%)		
<45 years	13,116 (10.39%)	64 (0.97%)
45- <50 years	16,175 (12.82%)	232 (3.53%)
50- <55 years	18,143 (14.38%)	543 (8.26%)

Table 5.1 Characteristics of 126,206 men free from prostate cancer and 6,575 men who developed prostate cancer in the UK Biobank

55- <60 years	22,034 (17.46%)	1,136 (17.28%)
60- <65 years	30,637 (24.28%)	2,310 (35.13%)
>65 years	26,101 (20.68%)	2,290 (34.83%)
Townsend Deprivation Index (range 1-5; most to least affluent), n (%)		
1 (Most affluent)	25992 (20.59%)	1585 (24.11%)
2	25617 (20.30%)	1428 (21.72%)
3	25178 (19.95%)	1406 (21.38%)
4	24733 (19.60%)	1173 (17.84%)
5 (Most deprived)	24528 (19.43%)	976 (14.84%)
Missing/unknown	158 (0.13%)	7 (0.11%)
Ethnicity, n (%)		
White	119260 (94.50%)	6306 (95.91%)
Mixed race or other	1617 (1.28%)	56 (0.85%)
Asian or Asian British	3019 (2.39%)	63 (0.96%)
Black or Black British	1639 (1.30%)	119 (1.81%)
Missing/unknown	671 (0.53%)	31 (0.47%)
Alcohol consumption, n (%)		
<1 g/day	8493 (6.73%)	419 (6.37%)
1-10 g/day	27362 (21.68%)	1429 (21.73%)
10-20 g/day	26720 (21.17%)	1507 (22.92%)
>20 g/day	55023 (43.60%)	2867 (43.60%)
Non-drinkers	7964 (6.31%)	330 (5.02%)
Missing/unknown	644 (0.51%)	23 (0.35%)
Physical activity		
< 10 METhrs/wk	25367 (20.10%)	1256 (19.10%)
10 - <20 METhrs/wk	20127 (15.95%)	1068 (16.24%)
20 - <40 METhrs/wk	24778 (19.63%)	1346 (20.47%)
40 - <60 METhrs/wk	12663 (10.03%)	655 (9.96%)
≥ 60 METhrs/wk	20141 (15.96%)	1061 (16.14%)

Table 5.1 Characteristics of 126,206 men free from prostate cancer and 6,575 men who developed prostate cancer in the UK Biobank		
Missing/Unknown/PNTS	23130 (18.33%)	1189 (18.08%)
Level of education		
O level/GCSE/CSE	17254 (13.67%)	794 (12.08%)
A levels	6307 (5.00%)	298 (4.53%)
Professional qualification/NVQ/HND/HNC/Degree or other professional qualification	78752 (62.40%)	4062 (61.78%)
Missing/unknown	23893 (18.93%)	1421 (21.61%)
Family history of prostate cancer		
Yes	8517 (6.75%)	669 (10.17%)
No	117689 (93.25%)	5906 (89.83%)
Ever had a PSA test		
Yes	36210 (28.69%)	2851 (43.36%)
No	83033 (65.79%)	3412 (51.89%)
Missing/unknown	6963 (5.52%)	312 (4.75%)
Smoking status, n (%)		
Never	61322 (48.59%)	3079 (46.83%)
Former	48614 (38.52%)	2865 (43.57%)
Current	15641 (12.39%)	597 (9.08%)
Missing/unknown	629 (0.50%)	34 (0.52%)
BMI, n (%)		
<25 kg/m2	30860 (24.45%)	1627 (24.75%)
25- <30 kg/m2	62276 (49.34%)	3353 (51.00%)
30- <35 kg/m2	25279 (20.03%)	1288 (19.59%)
<35 kg/m2	7331 (5.81%)	289 (4.40%)
Missing/unknown	460 (0.36%)	18 (0.27%)
Fasting status		
< 3 hours	32116 (25.45%)	1553 (23.62%)
3-6 hours	85728 (67.93%)	4679 (71.16%)
> 6 hours	8362 (6.63%)	343 (5.22%)

Table 5.1 Characteristics of 126,206 men free from prostate cancer and 6,575 men who developed prostate cancer in the UK Biobank		
Diabetes		
Yes	8882 (7.04%)	416 (6.33%)
No	116720 (92.48%)	6135 (93.31%)
Missing/unknown/PNTS	604 (0.48%)	24 (0.37%)

Table 5.1 Characteristics of 126,206 men free from prostate cancer and 6,575 men who developed prostate cancer in the UK Biobank. SD, standard deviation; g/day, grams per day; MET, metabolic equivalent; GCSE, General Certificate of Secondary Education; CSE, Certificate of Secondary Education; NVQ, national vocational qualification, HNC, higher national certificate; PSA, prostate-specific antigen; BMI, body mass index; PNTS, prefer not to say.

5.4.2 Metabolites and prostate cancer risk

The associations of 249 metabolite concentrations, ratios, and percentages with prostate cancer risk are depicted in Figure 5.2. Seven metabolic features were associated with prostate cancer risk after correction for multiple testing (Figure 5.3, Appendix 5.1). The positive associations included docosahexaenoic acid (DHA) to total fatty acids percentage (HR per one standard deviation = 1.05, 95% CI 1.03-1.08, P-value = 9.3E-05, Padj = 0.023), degree of unsaturation (HR = 1.05, 95% CI 1.02-1.07, P-value = 4.1E-04, Padj = 0.034), phospholipids to total lipids in small low-density lipoprotein (LDL) percentage (HR = 1.06, 95% CI 1.02-1.09, P-value = 5.6E-04, Padj = 0.035), ratio of polyunsaturated fatty acids (PUFAs) to monounsaturated fatty acids (MUFAs) (HR = 1.05, 95% CI 1.02-1.08, P-value = 7.30E-04, Padj = 0.036), and PUFAs to total fatty acids percentage (HR = 1.05, 95% CI 1.02-1.08, P-value = 1.1E-03, Padj = 0.045). Two metabolite features were inversely associated with prostate cancer risk after multiple testing (Figure 5.3): glucose (HR = 0.95, 95% CI 0.93-0.98, P-value = 1.8E-04, Padj = 0.023) and the fatty acid feature monounsaturated fatty acids (MUFAs) to total fatty acids percentage (HR = 0.96, 95% CI 0.93-0.98, P-value = 1.3E-03, Padj = 0.045. Sensitivity analyses including additional adjustment for PSA testing were generally similar, and can be seen in Appendix 5.2.

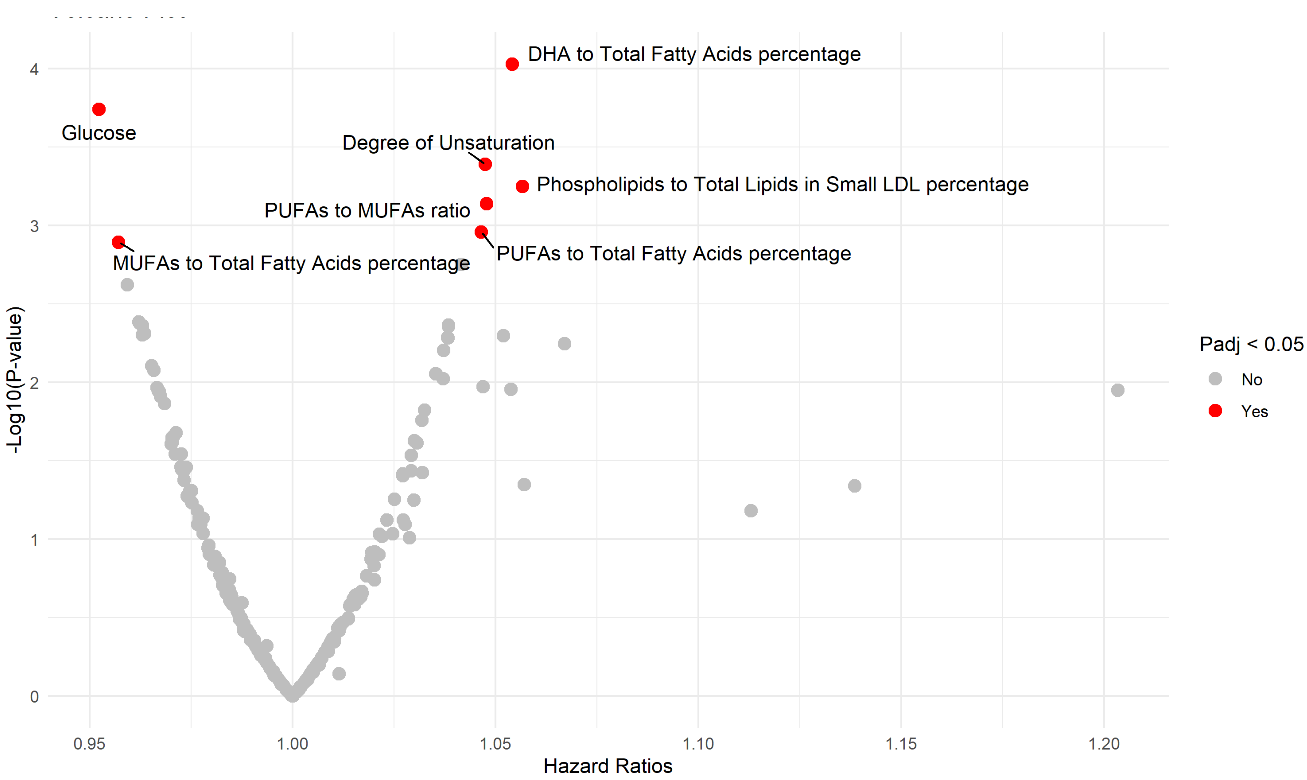


Figure 5.2 Volcano plot depicting the hazard ratio and -log10(p-value) of prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of 249 metabolic features. Metabolite features that associated with overall prostate cancer risk after correction for multiple testing (Padj < 0.05) are labelled and in red. LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol; MUFA, monounsaturated fatty acid; PUFA, polyunsaturated fatty acid; DHA, docosahexaenoic acid.

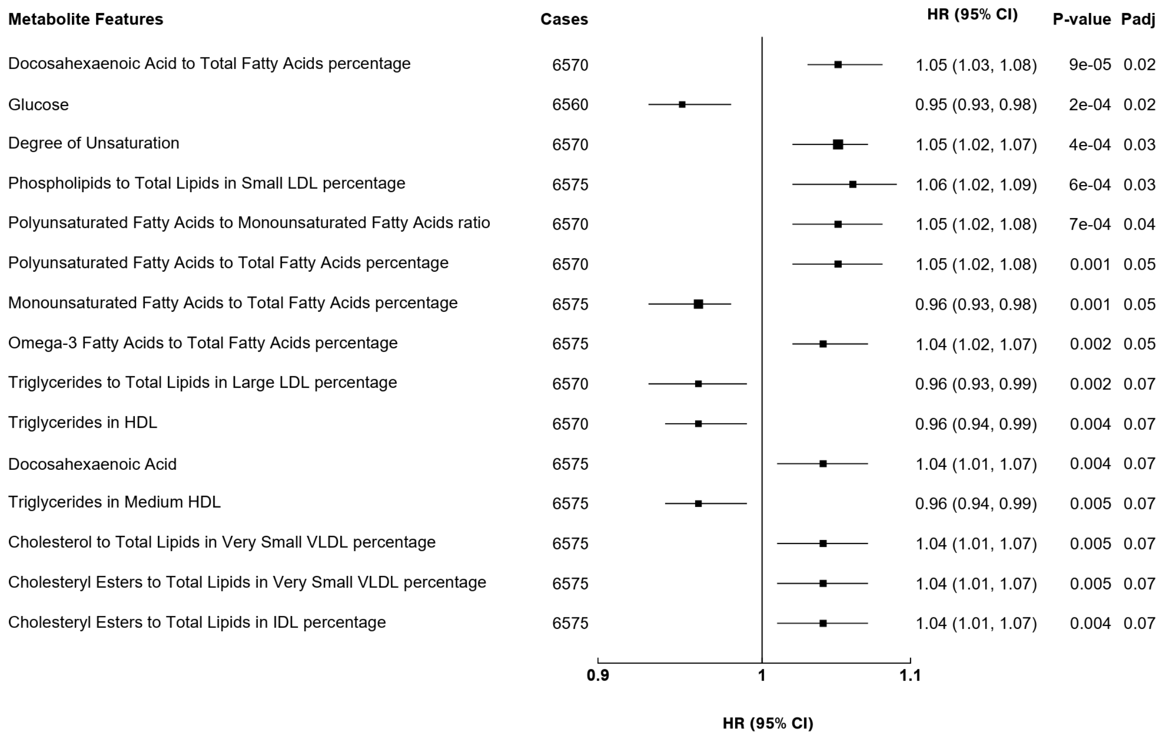


Figure 5.3. Hazard ratios (95% confidence interval) for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of 15 most significant metabolites or metabolic features. HR, hazard ratio; CI, confidence interval; LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol; IDL, intermediate-density lipoprotein cholesterol. Padj refers to P-value adjusted for multiple testing using the Benjamini–Hochberg false discovery rate (FDR) method.

In analyses stratified by time from blood draw to diagnosis (Table 5.2 and Appendix 5.3), the association of DHA to total fatty acids percentage with risk was generally stable across follow-up time group: < 3 years (HR = 1.03, 95% CI 0.97–1.09, P-value = 0.32, Padj = 0.85); > 7 years (HR = 1.05, 95% CI 1.01–1.09, P-value = 0.012, Padj = 0.44)

In contrast, the association of phospholipids to total lipids in small LDL percentage with disease risk appeared mainly driven by the < 3 years since blood collection subgroup (HR = 1.12, 95% CI 1.04–1.20, P-value = 0.0013, Padj = 0.034) in comparison to > 7 years (HR = 1.02, 95% CI 0.97–1.07, P-value = 0.39, Padj = 0.77).

Table 5.2. Hazard ratios and 95% CIs of top seven most significantly associated metabolite features with prostate cancer risk, stratified by time since blood collection																		
	< 3 years						3-7 years						>7 years					
Metabolite features	**Cases**	**HR**	**LL**	**UL**	**P-Value**	**Padj**	**Cases**	**HR**	**LL**	**UL**	**P-Value**	**Padj**	**Cases**	**HR**	**LL**	**UL**	**P-Value**	**Padj**
DHA to Total FA %	1307	1.03	0.97	1.09	0.3	0.9	2159	1.08	1.03	1.13	0.002	0.3	3104	1.05	1.01	1.09	0.01	0.4
Glucose	1303	0.95	0.90	1.01	0.1	0.5	2157	0.95	0.91	0.99	0.02	0.3	3100	0.96	0.92	0.99	0.02	0.4
Degree of Unsaturation	1307	1.02	0.96	1.08	0.5	1.0	2159	1.07	1.02	1.12	0.003	0.3	3104	1.04	1.01	1.08	0.02	0.4
PLs to Total Lipids in Small LDL %	1309	1.12	1.04	1.20	0.001	0.03	2159	1.07	1.01	1.13	0.01	0.3	3107	1.02	0.97	1.07	0.4	0.8
PUFAs to MUFAs ratio	1307	1.03	0.97	1.10	0.3	0.8	2159	1.06	1.01	1.11	0.02	0.3	3104	1.05	1.01	1.09	0.02	0.4
PUFAs to Total FA %	1307	1.04	0.97	1.10	0.3	0.8	2159	1.06	1.01	1.11	0.03	0.3	3104	1.04	1.00	1.09	0.03	0.5
MUFAs to Total FA %	1307	0.97	0.91	1.03	0.3	0.9	2159	0.95	0.90	0.99	0.02	0.3	3104	0.96	0.92	1.00	0.03	0.5

Table 5.2 Hazard ratios and 95% confidence interval for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of 7 most significant metabolites or metabolic features, stratified by time since blood collection. HR, hazard ratio; CI, confidence interval; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit; DHA, docosahexaenoic acid; FA, fatty acids; %, percentage; PL, phospholipid; LDL, low-density lipoprotein cholesterol; PUFA, polyunsaturated fatty acid; MUFA, monounsaturated fatty acid. Padj refers to P-value adjusted for multiple testing using the Benjamini–Hochberg false discovery rate (FDR) method.

In the sensitivity analyses with additional adjustment for diabetes (hazard ratios for all 249 metabolite concentrations ratios, and percentages shown in Figure 5.4), the associations of DHA to total fatty acids percentage (HR = 1.06, 95% CI 1.03-1.09, P-value = 4.9E-05, Padj = 0.012) and phospholipids to total lipids in small LDL percentage (HR = 1.07, 95% CI 1.03-1.10, P-value = 1.4E-04, Padj = 0.016) with risk of prostate cancer remained significant after FDR correction (full results shown in Appendix 5.4, results stratified by follow-up time in Appendix 5.5).

The glucose association (HR = 0.97, 95% CI 0.94-0.99, P-value = 0.018, Padj = 0.18) was attenuated after adjusting for diabetes, while the associations of degree of unsaturation (HR = 1.04, 95% CI 1.02-1.07, P-value = 8.5E-04, Padj = 0.95), ratio of PUFAs to MUFAs (HR = 1.04, 95% CI 1.02-1.07, P-value = 2.0E-03, Padj = 0.091), PUFAs to total fatty acids percentage (HR = 1.04, 95% CI 1.02-1.07, P-value = 2.2E-03, Padj = 0.091), and MUFAs to total fatty acids percentage (HR = 0.96, 95% CI 0.94-0.99, P-value = 4.3E-03, Padj = 0.10) with prostate cancer risk were slightly weaker after adjustment for diabetes (Figure 5.5).

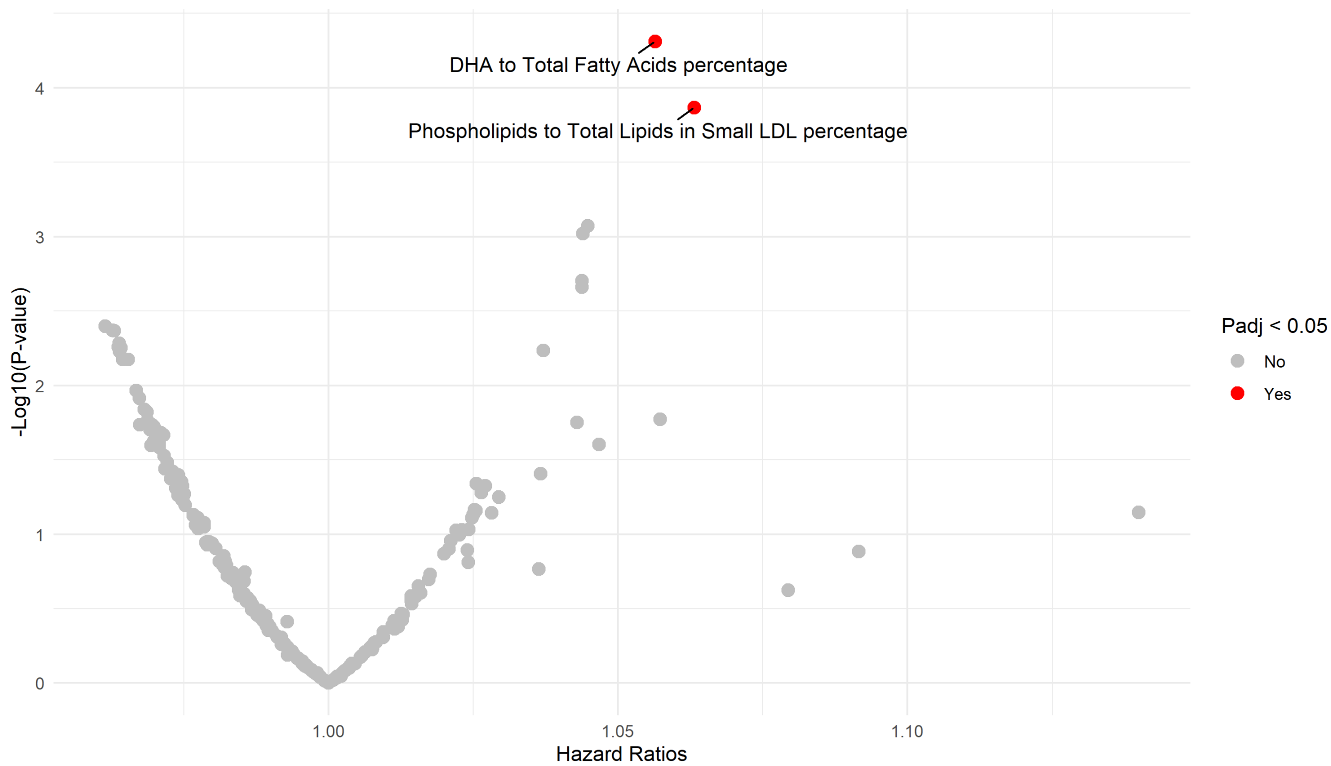


Figure 5.4 Volcano plot depicting the hazard ratios (95% confidence interval) and -log10(p-value) for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of 249 metabolite features, with additional adjustment for diabetes. Metabolite features that associated with overall prostate cancer risk after correction for multiple testing (Padj < 0.05) are labelled and in red. LDL, low-density lipoprotein cholesterol; DHA, docosahexaenoic acid.

Metabolite Features	Cases	HR (95% CI)	P-value	Padj
Docosahexaenoic Acid to Total Fatty Acids percentage	6570	1.06 (1.03, 1.08)	5e-05	0.01
Phospholipids to Total Lipids in Small LDL percentage	6575	1.06 (1.03, 1.10)	1e-04	0.02
Degree of Unsaturation	6570	1.04 (1.02, 1.07)	8e-04	0.06
Omega-3 Fatty Acids to Total Fatty Acids percentage	6570	1.04 (1.02, 1.07)	0.001	0.06
Polyunsaturated Fatty Acids to Total Fatty Acids percentage	6570	1.04 (1.02, 1.07)	0.002	0.09
Polyunsaturated Fatty Acids to Monounsaturated Fatty Acids ratio	6570	1.04 (1.02, 1.07)	0.002	0.09
Triglycerides in LDL	6575	0.96 (0.94, 0.99)	0.006	0.1
Triglycerides in HDL	6575	0.96 (0.94, 0.99)	0.004	0.1
Monounsaturated Fatty Acids	6570	0.96 (0.94, 0.99)	0.004	0.1
Docosahexaenoic Acid	6570	1.04 (1.01, 1.06)	0.006	0.1
Monounsaturated Fatty Acids to Total Fatty Acids percentage	6570	0.96 (0.94, 0.99)	0.004	0.1
Omega-6 Fatty Acids to Omega-3 Fatty Acids ratio	6569	0.97 (0.94, 0.99)	0.007	0.1
Triglycerides in Large LDL	6575	0.96 (0.94, 0.99)	0.006	0.1
Triglycerides in Medium LDL	6575	0.96 (0.94, 0.99)	0.007	0.1
Triglycerides in Large HDL	6575	0.96 (0.94, 0.99)	0.006	0.1

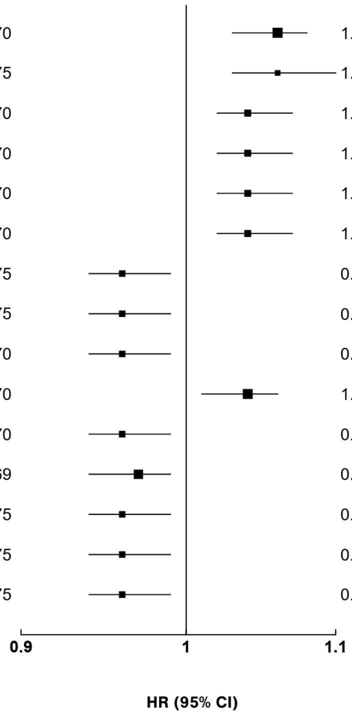


Figure 5.5 Hazard ratios (95% confidence interval) for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of 15 most significant metabolite features, after additional adjustment for diabetes. HR, hazard ratios; CI, confidence interval; LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol. Padj refers to P-value adjusted for multiple testing using the Benjamini–Hochberg false discovery rate (FDR) method.

5.4.3 Metabolite features and risk of death from prostate cancer

The associations of 249 metabolite concentrations, ratios, and percentages with prostate cancer death are depicted in Figure 5.6 (full results in Appendix 5.6). After multiple testing, three metabolites were associated with risk of death from prostate cancer; glycoprotein acetyls (HR = 1.25, 95% CI 1.14-1.4, P-value = 4.9E-07 , Padj = 1.2E-04) and citrate (HR = 1.18, 95% CI 1.08-1.28, P-value = 1.6E-04, Padj = 0.013) were positively associated with risk of prostate cancer death while higher concentrations of albumin were associated with lower risk of prostate cancer death (HR = 0.86, 95% CI 0.79-0.93, P-value = 1.1E-04, Padj = 0.013) (Figure 5.7).

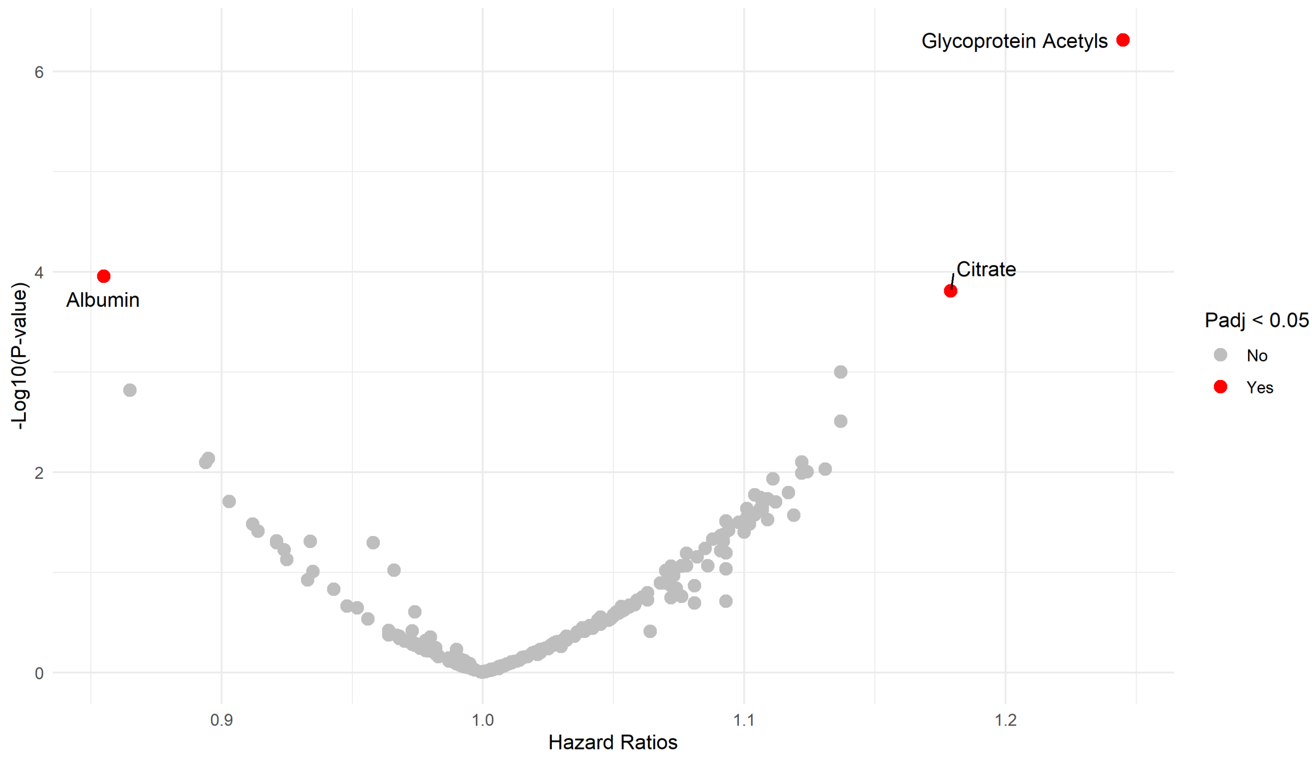


Figure 5.6 Volcano plot depicting the hazard ratios (95% confidence interval) and -log10(p-value) for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of 249 metabolites. Metabolite features that associated with risk of prostate cancer death after correction for multiple testing (Padj < 0.05) are labelled and in red.

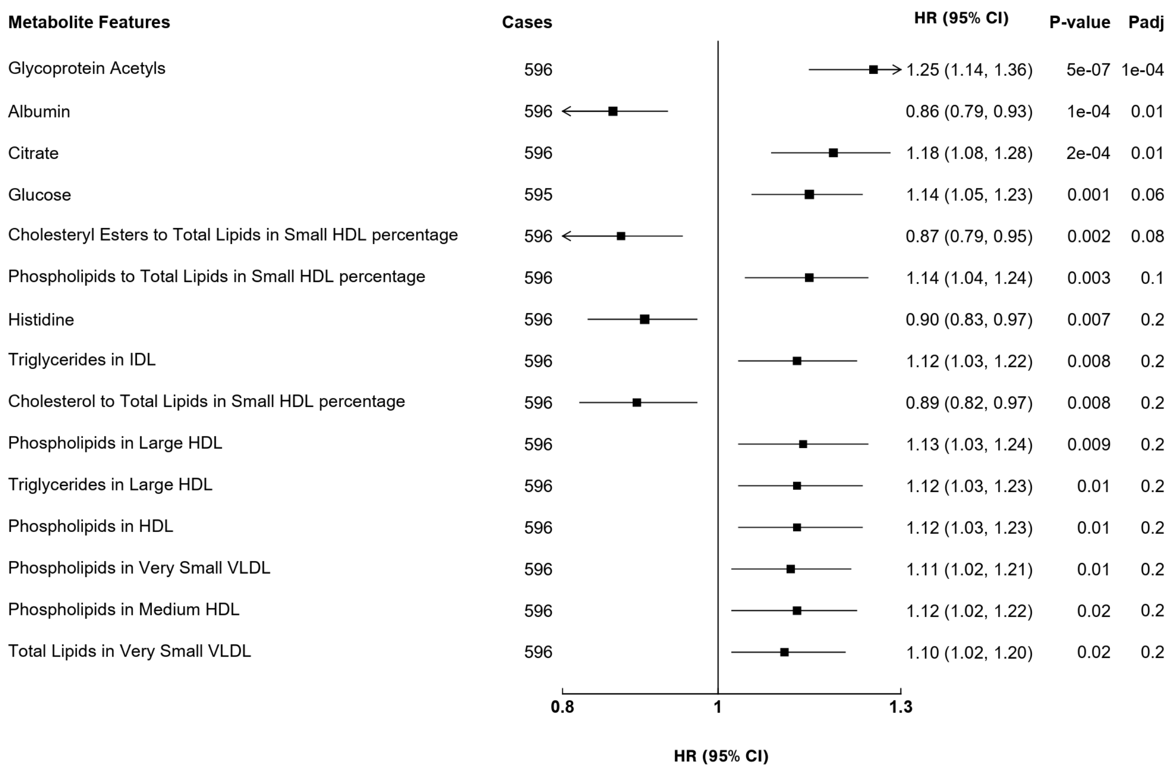


Figure 5.7 Hazard ratios (95% confidence interval) for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of 15 most significant metabolite features. HR, hazard ratios; CI, confidence interval; LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol; IDL, intermediate-density lipoprotein cholesterol; VLDL, very low-density lipoprotein cholesterol. Padj refers to P-value adjusted for multiple testing using the Benjamini–Hochberg false discovery rate (FDR) method.

Table 5.3 shows the regression dilution bias (RDB) -corrected hazard ratios for the top three most associated metabolites with risk of prostate cancer death. Regression dilution ratios (RDRs) for the repeat metabolite measurements were modest to low, indicating varying levels of within-person reproducibility (0.5 for glycoprotein acetyls, 0.4 for albumin, and 0.3 for citrate). After correction for regression dilution bias, the associations with prostate cancer death were strengthened (Table 5.3). For glycoprotein acetyls, the hazard ratio increased from 1.25 (95% CI: 1.14–1.36) to 1.51 (95% CI: 1.29–1.77). For albumin, the inverse association strengthened from 0.86 (95% CI: 0.79–0.93) to 0.65 (95% CI: 0.52–0.82). For citrate, the positive association was magnified from 1.18 (95% CI: 1.08–1.28) to 1.69 (95% CI: 1.29–2.21). These findings suggest that baseline-only estimates underestimated the strength of associations due to within-person variability; however, metabolites with lower RDRs and larger changes in RDB-corrected HRs (particularly citrate) should be interpreted more cautiously, as their corrected estimates are more strongly influenced by measurement error and short-term fluctuations.

Table 5.3 Correction of hazard ratios for regression dilution bias using repeat measures of metabolites for top 3 metabolites associated with prostate cancer death

Metabolite	RDR	New HR	New CI LL	New CI UL	Original HR	Original CI LL	Original CI UL
Glycoprotein acetyls	0.5	1.51	1.29	1.77	1.25	1.14	1.36
Albumin	0.4	0.65	0.52	0.82	0.86	0.79	0.93
Citrate	0.3	1.69	1.29	2.21	1.18	1.08	1.28

Table 5.3 Correction of hazard ratios and confidence intervals for regression dilution bias using repeat measures of metabolites for top 3 metabolites associated with prostate cancer death. RDR, regression dilution ratio; HR, hazard ratio; CI, confidence interval; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit

For results stratified by time since blood collection (Table 5.4), the glycoprotein acetyls (HR = 1.40, 95% CI 1.24–1.57, P-value = 6.4E-08, Padj = 1.6E-05) and albumin (HR = 0.79, 95% CI 0.71–0.88, P-value = 2.6E-05, Padj = 2.1E-03) associations with prostate cancer mortality was stronger in men who died ≤ 5 years since blood collection, compared to those who died from prostate cancer > 5 years since blood collection (glycoprotein acetyls: HR = 1.11, 95% CI 0.99–1.25, P-value = 0.084, Padj = 1.0; albumin: (HR = 0.93, 95% CI 0.83–1.04, P-value = 0.19, Padj = 1.0). Meanwhile, the association of citrate with prostate cancer mortality by time since blood collection appeared stable in the ≤ 5 years subgroup (HR = 1.18, 95% CI 1.05–1.33, P-value = 0.01, Padj = 1.0) (Table 5.4 and Appendix 5.7).

Table 5.4 Hazard ratios and 95% CIs of top three most significantly associated metabolite features with prostate cancer death, stratified by time since blood collection

	≤ 5 years						> 5 years					
Metabolite features	**Cases**	**HR**	**LL**	**UL**	**P-value**	**Padj**	**Cases**	**HR**	**LL**	**UL**	**P-value**	**Padj**
GlycA	290	1.40	1.24	1.57	6E-08	2E-05	306	1.11	0.99	1.25	0.1	1.0
Albumin	290	0.79	0.71	0.88	3E-05	0.002	306	0.93	0.83	1.04	0.2	1.0
Citrate	290	1.18	1.04	1.33	0.01	0.09	306	1.18	1.05	1.33	0.01	1.0

Hazard ratios and 95% confidence interval for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of three most significant metabolites or metabolic features, stratified by time since blood collection. HR, hazard ratio; CI, confidence interval; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit; GlycA, glycoprotein acetyls. Padj refers to p-value adjusted for multiple testing using the Benjamini–Hochberg false discovery rate (FDR) method.

The associations of the same three metabolites with risk of prostate cancer death were similar in the sensitivity analyses with additional adjustment for diabetes: glycoprotein acetyls (HR = 1.24, 95% CI 1.14-1.35, P-value =6.8E-07, Padj = 1.7E-04); citrate (HR = 1.17, 95% CI 1.08-1.28, P-value = 2.8E-04, Padj = 0.023); albumin (HR = 0.85, 95% CI 0.79-0.92, P-value = 9.3E-05, Padj = 0.012) (Figure 5.8; full results in Appendix 5.8 and stratified by follow-up time in Appendix 5.9).

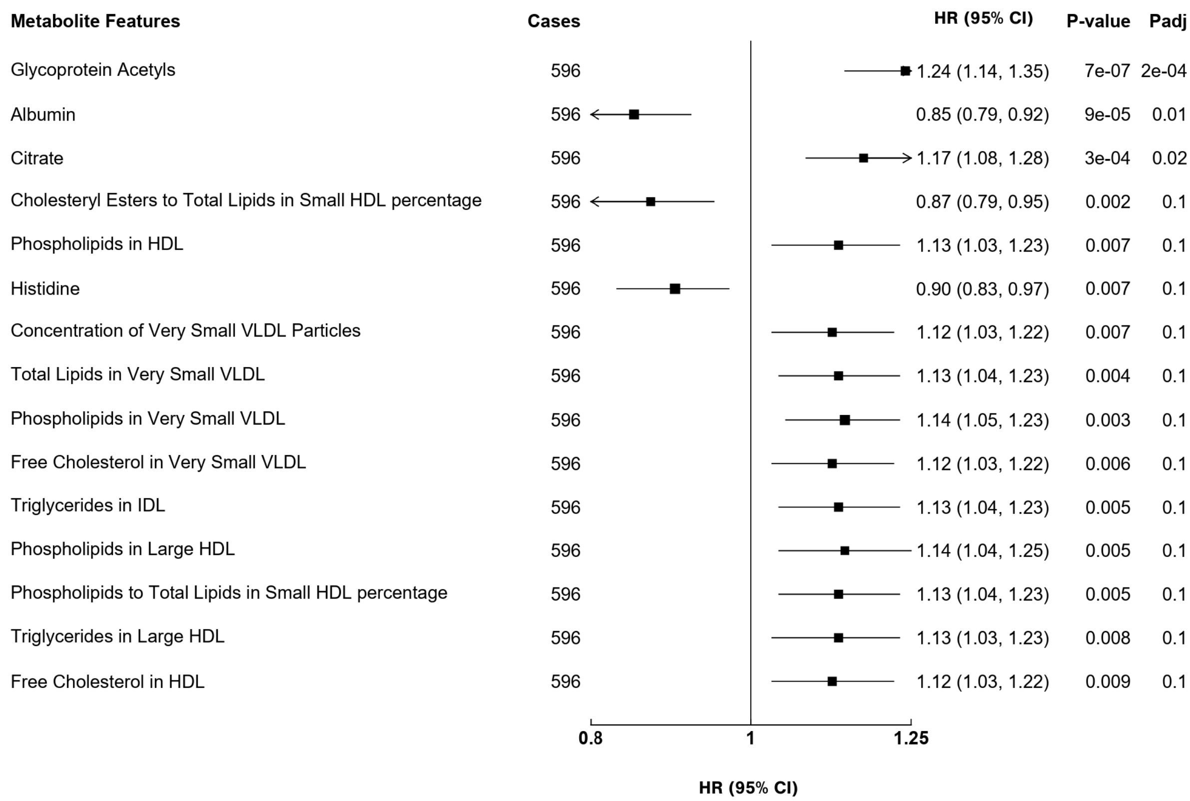


Figure 5.8 Hazard ratios (95% confidence interval) for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of 15 most significant metabolite features, after additional adjustment for diabetes. HR, hazard ratios; CI, confidence interval; LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol; IDL, intermediate-density lipoprotein cholesterol; VLDL, very low-density lipoprotein cholesterol. Padj refers to P-value adjusted for multiple testing using the Benjamini–Hochberg false discovery rate (FDR) method.

5.4.4 Amino acids and prostate cancer risk

There were no associations that were significant after multiple testing for amino acids and overall prostate cancer risk (Figure 5.9), or risk of prostate cancer death (Figure 5.10). For associations that were conventionally significant, the relative risk of prostate cancer per standard deviation increase in alanine was 0.97 (95% CI 0.95–1.00, P-value = 0.03, Padj = 0.2) and for tyrosine was 0.97 (0.97, 95% CI 0.95–1.00, P-value = 0.02, Padj = 0.1).

Men with higher concentrations of histidine (HR = 0.90, 95% CI 0.83–0.97, P-value = 0.007, Padj = 0.2) and glutamine (HR = 0.92, 95% CI 0.85–1.00, P-value = 0.05, Padj = 0.3) had lower risk of death from prostate cancer, before correction for multiple testing.

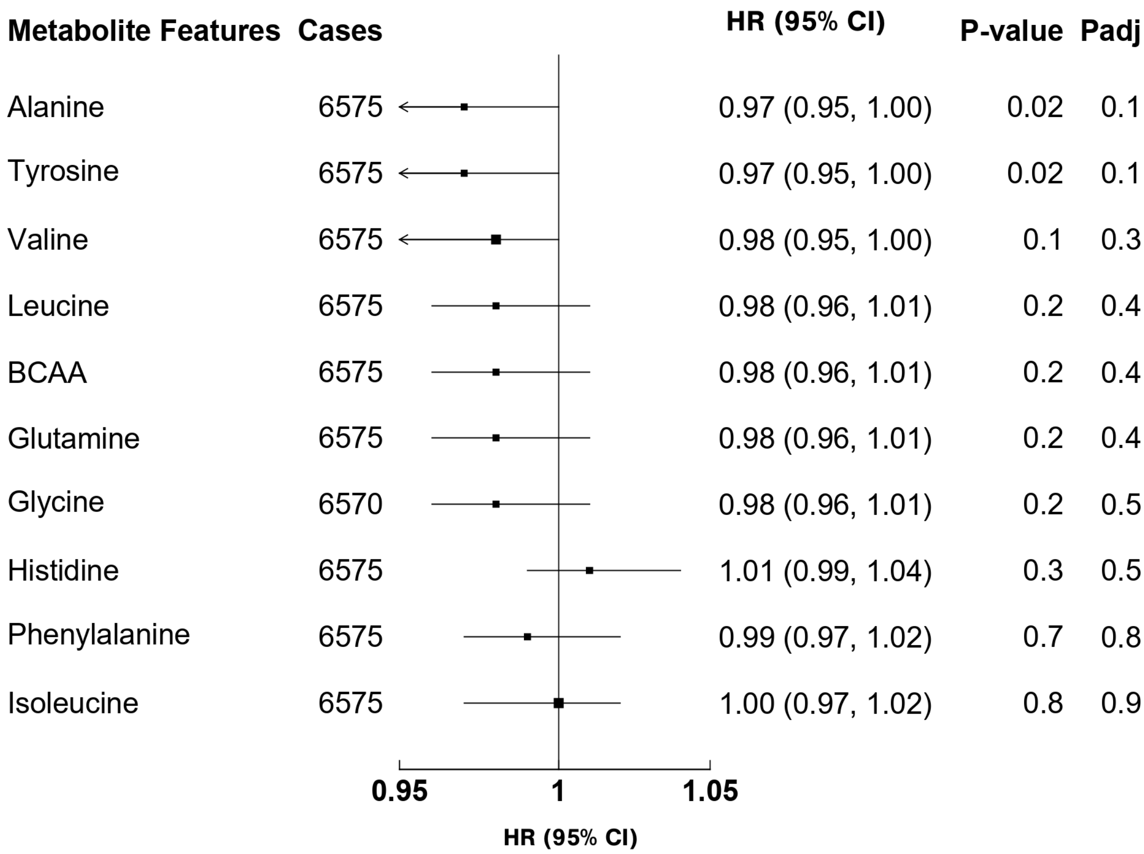


Figure 5.9 Hazard ratios (95% confidence interval) for prostate cancer associated with a one standard deviation increase in concentration of nine amino acids and one sum of amino acids. HR, hazards ratio; CI, confidence interval; BCAA, total concentration of branched-chain amino acids (leucine, isoleucine, and valine). Padj refers to P-value adjusted for multiple testing using the Benjamini–Hochberg false discovery rate (FDR) method.

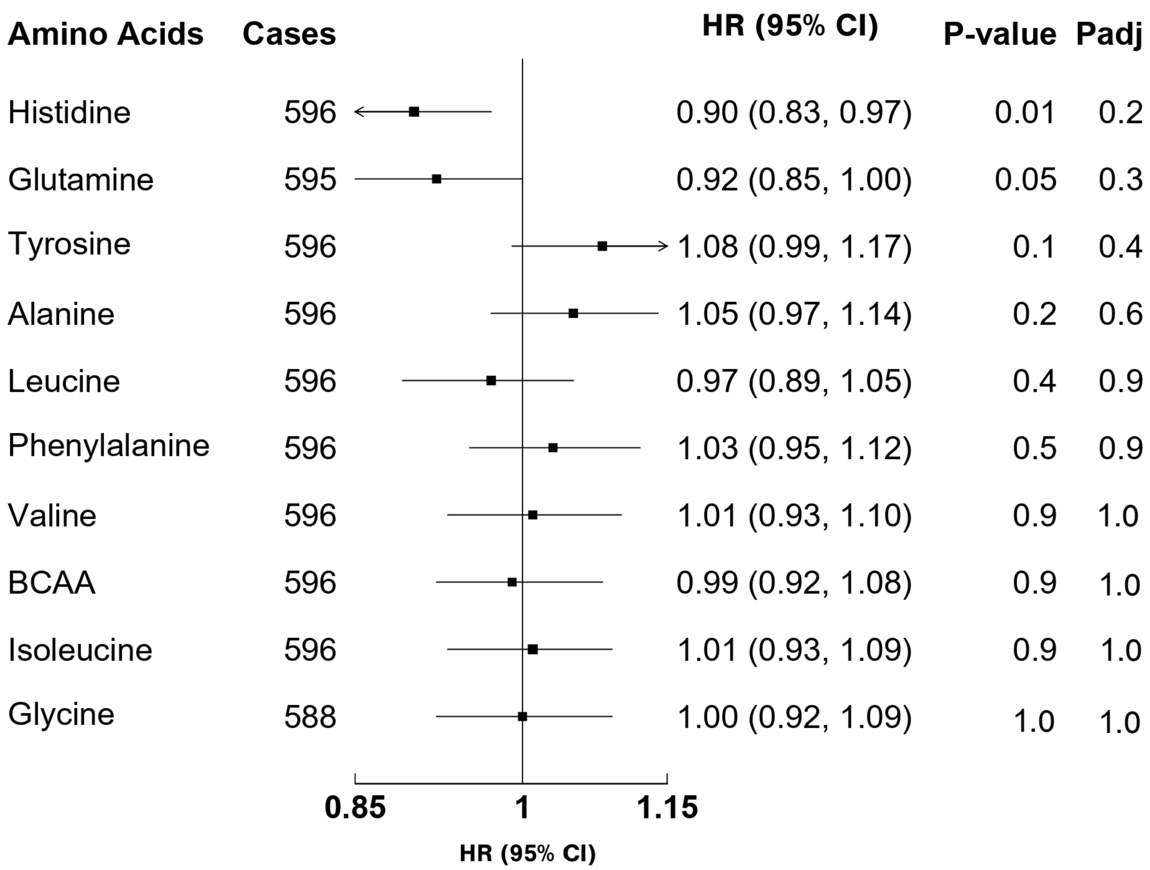


Figure 5.10 Hazard ratios (95% confidence interval) for prostate cancer death associated with a one standard deviation increase in concentration of nine amino acids and one sum of amino acids. HR, hazards ratio; CI, confidence interval; BCAA, total concentration of branched-chain amino acids (leucine, isoleucine, and valine). Padj refers to P-value adjusted for multiple testing using the Benjamini–Hochberg false discovery rate (FDR) method.

5.4.5 Cross-sectional secondary analyses

Table 5.5 presents the associations of fish intakes with geometric mean concentration of DHA and with DHA percentage to total fatty acids. Mean DHA concentrations increased with higher intakes of both oily and non oily fish (P-value <0.001 for both) (Table 5.5). Participants consuming oily fish ≥2 times per week had geometric mean DHA concentrations of 0.25 nmol/L (95% CI: 0.25–0.25), a 0.07 nmol/L increase from never to less than once a week. Men consuming non oily fish ≥2 times per week had DHA concentrations of 0.22 nmol/L (95% CI: 0.22–0.22), a 0.03 nmol/L higher than in those consuming non oily fish never to less than once a week. Likewise, the geometric mean of DHA to total fatty acids was higher with higher fish consumption; it was 2.19 nmol/L (95% CI: 2.18–2.20) among those who consumed oily fish ≥2 times per week and 1.92 nmol/L (95% CI: 1.91–1.93) among those who consumed non oily fish ≥2 times per week (Table 5.5).

Table 5.5 Intakes of oily fish and non oily fish in association with geometric mean concentration of DHA and mean DHA percentage to total fatty acids.

Fish intake	Number of men	Geometric mean DHA concentration (nmol/L; 95% CI)	P-value
Oily fish			
Never to less than once a week	57,358	0.18 (0.18-0.18)	p< 0.001
Once a week	45,577	0.21 (0.21-0.21)	p< 0.001
≥ 2 times a week	22,085	0.25 (0.25-0.25)	p< 0.001
Non oily fish			
Never to less than once a week	42,593	0.19 (0.19-0.19)	p< 0.001
Once a week	62,322	0.21 (0.21-0.21)	p< 0.001
≥ 2 times a week	20,233	0.22 (0.22-0.22)	p< 0.001
		Geometric mean percentage of DHA to total fatty acids (95% CI)	
Oily fish			
Never to less than once a week	57,358	1.57 (1.56-1.57)	p< 0.001
Once a week	45,577	1.83 (1.82-1.84)	p< 0.001
≥ 2 times a week	22,085	2.19 (2.18-2.20)	p< 0.001
Non oily fish			
Never to less than once a week	42,593	1.65 (1.64-1.65)	p< 0.001
Once a week	62,322	1.79 (1.78-1.79)	p< 0.001
≥ 2 times a week	20,233	1.92 (1.91-1.93)	p< 0.001

Table 5.5 Intakes of oily fish and non oily fish in association with geometric mean concentration of DHA and mean DHA percentage to total fatty acids. Intakes of oily fish and non oily fish in association with geometric mean concentrations of DHA and mean DHA percentage to total fatty acids. All models are adjusted for age, fasting status, body mass index, height, alcohol intake, ethnicity, level of education, smoking status, and weekly physical activity. CI, confidence interval; DHA, docosahexaenoic acid; p, P-value.

Participants in the highest quintile of DHA intake had 1.3 nmol/L (95% CI: 1.2–1.4) higher circulating IGF-I concentrations than those in the lowest quintile (Table 5.6). Similarly, higher percentages of DHA relative to total fatty acids were associated with elevated IGF-I levels (Table 5.6). Mean IGF-I concentration for those in the highest fifths of DHA percentage was

21.9 nmol/L (95% CI: 21.8–22.0), compared to 20.6 nmol/L (95% CI: 20.5–20.6) in men the lowest quintile (p<0.001).

Table 5.6 Associations of DHA concentration and DHA percentage to total fatty acids with geometric mean concentrations of IGF-I			
Exposure (fifths)	**Number of men**	**Geometric mean IGF-1 concentration (nmol/L; 95% CI)**	**P-value**
DHA		IGF-1	
1	23,909	20.6 (20.5-20.6)	p< 0.001
2	23,903	21.1 (21.0-21.2)	p< 0.001
3	23,906	21.3 (21.2-21.3)	p< 0.001
4	23,908	21.5 (21.4-21.6)	p< 0.001
5	23,902	21.9 (21.8-21.9)	p< 0.001
DHA percentage to total fatty acids		**Geometric mean IGF-1 concentration (nmol/L; 95% CI)**	
1	24,208	20.6 (20.5-20.6)	p< 0.001
2	24,056	21.1 (20.1-21.1)	p< 0.001
3	23,924	21.3 (21.2-21.4)	p< 0.001
4	23,803	21.5 (21.4-21.5)	p< 0.001
5	23,537	21.9 (21.9-22.0)	p< 0.001

Table 5.6 Associations of DHA concentration and DHA percentage to total fatty acids with geometric mean concentrations of IGF-I. All models are adjusted for age, fasting status, body mass index, height, alcohol intake, ethnicity, level of education, smoking status, and weekly physical activity. CI, confidence interval; DHA, docosahexaenoic acid; IGF-1, insulin-like growth factor-1; p, P-value.

Fish intake was associated with an increased likelihood of having ever previously undergone a PSA test (Table 5.7). Compared to individuals consuming oily fish less than once per week, those consuming oily fish ≥ 2 times per week had a 17% higher likelihood of prior PSA testing (OR = 1.17, 95% CI: 1.13–1.21, P-value < 0.001). A similar though smaller magnitude trend was observed for non-oily fish intake, where individuals consuming non-oily fish ≥2 times per week had a 9% higher likelihood of PSA testing (OR = 1.09, 95% CI 1.05–1.13, P-value < 0.001).

Table 5.7 Odds ratios of the associations of oily fish and non oily fish intakes with having previously undergone a PSA test.			
Fish intake	**Number of men**	**Odds ratio (95% CI)**	**P-value**
Oily fish			p< 0.001
Never to less than once a week	57,358	1.00	
Once a week	45,577	1.07 (1.04-1.10)	
≥ 2 times a week	22,085	1.17 (1.13-1.21)	
Non oily fish			p< 0.001
Never to less than once a week	42,593	1.00	
Once a week	62,322	1.05 (1.02-1.08)	
≥ 2 times a week	20,233	1.09 (1.05-1.13)	

Table 5.7 Odds ratios (95% confidence interval) of the associations of oily fish and non oily fish intakes with having previously undergone a PSA test. All models are adjusted for age, fasting status, body mass index, height, alcohol intake, ethnicity, level of education, smoking status, and weekly physical activity. CI, confidence interval; PSA, prostate-specific antigen; p, P-value.

Compared to individuals consuming oily fish less than once per week, those consuming oily fish ≥ 2 times per week had no significant difference in prostate cancer risk (HR = 0.99, 95% CI = 0.93–1.06) (Table 5.8). Similarly, non-oily fish intake was not associated with prostate cancer risk, as those consuming non-oily fish ≥ 2 times per week had a hazard ratio of 1.01 (95% CI 0.94–1.09) compared to the lowest intake category.

Table 5.8 Association between categories of fish intake and prostate cancer risk			
	Prostate cancer cases		
Fish intake	**Number of men**	**HR (95% CI)**	**P-value**
Oily fish			0.5
Never to less than once a week	2524	1.00	-
Once a week	2449	1.01 (0.96-1.07)	0.7
≥ 2 times a week	1245	0.99 (0.93-1.06)	0.7
Non oily fish			0.04
Never to less than once a week	1885	1.00	-
Once a week	3340	1.06 (1.00-1.12)	0.03
≥ 2 times a week	998	1.01 (0.94-1.09)	0.7

Table 5.8 Association between categories of fish intake and prostate cancer risk. Relative to a fish intake of less than once a week, stratified by region and Townsend Deprivation Index with age as the underlying time variable. Adjusted for fasting status, body mass index, height, alcohol intake, family history of

Table 5.8 Association between categories of fish intake and prostate cancer risk
prostate cancer, ethnicity, level of education, smoking status, and weekly physical activity. CI, confidence interval; HR, hazard ratio per one standard deviation from Cox regression models.

Table 5.9 presents the associations of glucose concentrations with geometric mean concentration of free testosterone. Higher glucose concentrations were associated with lower circulating free testosterone. Men in the highest quintile of glucose had 8% lower mean free testosterone compared with those in the lowest quintile (p-het<0.001).

Table 5.9 Associations of glucose concentration with geometric mean concentrations of free testosterone

Exposure (fifths)	Number of men	Geometric mean free testosterone concentration (nmol/L; 95% CI)	P-het
Glucose		Free Testosterone	p<0.001
1	22,082	205.9 (205.1-206.7)	
2	22,087	204.6 (203.8-205.4)	
3	22,013	202.0 (201.2-202.9)	
4	22,011	197.9 (197.0-198.7)	
5	21,848	189.8 (188.9-190.7)	

Table 5.9 Associations of glucose concentration with geometric mean concentrations of free testosterone. All models are adjusted for age, fasting status, body mass index, height, alcohol intake, ethnicity, level of education, smoking status, and weekly physical activity. CI, confidence interval; p-het, P-value for heterogeneity.

5.5 Discussion

This is the largest prospective study to date of these 249 circulating metabolite features in relation to prostate cancer incidence. It also distinguishes itself from prior metabolite-prostate cancer work in the UK Biobank by including subgroup analyses of time since blood collection and assessing risk of prostate cancer death [67]. In this chapter, I report that men who had a higher percentage of DHA to total fatty acids were at increased risk of developing overall prostate cancer, as well as men with a higher percentage of phospholipids to total lipids in small LDL. Both metabolite feature associations displayed heterogeneity by follow-up time, with the former demonstrating a consistent risk association in longer time since blood collection, suggesting a possible aetiological role in disease risk, and the latter driven by shorter time since blood collection.

Four other fatty acid metabolite features including degree of unsaturation (positive), ratio of PUFAs to MUFAs (positive), PUFAs to total fatty acids percentage (positive), and MUFAs to total fatty acids percentage (inverse), as well as glucose (inverse), were also associated with risk of prostate cancer; however, weaker associations with risk in the sensitivity analyses (especially for glucose) suggested that those results may have been due to confounding by diabetes.

Higher levels of glycoprotein acetyls and citrate, and lower levels of albumin were associated with risk of death from prostate cancer; the glycoprotein acetyls and albumin associations appeared to be driven by the first 5 years since blood collection, suggesting possible reverse causality. Meanwhile, the citrate association displayed little heterogeneity by time since blood collection, suggesting a potential aetiological role.

5.5.1 Metabolite features and prostate cancer risk

The previous study in UK Biobank [67] with 1,436 incident cases (with a median follow-up of 11.3 years versus 6,575 incident cases in the current analyses that were diagnosed over a median of 12.5 years) investigated associations between these metabolites and prostate cancer risk, and found lower levels of glucose, citrate, and lipoproteins (including several phospholipids) in men who were subsequently diagnosed with prostate cancer compared to those who were not [67]. In addition to having a larger sample size of 6,575 total incident prostate cancer cases, the current extended analyses differed from the previous study by investigating associations by time from blood collection to prostate cancer diagnosis, including sensitivity analyses that adjusted for diabetes, investigating prostate cancer death as an outcome, and including metabolite percentages as exposure variables (in addition to absolute metabolite concentrations, thus extending the number of metabolite features from 168 to 249). The current analyses showed a similar inverse association to that previous reported between glucose and overall prostate cancer risk, however, after additional adjustment for diabetes, we found the association to be attenuated. The metabolite exposure variable most strongly associated with prostate cancer in the current analyses, i.e. DHA to total fatty acids percentage, was not investigated in the previous study. In the previous study, a null association was reported between DHA and prostate cancer risk, and no other fatty acids were found to be associated with risk after accounting for multiple testing.

While the previous study focused exclusively on absolute concentrations of metabolites, in this analysis percentages were also investigated. This dual approach provides a more comprehensive understanding of DHA's role in prostate cancer risk. Exclusively assessing the association of DHA concentration with prostate cancer risk may not account for variation in total fatty acids, as individuals with higher overall fatty acid levels may also have higher DHA levels without a meaningful difference in DHA's contribution to risk. Including percentage of DHA to total fatty acids additionally accounted for these variations, providing a clearer picture of DHA's role within the broader fatty acid profile. Similarly, while the percentage of phospholipids to total lipids in small LDL (in which higher levels were found to be positively associated with overall prostate cancer in this chapter) demonstrates a lipoprotein-related finding, it represents a relative measure rather than an absolute concentration [67].

Regarding the higher prostate cancer risk observed in men with higher DHA levels relative to total fatty acids, some comparable associations have been reported in previous cohorts [160]. DHA is a long-chain omega-3 polyunsaturated fatty acid (PUFA), meaning it is a long-chain PUFA with one of its double bonds located 3 carbon atoms from the methyl end [161]. A case-cohort study in the Selenium and Vitamin E Cancer Prevention Trial (SELECT) found that men with higher blood concentrations of long chain omega-3 PUFAs were at increased risk of low-grade and overall prostate cancer [71]; likewise, a nested case-control study in the Prostate Cancer Prevention Trial (PCPT) found that men with higher DHA levels were 2.5 times more likely to develop high-risk prostate cancer than men with lower levels [70]. A 2014 individual participant data meta-analysis of over 5,000 cases and 6,000 controls, which included data from SELECT, PCPT, and the EPIC cohort, reported that men with higher circulating docosapentaenoic acid (DPA) (another long-chain polyunsaturated omega-3 fatty acid) had a slightly elevated risk of developing prostate cancer [69]. In subgroup analyses, there was heterogeneity by grade, in which the association of DPA with risk was driven by low-grade disease. Risk associations in subgroup analyses by stage, aggressive disease, and time to diagnosis appeared stronger in the localised, non-aggressive, and less than 3 years groups, though the formal test for heterogeneity in these three subgroup analyses was not statistically significant [69].

The role of lipid dysregulation in tumour aetiology has previously been investigated, and may provide a biological context to the findings reported in this chapter. In a 2018 review of prostate cancer studies by Lima et al., metabolic alterations in fatty acid profiles were reported in serum and plasma samples from prostate cancer patients [137]. These changes are thought to reflect underlying disruptions in lipid metabolism, including increased lipogenesis, which supports tumour growth by fuelling cell proliferation and modulating intercellular signalling pathways [162]. A key component of this altered metabolic state is enhanced fatty acid synthesis and lipid β-oxidation, which provide energy and biosynthetic material for abnormal cell proliferation. DHA, a long-chain omega-3 polyunsaturated fatty acid, may be involved in this metabolic reprogramming. Although DHA is typically considered anti-inflammatory [163], observational evidence suggests that its role in prostate cancer is complex [70]. Altered DHA levels observed in prostate cancer may reflect a shift in fatty acid metabolism, possibly through incorporation into membrane lipids or modulation of β-oxidation activity [137,162].

There were no associations between amino acids and prostate cancer risk in this study that passed correction for multiple testing. In a nested case-control analysis in the SU.VI.MAX cohort, another prospective study (171 incident cases in a cohort of 5,141 men followed for an average of 13 years) that used NMR to quantify metabolites, six amino acids were found to be positively associated with overall prostate cancer risk, including tyrosine (OR = 1.40, 95% CI 1.06–1.85], p-value = 0.02), as well as creatine and albumin lysyl; men with lower levels of urea had a higher risk of developing prostate cancer. None of the metabolite-prostate cancer associations from the SU.VI.MAX study met correction for multiple testing in the current analyses [15]. It should be noted that in contrast to Nightingale Health's targeted metabolomics approach, the SU.VI.MAX study used an untargeted NMR metabolomics approach, and therefore the metabolites investigated only partially overlap.

I reported that the inverse association between glucose and prostate cancer risk was attenuated after adjustment for diabetes status, highlighting the potential role of diabetes as either a potential confounder or mediator in this relationship. Previously, a systematic review and meta-analysis (that included ten prospective cohort studies, one nested case-control study, one case-cohort study and three case-control studies) investigated fasting glucose in relation to risk of prostate cancer. The authors reported an inverse association in the subgroup of cohort studies [164], in line with my findings (before adjustment for diabetes). Meanwhile, several meta-analyses have reported an association between the presence of diabetes, which is associated with hyperglycaemia, and reduced prostate cancer risk [155,156]. Elevated glucose, and decreased testosterone and insulin in men with diabetes, may explain the inverse association between glucose and prostate cancer risk reported in this chapter. There may be a detrimental effect of hyperglycaemia on testosterone production, as lower levels of testosterone are found in men with diabetes [165]; meanwhile, epidemiological [158] and experimental studies have presented strong evidence supporting the role of androgens in the development of prostate cancer [154–156,166]. Consistent with these studies, I report in the cross-sectional analyses that increased glucose concentrations were associated with a reduction in circulating free testosterone. Diabetes influences insulin levels in a complex manner, with early-stage hyperinsulinemia transitioning to hypoinsulinemia as pancreatic function declines [3,152]. In summary, these findings warrant consideration of the complex interplay between glucose metabolism, diabetes, and prostate cancer risk when interpreting associations between metabolic biomarkers and prostate cancer outcomes.

5.5.2 Metabolite features and prostate cancer death

In analyses of prostate cancer death, glycoprotein acetyls, albumin, and citrate showed associations that were stronger after correction for regression dilution bias. While these findings may indicate that baseline-only estimates may underestimate the true magnitude of associations, they should be interpreted cautiously. The regression dilution ratios (RDRs) for albumin and citrate in particular were low (0.4 and 0.3, respectively), suggesting considerable within-person variability. As a result, the corrected estimates for these metabolites are likely to be less reliable, and the apparently stronger associations, particularly for citrate, should be regarded with caution and not over-interpreted.

Regarding the association between glycoprotein acetyls and prostate cancer mortality, glycoprotein acetyls have previously been identified as a marker of systemic inflammation [167,168]. While glycoprotein acetyls are relatively novel NMR-identified biomarkers, and there are limited prospective findings to establish an association with prostate cancer mortality, a recent review of metabolic markers and prostate progression reported higher levels in men with prostate cancer [169]. Plasma samples were profiled from 41 South African men with high risk, aggressive prostate cancer, and NMR spectroscopy determined that these men had higher plasma glycoprotein acetyls [170]. Similarly, when stratifying by follow-up time, I reported higher circulating glycoprotein acetyls in men who died within five years of blood collection, suggesting possible reverse causality.

There is limited ability in the literature to compare the positive association between higher circulating citrate and risk of prostate cancer death to other studies that have data on advanced stage and grade. Previously in the UK Biobank, the authors reported that higher circulating citrate contributed to lower overall prostate cancer risk [67]; in line with this, I found a conventionally significant inverse association between citrate and overall prostate cancer risk within three years of blood collection (Appendix 5.3)**.** The only observational study that reported a significant association between citrate and risk of more advanced disease was the ATBC cohort, though in an opposite direction from this chapter's findings. In a case-control study of 200 matched pairs, lower levels of citrate were associated with more aggressive tumour subtypes [171]. However, authors in the PLCO cohort did highlight a conventionally significant positive association between citrate and aggressive prostate cancer in the 10-20 years since blood collection [9]. A large amount of citrate is produced and accumulated by normal prostate epithelial cells, and is then secreted as a major component of prostatic fluid. In comparison to normal prostate tissue, prostate cancer tissue often has decreased concentrations of citrate [68,172–174]. One possible mechanism for the current findings is that higher circulating citrate concentrations may reflect broader metabolic alterations occurring in advanced disease. Emerging research suggests that tumour cells may import extracellular citrate for energy and lipid synthesis, potentially contributing to more aggressive disease [68,173]. The observed positive association between baseline plasma citrate and prostate cancer mortality may therefore indicate a broader metabolic phenotype predisposing to tumour progression, perhaps for example due to bone degradation or metastatic disease, rather than a causal risk factor for prostate cancer initiation [175].

I found that lower concentrations of circulating albumin were associated with higher risk of prostate cancer death, which was driven by the first 5 years since blood collection, suggesting potential reverse causality. The previous UK Biobank study reported a positive association with albumin and risk of overall prostate cancer, and in an EPIC-Heidelberg case-cohort study (n=2,739, including 761 total cancer deaths,1,273 total men, and 554 prostate cancer cases), albumin was inversely associated with overall cancer death, but not with incident prostate cancer [176]. Lower albumin levels in aggressive prostate cancer patients have been well documented, likely attributed to an inhibition of albumin synthesis caused by systemic inflammation, malnutrition, and an elevated rate of albumin turnover driven by tumours [177–179]. Decreased albumin may contribute to cancer progression and poorer outcomes, and indicate poor nutrition, inflammation, and short survival within the body [177–179]. Thus, the current study's association may be driven by pre-existing aggressive cancer within 5 years of blood collection.

5.5.3 Comparing these results to Chapter 4 (EPIC)

The association of phospholipids to total lipids in small LDL percentage with higher disease risk was likely driven by tumour-related metabolic alterations up to 5 years before diagnosis. This metabolite feature describes the relative phospholipid content of small LDL particles, primarily composed of phosphatidylcholines. Meanwhile, in the previous chapter, I reported that higher concentrations of a metabolite pattern characterized by 64 phosphatidylcholines were associated with lower advanced disease risk; this association was stronger in the shorter period since blood collection, suggesting plasma metabolic profile changes up to a decade before diagnosis of advanced stage disease. Despite differences in direction, both associations were strongest shortly after blood collection, suggesting reverse causality in both cases. The contrasting directions likely reflect differences in the biological context measured (total circulating lipids versus lipoprotein composition) and tumour subtype outcomes. These complementary findings further support the hypothesis that lipid alterations may reflect early disease processes or undetected prostate cancer, rather than an aetiological role.

5.5.4 Cross-sectional secondary analyses investigating correlates of DHA

I conducted the cross-sectional analyses in this chapter to provide greater dietary and metabolic context to the DHA to total fatty acids percentage association with prostate cancer risk highlighted in this chapter. The results demonstrated a positive association between higher intakes of oily fish and mean blood DHA, which is in line with known evidence that fish oils contain high levels of omega-3 long chain polyunsaturated fatty acids, including DHA [180]. Regarding how this relates to more established potential prostate cancer risk factors, circulating IGF-1 has previously been associated with higher risk of prostate cancer, both in Mendelian randomisation analyses [181] and in a pooled analysis of 17 prospective studies [48]. The cross-sectional results in this chapter indicated a strong DHA-IGF-1 association, which may provide greater physiological context to this chapter's DHA-related prospective finding. A recent cross-sectional study also found that higher intakes of both oily and non-oily fish led to increased circulating insulin-like growth factor 1 (IGF-1) concentrations, and an association between intake of polyunsaturated fat long chain omega-3 fatty acids and higher IGF-1 [152], in line with the association reported in this chapter between DHA and IGF-1. However, in this chapter's prospective analyses of fish intake and overall prostate cancer risk, there was no association with risk. This is consistent with the results of the most recent meta-analysis evaluating fish intake and prostate cancer risk; out of 22 prospective cohort studies, with a follow-up period ranging from six to 33 years, there was no significant association with overall prostate cancer risk (nor for advanced or localised prostate cancer). However, the authors did report an association between higher fish intake and lower prostate cancer risk, including that a daily increase in fish intake was associated with 12% lower risk of prostate cancer death in dose-response analyses [182].

The cross-sectional analyses also provide additional context to a potential role of detection bias (see section 6.4.2.2. Detection bias); higher fish intake was positively correlated with undergoing a PSA test. Health-conscious individuals who consume more fish (which is correlated with circulating DHA) are more likely undergo a PSA screening test, thus potentially biasing the association of higher circulating DHA with higher prostate cancer incidence [69].

5.6 Strengths and Limitations

This study benefits from a prospective study design and large sample size, which strengthens the reliability of the analyses. By measuring the metabolites before the onset of prostate cancer, this design enhances the ability to establish the temporal sequence of events, reducing potential biases and improving the inference of a causal relationship [183].

Adjusting for potential confounders enhanced the robustness of the analysis by isolating the independent effects of the predictor variables, thereby reducing the likelihood of residual confounding. This ensured that the observed associations are more likely to reflect true relationships rather than the influence of unaccounted for variables. Moreover, adjusting for diabetes in the sensitivity analyses reduced the risk of confounding, particularly for strongly correlated metabolites like glucose.

The potential influence of reverse causation was assessed in risk analyses that stratified by follow-up time, which was not previously done in the 2024 UK Biobank study [67].

Although UK Biobank blood samples underwent minimal local processing that could affect labile molecules, subsequent centralised, high-throughput processing using standardised protocols helped minimise variability and ensured consistent, high-quality samples across the cohort (see section 6.4.2.3.3.1 Processing) [184].

Dietary and lifestyle covariate data were self-reported, which can result in some measurement error due to misreporting of food consumption or recall bias. Using a single dietary assessment method at one time point, such as a food frequency questionnaire, may lead to measurement errors and bias. However, several studies have supported the reproducibility and validity of the food frequency questionnaire method [185,186].

In this study, I investigated circulating metabolites using targeted NMR. While targeted metabolomics provides reliable, high-throughput quantification of a defined set of metabolites, it has its limitations, as it only measures known metabolites and provides a partial view of the metabolome. The pre-selected metabolite panel may miss important metabolic alterations relevant to prostate cancer risk and progression. However, the NMR-based method offers distinct advantages, providing better overall precision than untargeted metabolomics [55,66]. The use of a standardized panel of pre-defined metabolites ensures consistent, quantitative measurement across all participants in a study, allowing for robust within-cohort comparisons. Additionally, this standardization enhances the reproducibility and comparability of findings across studies that use the same platform.

The UK Biobank does not provide information on tumour subtypes; however, in addition to risk of overall prostate cancer, risk of death from prostate cancer was also investigated as a proxy for more aggressive disease.

The majority of participants available for these analyses were of White ethnicity (which is representative of the general UK population [187]), and the current study cannot assess metabolite-prostate cancer risk in other ethnicities. Finally, the mean follow-up time available for all participants was 11.6 years, which may not be long enough to detect associations, as prostate cancer can take decades to develop and progress.

5.7 Conclusion

In conclusion, this prospective study, which is the largest to date of these metabolites and prostate cancer, demonstrates that men with a higher DHA to total fatty acid percentage had a higher risk of prostate cancer. The association of higher phospholipids to total lipids in small LDL percentage with higher disease risk was likely caused by pre-clinical cancer alterations. Men with higher plasma albumin and lower glycoprotein acetyls had a lower risk of prostate cancer mortality within the first five years of blood collection, suggesting early metabolic shifts in response to aggressive disease up to five years before mortality. Triangulating evidence via complementary approaches is warranted to establish causality and explore potential mechanisms underlying these associations. Future work may include studies with a larger sample size, including pooled analyses of several cohorts, MR genetic studies (see section 6.5.4 Genomics), and exploring potential mechanisms including lipid metabolism and inflammation.

6. Discussion

6.1 Overview

This thesis assessed plasma metabolomics in relation to prostate cancer epidemiology, aiming to characterise the correlates of plasma metabolic profiles in men, and to investigate metabolomic associations with prostate cancer risk. This was done by first investigating whether there were potentially modifiable dietary and lifestyle correlates of metabolites previously found to be associated with prostate cancer in the EPIC cohort, then robustly further assessing metabolite-prostate cancer associations using new data from EPIC and the UK Biobank.

6.2 Summary of Findings

- In Chapter 3, cross-sectional analyses using 3,042 men in the EPIC cohort (split into 2,524 in the discovery set and 518 in the validation set) examined 24 dietary and lifestyle correlates of three metabolite patterns (and the individual metabolites loading on those patterns). Intakes of alcohol, total fish products, and its subsets total fish and lean fish were positively associated with Pattern 1 (64 phosphatidylcholines [including PC aas and PC aes], and two hydroxysphingomyelins). BMI was positively associated with Pattern 2, which appeared to be driven by a strong positive BMI-glutamate association. Finally, both BMI and fatty fish were inversely associated with Pattern 3 (eight lysophosphatidylcholines). In summary, these analyses show that fish and its subtypes, alcohol, and BMI are associated with metabolite patterns and several individual metabolites, which have in turn been linked to lower risk of aggressive prostate cancer. While the directionality of the dietary/lifestyle factor-metabolite associations cannot be established due to the cross-sectional design, this study identifies modifiable correlates of metabolites and metabolite patterns that have been a key focus of recent prostate cancer metabolomics research. These findings can help contextualise future research investigating relevant metabolic pathways potentially associated with early disease processes.

- Next, this thesis aimed to further investigate associations between metabolites and prostate cancer risk by following up on the previous prospective study in EPIC; the prior nested case-control study (n=3,057 matched pairs) had smaller sample sizes and limited statistical power for analyses in the subgroups, especially when stratifying analyses by time since blood collection to account for reverse causation. Incorporating an additional 1,330 matched pairs, Chapter 4's extended nested case-control study (n=4,387 matched pairs) increased the reliability of estimates of the associations of 148 individual metabolite concentrations and the three metabolite patterns with prostate cancer risk in aggressive tumour subtypes. Overall, there were no statistically significant associations of specific metabolites or metabolite patterns with overall, aggressive, or high grade prostate cancer. Six phosphatidylcholines were inversely associated with advanced prostate cancer diagnosed at or within 10 years of blood collection. Metabolite Patterns 1 (64 phosphatidylcholines [PC aas and PC aes] and three hydroxysphingomyelins) and 2 (two acylcarnitines, glutamate, ornithine, and taurine) were also inversely associated with advanced prostate cancer; when stratified by follow-up time, these associations were observed for diagnoses at or within ten years of recruitment but were weaker after longer follow-up. Pattern 3 (eight lysophosphatidylcholines) was associated with prostate cancer death. These results suggest that the plasma metabolite profile changes in response to the presence of prostate cancer up to a decade before detection of advanced stage disease. The analyses in this chapter represent the largest study to date of these metabolites and prostate cancer risk in the EPIC cohort.
- For the final project of this thesis, I aimed to more broadly capture the metabolome and investigate its associations with prostate cancer risk and mortality, by incorporating data from an additional metabolomics platform and cohort, representing the largest prospective

study to date of metabolites and prostate cancer risk. This prospective study of 126,206 men in the UK Biobank (6,575 prostate cancer cases) assessed the association of 249 metabolites with overall prostate cancer risk, as well as with risk for death from prostate cancer. Men who had a higher percentage of DHA to total fatty acids had a higher risk of prostate cancer, which displayed a consistent association with risk when stratified by follow-up time. In addition, a higher percentage of phospholipids to total lipids in small LDL was associated with higher risk of prostate cancer, with the greatest magnitude of association in the first 3 years since blood collection, suggesting metabolic alterations from an existing tumour. Men with higher circulating citrate concentrations had an increased risk of prostate cancer death. Finally, higher concentrations of glycoprotein acetyls and lower concentrations of albumin were associated with increased risk of death from prostate cancer, with the strongest association within 5 years of blood collection, likely indicating pre-clinical changes in blood metabolite profiles in response to an existing tumour. This is the first study of these metabolites and prostate cancer in the UK Biobank that includes robust subgroup analyses by follow-up time, and investigates prostate cancer mortality as an outcome. In cross-sectional analyses to further contextualise results, I also report that higher oily fish intake was strongly associated with higher plasma DHA concentrations, and that plasma DHA concentration was strongly positively associated with circulating IGF-I level (which is an established prostate cancer risk factor).

6.3 Findings in Context

6.3.1 Correlates of metabolites

This thesis has helped contribute to the wider current work in metabolomics by providing a greater understanding of the relationship between potentially modifiable dietary and lifestyle factors and several metabolites and metabolite patterns. Using two distinct metabolomics platforms, both Chapters 3 (EPIC) and 5 (UK Biobank) commonly found that dietary fish intake is reflected in plasma lipid profiles. In particular, fish intake had a considerable association with higher concentrations of both plasma phosphatidycholines (PC aas and PC aes) and DHA. Many phosphatidylcholines, including PC aas, contain DHA as one of their two ester-linked fatty acid side chains, and thus can be carriers of DHA, supporting the consistent findings across both chapters [188]. These results, especially in EPIC, are consistent with another cross-sectional study using the Biocrates platform in the UK Fenland Study cohort (n=10,806); the authors reported that fish consumption associated with higher phospholipids, including PC aas and PC aes, which are fundamental components of cell membranes and plasma lipoproteins [80]. While the study design investigating correlates of metabolites in this thesis prevents drawing conclusions on the directionality of associations, these results are further supported by dietary intervention studies. A meta-analysis of 34 fish intake dietary intervention studies reported significant increases in plasma DHA (as well as other long-chain omega-3 polyunsaturated fatty acids) in 31 studies post-intervention, especially with oily fish consumption [189]. Two randomised controlled trials in the meta-analysis also reported increases in DHA-containing phosphatidylcholines and other phospholipids with increasing doses of weekly oily fish consumption [189]. The findings reported in Chapters 3 and 5, supported by consistent associations across cohorts and dietary intervention studies, provide strong support for an effect of fish intake on circulating lipid metabolite concentrations. In Chapter 5, I also identified a positive correlation between circulating DHA and the prostate cancer risk factor IGF-1. Previous cross-sectional studies in the UK Biobank [152] and smaller cohorts [190] have also reported a positive association between fish intake and IGF-1, suggesting the possibility of a shared metabolic pathway between fish intake, DHA, and IGF-1. This work highlights modifiable dietary and lifestyle correlates of the metabolic landscape, which includes plasma metabolite profiles that may reflect preclinical prostate cancer, offering further context into tumour-associated pathways and the surrounding metabolic environment.

Alcohol intake and BMI were also correlated with blood metabolic profiles in this thesis, and notably in this thesis I found a particularly strong BMI-glutamate association. Several cross-sectional studies across cohorts and platforms, including in EPIC (Biocrates) [191,192], the Framingham Heart Study (custom LC-MS) [193], TwinsUK (Metabolon and Biocrates) [97,98], the Cooperative Health Research in the Region of Augsburg (KORA) (Biocrates) [98], and CARLA cohorts (Biocrates) [99] have reported a similar correlation between higher alcohol intake and higher circulating phosphatidylcholines and other lipids. These include PC aas [97–99,191–193], PC aes [191–193], hydroxysphingomyelins [191], and metabolite patterns in men characterised by PC aas, and a combination PC aas and PC aes [192]. In line with my findings of elevated phosphatidylcholine concentrations correlating with alcohol intake, the prospective Framingham Heart Study (n=2,428) reported positive associations between long-term alcohol consumption (measured over approximately 20 years) and phosphatidylcholines, including PC 32:1 and PC 34:1 [193]. However, a prospective study in the CARLA cohort (n=1,030) found decreased phosphatidylcholines, specifically four PC aes, over a four year period of alcohol consumption [99]. These longitudinal results may indicate potential longer-term metabolic responses to alcohol intake that cannot be concluded from cross-sectional correlations. Prospective studies in other cohorts with large sample sizes, long follow-up, and platforms using measurement methods besides LC-MS are warranted to validate and better understand the alcohol-phosphatidylcholine associations. A combination of Mendelian randomisation, longitudinal cohort studies with repeated measurements, and controlled intervention studies can also be complementary to further elucidating these metabolic associations in the future.

Several cohorts have identified a similar strong BMI-glutamate positive correlation in previous cross-sectional work, providing external validation to the results reported in this thesis [102,194]. Furthermore, two Mendelian randomisation analyses support a causal effect of BMI on circulating glutamate levels, using BMI-associated genetic variants from large-scale genome-wide association studies (GWAS) conducted by the Genetic Investigation of Anthropomorphic Traits (GIANT) consortium [83,195]. Thus, the findings presented in this thesis are consistent with prior research, and provide additional observational evidence that complements existing causal inferences between BMI and circulating glutamate levels. Previously in EPIC, Pattern 2 (which included a positive glutamate loading) was inversely associated with prostate cancer risk [4], which is also consistent with the observed inverse association of BMI with overall risk of prostate cancer diagnosis [196]. These results may contribute to a broader understanding of how BMI and systemic metabolic factors may interact with the prostate cancer tumour microenvironment. Nevertheless, it also warrants consideration that glutamate is a marker of higher BMI, which may affect timing of diagnosis (see section 6.4.2.2. Detection bias), rather than being an aetiological risk factor. In a recent MR analysis of glutamate and prostate cancer risk, where instruments for glutamate were obtained from a large-scale GWAS, and for prostate cancer from PRACTICAL, glutamate was not found to be associated with risk [197]. However, assessing the genetic association between glutamate and prostate cancer risk is challenging; in the study, there was only one SNP associated with glutamate, and the authors were unable to test for horizontal pleiotropy in sensitivity analyses (to ensure genetic instruments only influence the outcome through the exposure of interest; see section 6.5.4. Genomics) [197]. Overall, while my reported results further validate the observational association between circulating glutamate and BMI, understanding glutamate's potential involvement in prostate cancer pathways requires cautious interpretation and future studies with more robust genetic instruments for glutamate.

6.3.2 Prospective findings

There were several metabolite feature-prostate cancer associations that were of greatest magnitude (or only observed) for cases diagnosed soonest after blood collection compared to longer after blood collection in Chapters 4 and 5, which is suggestive of possible reverse causation. This is particularly apparent for associations of several metabolites with more aggressive tumour subtypes/mortality, including lipid-associated metabolic features like glycerophospholipids (PC aas, PC aes, and lysophosphatidylcholines) and percentage of phospholipids to total lipids in small LDL cholesterol, as well as the inflammatory biomarkers glycoprotein acetyls and albumin.

Meanwhile, the association of higher circulating citrate with higher risk of prostate cancer death was similar in subgroup analyses stratified by time since blood collection, which may be consistent with a more stable aetiological role. There are inconsistencies in the literature regarding the relationship between circulating citrate and prostate cancer risk by more advanced stage and grade. One prospective cohort reported an association between citrate and risk of advanced prostate cancer, though in contrast to my results, the association was inverse [171]. However, another cohort did highlight a positive association between citrate and aggressive disease that was strongest in cases diagnosed after longer follow-up, though this association was only conventionally significant [9]. Citrate is essential to the production of metabolic energy, and is involved in the synthesis of fatty acids and cholesterol, which may have a role in tumour development [198]. The positive association between citrate and mortality reported in this thesis may be indicative of a broader metabolic phenotype relating to tumour development, such as bone degradation or metastatic disease [172–174], rather than a causal risk factor for prostate cancer initiation. It is important to acknowledge that the observed association in this thesis could also be due to chance, owing to the smaller sample size in the analyses stratified by prostate cancer death, and requires validation from other prospective cohorts with larger sample sizes and prostate cancer death as an outcome. To date, no single SNP has been consistently associated with circulating citrate concentrations, though genetic variants that influence related metabolic pathways such as citrate mitochondrial function or citrate transport may be relevant. Additional research, such as targeted GWAS focusing on citrate concentrations (genes like *LC13A5* [Solute Carrier Family 13 Member 5], *SLC25A1* [citrate transporter], or *SLC39A1* [Prostate-Specific Citrate Metabolism ZIP1]) may help pinpoint genetic factors associated with citrate levels in future work, to better understand the causal role of citrate in more aggressive disease [198].

In the prospective study in the UK Biobank, I also highlighted a potential role of fatty acids, particularly a high percentage of the omega-3 fatty acid DHA, in overall prostate cancer risk. The associations were stable when stratified by follow-up time, consistent with a possible role in aetiology. This aligns with several observational studies [70,71] and an individual participant meta-analysis of data on 5,098 cases from seven studies [69] in other cohorts (including EPIC, SELECT, and PCPT), which have similarly reported associations between higher circulating DHA and other long-chain omega-3 polyunsaturated fatty acids and increased prostate cancer risk, including for overall [69,71], low-grade [69,71], and high-risk [70] disease. In the meta-analysis, the absence of heterogeneity by follow-up time further supports a potential aetiological role [69]. The metabolic feature percentage DHA to total fatty acids reported in this thesis cannot be validated against the previous UK Biobank study, however, as the authors did not include relative percentages as exposure variables [67]. Regarding genetic evidence, a Mendelian randomisation (MR) analysis using data from the PRACTICAL consortium and UK Biobank reported an inverse association between blood DHA and prostate cancer risk [199], but this should be interpreted with caution. The MR estimate was based on a single genetic instrument that explained less than one percent of the variance in DHA levels, potentially limiting both precision and causal interpretation [199]. Moreover, my analysis assessed percentage DHA to total fatty acids measured in mid-to-late adulthood (potentially more reflective of recent and modifiable exposures), while the MR approach assessed lifelong, genetically influenced DHA levels, and examined absolute concentrations. In summary, my finding that a higher percentage of DHA to total fatty acids is associated with increased prostate cancer risk is consistent with several prior observational studies [69–71], and the stability of associations by follow-up time suggests a possible aetiological role of this metabolic feature. Differences in detection, however, are also important to consider when interpreting fatty acid associations with risk (see 6.4.2.2. Detection Bias). In brief, men with higher DHA are more likely to eat fish and adhere to other health conscious behaviours, such as undergo a PSA test, resulting in increased prostate cancer diagnoses in these men compared to the general population [69]. Without stage and grade data, interpretation of this finding is inherently limited. Future studies should prioritise the inclusion of percentage-based fatty acid measures, which may enhance epidemiological interpretation by accounting for the relative contribution of individual fatty acids to overall lipid profiles. Cohort studies with larger sample sizes by follow-up time and tumour subtype information, as well as MR studies incorporating more comprehensive genetic instruments, are warranted to further clarify the temporality, causal nature, and underlying metabolic pathways for this association.

In the recent systematic review and meta-analysis of 12 pre-diagnostic circulating metabolomic studies and prostate cancer risk, several metabolites were associated with overall, high-risk, and risk of death from prostate cancer, including those reporting in the previous work in EPIC, and across a range of MS (including Metabolon and Biocrates) and NMR platforms [75]. Building on and extending this published literature, the work in this thesis contributes one of the largest datasets to date for metabolite associations with prostate cancer risk. Including an overlap of 3,057 EPIC cases from *Schmidt et al.* (using the Biocrates platform) [4] used in both the meta-analysis summary results and this thesis, I present 10,962 prospective cases in total between the EPIC and UK Biobank chapters (compared to 6,431 total cases in the meta-analysis), which included 4,387 Biocrates-measured cases (compared to 3,834) and 6,575 NMR-measured cases (compared to 171) [75].

Several findings from the recent meta-analysis summary results [75] are consistent with those that I report in this thesis. In particular, in the meta-analysis nine phosphatidylcholines and one carnitine (C18:2) were found to be inversely associated with high- to very high-risk prostate cancer; taurine and glutamate associated with risk of prostate cancer death (though positively, while my finding was inverse), and 1-linoleoyl-GPC (also known as lysophosphatidylcholine C18:2) inversely associated with lethal prostate cancer [75]. In line with the meta-analysis results, I reported in the EPIC nested case-control analyses (Chapter 4) associations with prostate cancer risk and death. Specifically, I found that phosphatidylcholines were inversely associated with advanced prostate cancer, and acylcarnitine C18:2, glutamate, and taurine were among the metabolites contributing to a pattern inversely associated with advanced disease. However, the associations I found were stronger in men diagnosed closer to blood collection, suggesting possible reverse causation. Furthermore, in the UK Biobank analyses (Chapter 5), my results also supported a possible role of lipid metabolism dysregulation (specifically a higher percentage of phospholipids to total lipids in small LDL) in tumour development. Similar to the Chapter 4 results, the association was strongest within 3 years of blood collection. Though the systematic review did not explicitly stratify by time since blood collection, some included studies had long follow-up periods [75], highlighting the need for future work to examine temporal trends in more detail.

The authors of the recent review and meta-analysis suggested that metabolites associated with higher risk and lethal disease, including glycerophospholipids and amino acids, may be useful in a clinical context as biomarkers for early detection of disease with poor prognosis. While their findings suggest a possible association of lipid and amino acid metabolic pathways with prostate cancer risk, the magnitude of associations observed were generally modest [75]. In the systematic review, relative risks were frequently close to 1.0, emphasising the need for caution in regards to clinical applications [75]. In order to be useful for early detection, these metabolite biomarkers would need to provide meaningful prognostic information beyond already established screening tools like PSA testing [200]. For example, future work should investigate if these metabolites provide additional predictive information by incorporating them into multivariable models adjusting for PSA, and using methods suited for high-dimensional data, such as lasso or elastic net regression to perform variable selection while minimizing overfitting.

6.3.3 Comparison of the cohorts

Table 6.1 compares characteristics of the EPIC and UK Biobank datasets. The EPIC and UK Biobank cohorts differ in several important respects. Recruitment occurred more than a decade apart, with EPIC men enrolled primarily in the 1990s and UK Biobank participants between 2006 and 2010. This temporal difference coincides with changes in diagnostic practices, particularly the introduction and wider use of PSA testing in the 2000s [2,28–30]. As such, the slightly higher incidence rate of prostate cancer observed in the UK Biobank may partly reflect greater detection through PSA screening rather than true differences in underlying disease risk.

While sociodemographic and lifestyle factors were generally comparable, men in EPIC had lower educational attainment and higher smoking prevalence than those in UK Biobank (Table 6.1). Despite these variations, mortality rates from prostate cancer were broadly similar across the two cohorts, suggesting that differences in incidence may be more strongly driven by diagnostic practices than by fatal disease burden. In particular, findings from UK Biobank should be interpreted in light of the higher prevalence of PSA testing and the potential for detection bias, whereas results from EPIC may reflect a setting with less PSA-driven diagnosis.

Table 6.1 Characteristics of men in the EPIC and UK Biobank studies		
Characteristic	**EPIC cohort (n=4,387 matched sets)**	**UK Biobank (n=126,206)**
Age at blood collection, mean years (SD)	57.7 (6.50)	56.8 (8.19)
Median year of recruitment (range)	1995 (1992-1999)	2008 (2006-2010)
Alcohol consumption, n (%)		
<1 g/day	414 (4.71%)	8493 (6.73%)
1-10 g/day	2185 (24.9%)	27362 (21.68%)
10-20 g/day	1908 (21.7%)	26720 (21.17%)
>20 g/day	3692 (42.1%)	55023 (43.60%)
Non-drinkers	563 (6.42%)	7964 (6.31%)
Physical activity		
UKB: < 10 METhrs/wk EPIC: Inactive	1739 (19.8%)	25367 (20.10%)
UKB: 10 - <20 METhrs/wk EPIC: Moderately inactive	2712 (30.9%)	20127 (15.95%)
UKB: 20 - <40 METhrs/wk EPIC: Moderately active	2058 (23.5%)	24778 (19.63%)
UKB: ≥ 40 METhrs/wk EPIC: Active	2135 (24.3%)	32,804 (25.9%)
Level of education		
UKB: O level/GCSE/CSE EPIC: Primary or equivalent	3252 (37.1%)	17254 (13.67%)
UKB: A levels EPIC: Secondary	3018 (34.4%)	6307 (5.00%)
Professional qualification/NVQ/HND/HNC/Degree or other professional qualification	2217 (25.3%)	78752 (62.40%)
Smoking status, n (%)		
Never	2767 (31.5%)	61322 (48.59%)
Former	3595 (41.0%)	48614 (38.52%)
Current	2321 (26.5%)	15641 (12.39%)

Table 6.1 Characteristics of men in the EPIC and UK Biobank studies		
BMI, n (%)		
<25 kg/m2	2520 (28.7%)	30860 (24.45%)
25- <30 kg/m2	4611 (52.6%)	62276 (49.34%)
30- <35 kg/m2	1234 (14.1%)	25279 (20.03%)
<35 kg/m2	180 (2.05%)	7331 (5.81%)
Prostate cancer incidence rate (per 100,000 person-years)	353	451
Prostate cancer death (per 100,000 person-years)	47	42

Table 6.1 Characteristics of men in the EPIC (n=4,387 matched sets) and UK Biobank (n=126,206) cohorts. EPIC prostate cancer incidence and death rates are found from relevant published work [50,201], since this information was not available in nested case-control data. UK Biobank incidence and death rates are calculated based on the entire cohort (not just NMR data) for comparability with EPIC information which was from the whole cohort. SD, standard deviation; g/day, grams per day; MET, metabolic equivalent; GCSE, General Certificate of Secondary Education; CSE, Certificate of Secondary Education; NVQ, national vocational qualification, HNC, higher national certificate; BMI, body mass index.

6.4 Methodological considerations

Strengths of the research reported in this thesis include robust, well-described study populations with extensive covariate data, comprehensive subgroup analyses, and the utilisation of both individual metabolites from two different platforms and metabolite patterns in risk associations. However, certain methodological limitations remain that make it challenging to entirely exclude the potential influence of chance, bias, and confounding on the findings.

6.4.1 The role of chance

When analysing high-dimensional data, such as numerous metabolites, a large number of statistical tests increases the likelihood of Type 1 errors. Statistical analyses in this thesis were designed to minimize Type 1 error due to chance while maintaining sufficient statistical power to detect meaningful associations.

6.4.1.1 Discovery-validation approach

In the cross-sectional analyses described in Chapter 3, I employed a discovery-validation set approach to ensure the robustness and reproducibility of the findings. This method involved splitting the data into a discovery set to identify potential associations (EPIC prostate cancer study controls, n = 2,524; 83% of participants), then validating those findings in a separate validation set (EPIC kidney, liver, and colorectal cancer nested case-control study controls, n = 518; 17% of participants). This approach helped ensure that significant findings were replicable across different datasets. Incorporating an independent validation set provided stronger evidence of associations and reduced false positives.

6.4.1.2 False discovery rate

In the discovery set in Chapter 3, after determining associations between diet/lifestyle factors and metabolite patterns, I applied FDR using the Benjamini-Hochberg method. This approach allowed me to maintain a reasonable balance between discovering true associations and controlling for false positives in the discovery phase, ensuring that the associations identified were not likely due to random chance. Associations with a Padj value below the FDR threshold of 0.05 were then validated in the independent validation set. Using an independent dataset for validation provided additional confidence that the relationships were robust and reproducible, minimizing the risk of overestimating the significance of potential associations.

For all associations of individual metabolites with prostate cancer risk in EPIC and the UK Biobank, FDR was applied for each subgroup analysis using the Benjamini-Hochberg method at a threshold of 0.05, to correct for multiple testing. All associations that passed the significance threshold of $P_{adj} < 0.05$ were considered statistically significant.

6.4.1.3 Treelet transform

6.4.1.3.1 Chapters 3 and 4

Treelet transform identifies localised clusters of correlated metabolites, in contrast to principal component analysis (PCA), which generates global linear combinations. By grouping metabolites that are statistically correlated into sparse treelets, this approach can reduce noise sensitivity and enhance interpretability, making it easier to link metabolite patterns to specific pathways potentially involved in prostate cancer risk [73,94]. The same metabolite patterns derived in *Schmidt et al.* (2020) [4] were used in both chapters, in order to compare findings in this thesis with previous prostate cancer metabolomics results in EPIC.

6.4.1.4 Subgroup analyses

To investigate potential differences in associations between metabolic profiles and prostate cancer risk across groups, subgroup analyses were run. These included EPIC analyses stratified by time to diagnosis, as well as by tumour aggressiveness. While these subgroup analyses can offer insights into how metabolite associations with prostate cancer may vary across different groups or contexts, they contain smaller sample sizes, which may have reduced statistical power and the increased potential for chance findings.

6.4.2 The role of bias

6.4.2.1 Selection of Participants

There is some selection bias in both the EPIC and UK Biobank cohorts, especially in regards to ethnicity [67,81]. Ethnic group is associated with prostate cancer risk, so conducting analyses on largely White populations may impact the generalisability of findings to non-White populations [29,187,202,203]. The EPIC cohort provides limited information on ethnic diversity in participants, though it is expected that participants are primarily of European descent [81]. Similarly, 94.6 percent of participants in the UK Biobank cohort are of White ethnicity; while this is comparable to the 2001 UK Census (94.5 percent) [187], according to 2021 census data, 82 percent of people in England and Wales are White, and 18 percent identify as Black, Asian, mixed, or other ethnic group [204]. Furthermore, UK Biobank participants are more likely to own their own property, and less likely to have a mortgage or loan, live in rental accommodation, or be a current smoker, in comparison to the general population, suggesting socioeconomic bias in the dataset. On average, UK Biobank participants are more health-conscious than the general population [187].

While the lack of diverse ethnic representation alone may not be a limitation when investigating exposure-disease associations, particularly in such a large dataset, and does not affect comparisons between cases and controls within the data, the UK Biobank and EPIC sample sizes are very limited in regards to the role of non-White ethnicity in metabolite-prostate cancer risk associations [81,187].

6.4.2.2 Detection Bias

One challenge in evaluating exposure-prostate cancer risk associations is the role of detection bias, whereby certain characteristics may affect the likelihood of certain individuals or groups to be screened, be present for testing, and/or have their cancer detected. This can lead to apparent associations of biomarkers with prostate cancer risk, when in fact individuals with a related characteristic are either more or less likely to be screened compared to other groups [69]. For example, higher fish consumption is often associated with other health-conscious behaviours, including undergoing PSA testing [69]. In the UK Biobank analyses (Chapter 5), I reported a positive correlation between fish intake and plasma DHA, as well as a correlation between higher fish intake and increased likelihood of having a PSA test compared to the general population (and as a result, a higher chance of being diagnosed with prostate cancer).

Thus, it is possible that the positive DHA-prostate cancer risk association could be partially explained by health-conscious men with a high fish intake also having an increased likelihood of having a PSA test. Detection bias may also play a role in observed associations when a biomarker is both linked to likelihood of being tested and having a positive result on a PSA test. For instance, a higher BMI has been associated with lower prostate cancer risk. When evaluating this association, it is important to note that higher BMI is associated with lower PSA levels, as well as decreased sensitivity in prostate exams and biopsies due to enlarged prostates in obese men [205]. Notably, cancers diagnosed via PSA tests are often distinct from those that are clinically significant [206]. Delayed tumour detection can result in worse prognoses, offering a possible consideration when assessing why higher BMI is associated with lower overall prostate cancer incidence, but higher risk of mortality in recent literature [205]. Ultimately, the role of detection bias is important to consider when interpreting observed associations, as it may influence the apparent relationship between exposures and prostate cancer risk [69].

6.4.2.3 Measurement of Metabolites

6.4.2.3.1 Assay platform

This thesis investigated metabolites quantified using targeted metabolomics methods, which identify specific metabolites of interest. While this approach is valuable for its precision and focus on metabolites and metabolite classes selected *a priori*, it is also inherently limited. Specifically, it restricts the analysis to a predefined subset of the metabolome, only measuring known metabolites and providing a partial view of the entire metabolomic landscape. Furthermore, although targeted platforms often provide more consistent and reproducible data compared to untargeted approaches, not all measurements are fully quantitative [207,208]. For example, in Biocrates, absolute quantification using internal standards is typically limited to specific metabolite classes (e.g., amino acids), while many other metabolites are measured semi-quantitatively (e.g. glycerophospholipids) [57]. Nevertheless, having pre-defined metabolites allows for quantifiable comparisons between control and case groups, offering insights into how metabolite levels shift in response to changing physiological states. In comparison, untargeted metabolomics can detect a broader array of compounds, including unknowns, but faces challenges such as lower precision and difficulties in metabolite identification due to the lack of reference standards [207–211].

6.4.2.3.2 Reproducibility

Only one blood sample was available per participant, which may limit the reliability of long-term estimates. Previous research has shown that targeted metabolites generally demonstrate good reproducibility over time [59,212], though it has been reported that non-fasting samples of phosphatidylcholines, acylcarnitines, and sphingomyelins may be less stable over a two year period [148]. As such, having multiple samples over time would provide more reliable data for assessing the stability of metabolites and their use as long-term risk biomarkers, since unstable metabolites could lead to exposure misclassification and attenuation of true associations with disease outcomes. [11,148].

6.4.2.3.3 Pre-analytical considerations

There may be pre-analytical factors to consider when analysing data from measurements in blood samples. For instance, different sample handling procedures between study centres in EPIC (Chapters 3 and 4), the use of non-fasting blood samples, time of day at blood collection, process delay, and freeze-thaw cycles may impact measured metabolite levels (Chapters 3-5), as well as lead to heterogeneity in metabolite concentrations between different centres and studies (Chapter 3) [144–146]. In Chapter 3, to address potential heterogeneity in metabolite concentrations between the four cancer sub-studies arising from differences in pre-analytical conditions (such as sample handling or processing delays), I applied a previously developed dedicated statistical pipeline before statistical analysis [86], which has demonstrated high reproducibility even across datasets measured on instruments [79].

Fasting status was considered across all chapters to minimise potential bias due to its influence on certain metabolite levels. In Chapters 3 and 5, models were adjusted for fasting status and time of day at blood collection, while in Chapter 4, cases and controls were matched on study centre and fasting status to reduce potential confounding from pre-analytical differences [146]. The UK Biobank (Chapter 5) did not require participants be fasting at recruitment, in order to facilitate recruitment and ensure compliance in the large-scale study. While the scale of the cohort (approximately 126,000 participants) helps mitigate the impact of the variability from non-fasting status, I included fasting status as a covariate to control for its potential influence on metabolite levels. Although fasting is a common consideration in metabolomics studies, several studies, including a 2019 review of pre-analytical conditions, have suggested a limited role of fasting status in metabolite variation [144,146,213,214]. Among the metabolites measured in EPIC and UK Biobank, amino acids, hexoses, and some lipids are most affected by fasting status [214], and potentially plasma fatty acids (including DHA) [215].

6.4.2.3.3.1 Processing

Blood samples were minimally processed locally before being transported to a central facility for processing, which could cause degradation of cellular components and labile molecules such as lipids. However, the use of standardized collection and processing protocols reduces variability by ensuring consistency across all samples, including through centralised high-throughput processing, which improves quality control. To prevent freeze-thaw degradation of samples and preserve long-term stability, blood samples were stored in ultra-low temperatures ranging from -80°C and -180°C, and were also aliquoted into smaller volumes to avoid repeated freeze-thaw cycles [81,184].

6.4.2.4 Measurement of Food Intake

Chapters 3 and 5 in this thesis both used data on food intake exposures, including 23 dietary variables in Chapter 3, and on intake of oily and non oily fish intake in Chapter 5. More information in available in the EPIC study on estimated nutrient intake, including the reporting of grams per day, energy intake, and portion sizes, while the UK Biobank inquired about food frequency only. While the latter is less detailed, there is also less recall burden for participants, possibly reducing the likelihood of portion misreporting. Please see below for more details:

6.4.2.4.1 EPIC

In EPIC, the dietary intake data, which were used in the analyses in Chapter 3, was obtained using detailed food frequency questionnaires (FFQs). While comprehensive, dietary date from an FFQ is subject to well-recognised limitations including recall bias, misreporting, and errors related to the food composition tables used. Measurement error in self-reported dietary intake can attenuate associations between diet and metabolite levels or disease risk, particularly if intake is misclassified or crudely estimated. Despite this, numerous pilot and cross-sectional studies have supported the reproducibility and validity of the food frequency questionnaire method [79,86,130].

6.4.2.4.2 UK Biobank

Dietary data in the UK Biobank were collected via a touchscreen food frequency questionnaire at baseline, which has shown moderate to substantial reproducibility over a four-year period [186]. Validation against repeated web-based 24-hour dietary assessments confirmed that the questionnaire reliably ranks participants according to their intake of key food groups, including fish [186]. While the questionnaire is limited in its ability to estimate the exact quantities of foods or nutrients consumed and may be affected by self-reporting bias, it performs well in classifying individuals how frequently different foods are consumed (e.g. once per week, 2–4 times per week). This makes it suitable for prospective analyses that rely on comparing participants across broad intake categories. Therefore, the associations observed in this study between diet, circulating metabolites, and prostate cancer risk are unlikely to be substantially biased by dietary misclassification [186].

6.4.2.5 Follow-up and Categorisation of Prostate Cancer

Loss to follow-up was likely minimal in all three studies, as participants in both EPIC [81] and UK Biobank [216] were flagged for cancer and mortality outcomes through high-quality, population-based national registries, making substantial bias from loss to follow-up unlikely. Moreover, any losses would be expected to occur at random with respect to baseline metabolite concentrations, and thus would more likely dilute associations rather than create spurious findings.

Categorisation of prostate cancer may present a source of potential bias by introducing misclassification of disease subtypes, particularly in EPIC. In both EPIC and UK Biobank, prostate cancer cases were primarily identified through cancer registry data using ICD-10 codes [81,216]. In EPIC, detailed data on tumour stage and grade were available for a subset of prostate cancer cases, allowing stratification by tumour aggressiveness. However, inconsistencies in staging across centres and missing data in some regions may have led to misclassification, particularly in subtype analyses. While cases and controls were matched by centre for the main risk analyses, misclassification of aggressive disease within centres could bias subtype-specific associations. For example, if aggressive cases were incorrectly categorised as non-aggressive, and this misclassification was unrelated to metabolite levels, it would likely dilute observed associations and bias estimates toward the null in analyses stratified by tumour subtype. In the UK Biobank, no tumour aggressiveness markers such as grade, stage, or Gleason score were available; use of prostate cancer death as an outcome, determined via death certificates, likely provides high specificity, but if used as a proxy for aggressive disease will miss non-fatal yet aggressive cases. Because this misclassification is unlikely to be related to metabolite levels, it would likely reduce the contrast between groups, making it harder to detect true associations and thereby biasing risk estimates toward the null in subtype-specific analyses.

6.4.2.6 Reverse Causality

Reverse causality is a potential source of bias in prospective cohort studies, when preclinical or undiagnosed prostate cancer may affect metabolite concentrations at baseline. Rather than reflecting exposures that contribute to the development of cancer, altered metabolite levels may instead reflect early biological alterations caused by the presence of disease. This is particularly relevant for prostate cancer, where tumours can remain asymptomatic and undiagnosed for years, and for metabolites that are sensitive to systemic changes such as inflammation or altered energy metabolism [136,137].

To assess the potential impact of reverse causality, both studies included time-stratified analyses based on the interval between blood collection and prostate cancer diagnosis. In EPIC, analyses were stratified by time to diagnosis of <10 years versus ≥10 years. In the UK Biobank, analyses were stratified by time to diagnosis (<3, 3–7, and >7 years) for prostate cancer risk, and by ≤5 years and >5 years for prostate cancer death. This allowed for the investigation of whether associations between metabolites and prostate cancer risk (or risk of death) were stronger in the period closer to diagnosis, which would be more likely to reflect disease-driven metabolic changes. Observed associations that persisted over longer follow-up periods would be less likely to reflect reverse causality and more likely to represent potential etiological relationships. Although this approach cannot fully eliminate the possibility of reverse causality, it can assess its potential influence and strengthen the interpretation of findings. The majority of significant associations reported in this thesis were stronger in groups with shorter follow-up time, suggesting possible reverse causality and the effects of preclinical cancer on the metabolome.

6.4.3 The Role of Confounding

Metabolite concentrations can be influenced by a range of dietary, lifestyle, demographic, and genetic factors. Some of these factors may also be associated with prostate cancer risk, creating the potential for confounding in analyses of metabolite–risk associations. Both the EPIC [81] cohort and UK Biobank [186,216] provide a range of unique, detailed information on several lifestyle factors that helped inform the covariates chosen in this thesis. In all three studies, key confounders were identified through a combination of considerations, including through running different models, observational analyses, and *a priori* based on previous literature and studies [2,4,50,79,98,98,103,115,127,192,215,217–220]. Covariates were then selected and adjusted for in statistical models. For example:

- BMI, smoking, and alcohol intake have been shown to influence circulating metabolite levels [79,98,103,192,221], and have also been linked (though inconsistently) to prostate cancer risk [196,222,223].
- Education and deprivation may be associated with both metabolite profiles (via lifestyle and health behaviours) [224] and differential access to prostate cancer screening or diagnosis [218].

- Age is an established prostate cancer risk factor [2], while marital status has previously shown an association with prostate cancer risk, perhaps due to married men practising healthier behaviours, and adherence to routine PSA testing [217].
- Variables such as time of day of blood collection, fasting status, and energy intake primarily are associated with metabolite concentrations [148,214,215], and could introduce bias if unevenly distributed between cases and controls.

Common covariates among all three studies included BMI, alcohol intake, smoking status, level of education, fasting status (matching factor in Chapter 4), and alcohol intake. Additional adjustment factors included age at recruitment (Chapters 3 and 4), energy intake (Chapter 3), time of day of blood collection (covariate in Chapter 3, matching factor Chapter 4), physical activity (Chapters 3 and 5), marital status (Chapter 4), family history of prostate cancer (Chapter 5), and height (Chapter 5). In addition, age was used as the underlying time variable in Cox proportional hazards models to control for age-related variation in prostate cancer risk, and models were stratified by region and Townsend deprivation score.

In the nested case-control study (Chapter 4), potential confounding was reduced through matching on key variables including age at blood collection, time of day, fasting status, study centre, and length of follow-up. These factors are either directly associated with metabolite levels, prostate cancer risk, or both, and could otherwise distort associations if unevenly distributed. By matching cases and controls on these variables, the study ensured comparability at baseline, minimized confounding by design, and reduced the need for extensive statistical adjustment, strengthening the validity of observed associations between metabolites and prostate cancer risk.

Chapter 5 included sensitivity analyses that adjusted for (a) diabetes and (b) history of PSA testing.

Diabetes mellitus is associated with the regulation of blood glucose levels [164]. Several studies have indicated that men with diabetes may have a lower risk of being diagnosed with prostate cancer [155,156,164]. In Chapter 5, glucose was associated with reduced overall prostate cancer risk, after correction for multiple testing. In order to determine whether the association with glucose was a result of residual confounding from diabetes, I conducted a sensitivity analysis in which all outcomes were adjusted for presence of diabetes. The association with glucose was attenuated after the adjustment.

The UK Biobank provides information about men's history of PSA testing [225]. PSA testing has an integral role in prostate cancer diagnosis, suggesting that observed associations between risk factors and prostate cancer could be affected by detection bias [226]. For instance, men who are screened are more likely to be diagnosed with prostate cancer. Also, healthier individuals may be more likely to choose to undergo PSA tests, which may lead to biased metabolite-prostate cancer risk associations. In the UK Biobank, 29 percent of men reported that they had ever had a PSA test, which is higher than a prior study in the general UK population, which was reported to be approximately six percent for men in the same age group [225]. In most Western populations, PSA testing is often correlated with healthy behaviours [225]. As a result, history of PSA testing was selected as an additional covariate for sensitivity analyses. PSA testing was treated as a sensitivity analysis rather than a main covariate because it is a detection factor rather than a fundamental biological confounder, affecting diagnosis rates but not metabolite levels. Adjusting for BMI, smoking, alcohol, and deprivation already accounted for major confounders, making PSA testing less necessary in primary models. Including PSA testing in sensitivity analyses allowed assessment of whether screening influenced metabolite associations without overcorrecting the main model. This approach strengthened the study's conclusions by ensuring that observed metabolite associations were not artificially driven by differences in PSA screening rates.

6.4.4 Summary of Limitations

This thesis contributed substantially to the literature on dietary/lifestyle correlates of metabolites, and on the association of metabolites with prostate cancer risk via the two largest prospective studies to date in the EPIC and UK Biobank cohorts, respectively. Nevertheless, there are some key methodological limitations. Both cohorts comprised individuals of mainly White ethnicity, and there was selection bias towards more health-conscious individuals. In addition, Self-reported dietary and lifestyle data may be subject to measurement error. While residual confounding is possible in all observational analyses, extensive adjustment for known lifestyle and demographic covariates reduces this chance

Despite large sample sizes in each chapter, subgroup analyses had smaller sample sizes, limiting the ability to detect modest associations. Tumour subtype misclassification in EPIC may have biased associations toward the null, as it is possible that some aggressive tumours were grouped with low-risk cases due to incomplete or inconsistent staging and grading across centres. This would reduce the contrast between groups and attenuate observed associations. In UK Biobank, the absence of tumour classification limited the ability to examine subtype-specific associations; using prostate cancer mortality as a proxy for aggressiveness captures only the most severe cases and may miss earlier or less fatal aggressive disease.

Finally, all metabolite measurements were based on a single blood sample per participant, which may not fully capture long-term metabolic profiles or account for intra-individual variability over time.

6.5 Future Research

6.5.1 Other metabolomics platforms

Expanding the range of metabolomics platforms used in prostate cancer research could help quantify a broader spectrum of metabolites in relation to disease risk. While NMR and MS can each be employed in both targeted and untargeted assays, they differ in their analytical strengths, with NMR offering high reproducibility and MS providing greater sensitivity, particularly for low-abundance metabolites. Despite these advantages, each technique alone offers only partial coverage of the metabolome [55,66]. Platforms such as Biocrates and Metabolon apply these technologies distinctly, using pre-determined panels, metabolite compound libraries, and workflows that prioritise different metabolite classes. As a result, relying on a single platform may overlook important metabolic signals. Future studies should consider integrating multiple, complementary platforms within the same analysis to enhance metabolite coverage, improve quantification, and ultimately detect metabolites that are currently underrepresented or absent in the current analyses. This approach could provide a more comprehensive view of metabolomic variation and help uncover novel metabolic pathways involved in prostate cancer risk.

6.5.2 Future work in the UK Biobank

Future analyses can investigate metabolite associations stratified by tumour characteristics, such as stage and grade, once this data becomes available in UK Biobank. This will allow for more direct comparisons with findings from EPIC, and better evaluation of potential differences in associations by disease aggressiveness.

In addition, while this thesis primarily focused on individual metabolites in UK Biobank (Chapter 5), future work can investigate the association of metabolite patterns with prostate cancer risk, similar to Chapter 4. Utilising a pattern-based approach may capture underlying metabolic processes more effectively than single-metabolite analyses, and better account for correlations between metabolites.

Once both tumour subtype data become available, and more participant metabolite data are released via the Nightingale platform, it would be valuable to investigate whether the metabolite patterns identified in EPIC (e.g. the three patterns originally derived by *Schmidt et al.* [4]) can be replicated using the NMR platform. This could be done, for example, by comparing how key loadings on each pattern (such as phosphatidylcholines) relate to NMR features like lipid or fatty acid profiles. Exploring this overlap using datasets where both metabolite platforms are available would help determine whether pattern-level associations observed in EPIC might be applicable to UK Biobank.

6.5.3 Effect of the tumour on the metabolome

6.5.3.1 Temporality of the association between metabolites and prostate cancer risk

A key challenge in interpreting metabolite–prostate cancer risk associations is establishing whether observed metabolic differences reflect causal risk factors for prostate cancer, or are instead consequences of early, undiagnosed tumour development. Several results from the time since blood collection-stratified analyses in Chapters 4 and 5 indicated that undetected prostate cancer could be altering the metabolome. Local responses that are detectable in the blood stream to a growing tumour, including inflammation, immune activation, or metabolic reprogramming well before clinical diagnosis could be affecting men's metabolic profiles [136–139,162,227]. This is particularly relevant for cancers like prostate cancer, which often have long latency periods and are frequently asymptomatic in early stages. Future studies can further follow up on this thesis's findings of the tumour's effects on the metabolome. These may include analyses with repeated metabolite measurements over time to track within-person changes prior to diagnosis, as well as more detailed clinical information at each time point, such as PSA levels, imaging findings, or biopsy results.

6.5.3.2 Tumour Tissue

Future studies can incorporate metabolomic profiling of prostate tumour tissue to better understand whether changes in circulating metabolites reflect an underlying tumour. Studies of circulating metabolites may not fully capture the tumour microenvironment. If certain metabolite alterations are detected both in tumour tissue and in blood samples close to diagnosis, it may suggest that the tumour is actively contributing to the circulating metabolite profile. This could help address the issue of reverse causality by identifying metabolites that are elevated due to tumour metabolism. Comparing blood and tumour metabolite profiles may clarify whether observed associations are systemic or tumour-specific, and offer insights into metabolic pathways involved in cancer aetiology [228–230].

6.5.4 Genomics

To better understand which metabolites may play a causal role in prostate cancer development, future research can aim to triangulate findings using complementary approaches. While observational associations can highlight metabolic associations, they are susceptible to confounding and reverse causality. Integrating genetic data through methods such as Mendelian randomisation (MR) offers a way to strengthen causal inference by using genetic variants as proxies for long-term metabolite levels, under the assumption that these variants are not influenced by confounding or early disease processes [231].

MR analyses can help prioritise metabolites for future epidemiological analyses by identifying those with genetic instruments that also associate with prostate cancer risk. Triangulating evidence across observational, genetic, and mechanistic studies increases confidence in identifying aetiological metabolites, and can inform research that explores the biological pathways through which these metabolites may influence prostate tumour carcinogenesis. At the time of writing, a study very recently used GWAS data to investigate 1,091 plasma metabolites and 309 ratios in relation to prostate cancer risk [232]. Studies such as this can further elucidate the directional association between genetic variation, metabolites, and prostate cancer risk [231,232]. Using genetic variants as proxies for metabolite levels in MR studies can help determine whether observed associations reflect potential causal relationships, rather than being consequences of pre-clinical disease.

Nevertheless, applying MR to metabolomics should be approached with caution. First, because many metabolites are highly correlated, there are very few SNPs that are uniquely associated with individual metabolites. As a result, many instruments used in MR analyses are pleiotropic, reflecting broader metabolic pathways rather than specific metabolites [83,233]. Furthermore, many SNPs associated with circulating metabolites also show pleiotropic associations with other traits, including lifestyle-related risk factors and various disease endpoints. This type of horizontal pleiotropy violates a core assumption of MR, in which genetic instruments influence the outcome only through the exposure of interest. Thus, it may be difficult to elucidate whether observed associations reflect a true causal effect of the metabolite or are instead driven by these alternative pathways, potentially biasing MR estimates and weakening causal inference [83,233]. A more comprehensive understanding of the genetic determinants of circulating metabolites, combined with ongoing methodological developments, may facilitate more reliable causal inference in future studies [83,233].

6.5.5 Collaborative work

Pooled analyses from large consortia can increase statistical power, including by enabling more detailed analyses by subgroup and tumour subtype, and supporting replication of findings across populations and platforms. This is especially useful for investigating associations across a range of ethnicities, as the findings from this study, which included data from EPIC and the UK Biobank, comprised mainly individuals of White ethnicity. One example of this is the

Consortium of Metabolomics Studies (COMETS), which includes 47 cohorts from Asia, Europe, North America, and South America, adding to a total of 136,000 participants and 17 platforms [234].

6.5.6 Proteomics as an alternative approach

Beyond metabolomics, proteomics offers another complementary layer of insight into prostate cancer mechanisms. Protein biomarkers are often more direct products of genetic variation compared to metabolites. In addition, protein biomarkers can serve as potential drug targets, and are increasingly being studied using Mendelian randomisation to strengthen causal inference. Multiplex proteomic platforms now allow the simultaneous measurement of thousands of proteins in large population cohorts, providing opportunities to combine proteomic and metabolomic data to further probe disease mechanisms [235,236].

6.6 Conclusion

In conclusion, this thesis contributes substantially to research on metabolomics and prostate cancer risk. Featuring the largest nested case-control and prospective cohort studies to date on metabolites and prostate cancer risk, and investigating a broad range of metabolites measured by two metabolomics platforms, the work reported here contributes to our understanding of metabolites and prostate cancer risk, including for more aggressive disease or death. A key finding is that men's plasma metabolite profiles start to change in response to advanced stage disease up to 10 years before diagnosis, and that several lipid-related metabolites are involved in these alterations. In addition, higher plasma DHA was associated with increased risk of prostate cancer, with stable risk associations by follow-up time, which may indicate a potential aetiological role that needs further prospective and genetic evidence. Finally, men's metabolic profiles, including those that have previously been associated with disease risk, are modified by alcohol and fish intake, and BMI. Future research can further investigate these findings through collaborative work, including the pooling of several studies with extended follow-up time and a combination of metabolomics platforms, and the triangulation of findings using complementary approaches like genetic MR studies to better assess causality.

References

1 Global Cancer Observatory. Cancer Today. Lyon, France: International Agency for Research on Cancer; 2020. https://gco.iarc.fr/today/online-analysis-table?v=2020&mode=cancer&mode_population=continents&population=900&populations=900&key=asr&sex=0&cancer=39&type=0&statistic=5&prevalence=0&population_group=0&ages_group%5B%5D=0&ages_group%5B%5D=17&group_cancer=1&include_nmsc=0&include_nmsc_other=1 (accessed 3 Mar 2022).

2 Bergengren O, Pekala KR, Matsoukas K, *et al.* 2022 Update on Prostate Cancer Epidemiology and Risk Factors—A Systematic Review. *European Urology*. 2023;84:191–206. doi: 10.1016/j.eururo.2023.04.021

3 Watts EL, Fensom GK, Smith Byrne K, *et al.* Circulating insulin-like growth factor-I, total and free testosterone concentrations and prostate cancer risk in 200 000 men in UK Biobank. *Int J Cancer*. 2021;148:2274–88. doi: 10.1002/ijc.33416

4 Schmidt JA, Fensom GK, Rinaldi S, *et al.* Patterns in metabolite profile are associated with risk of more aggressive prostate cancer: A prospective study of 3,057 matched case–control sets from EPIC. *Int J Cancer*. 2020;146:720–30. doi: 10.1002/ijc.32314

5 Fages A, Duarte-Salles T, Stepien M, *et al.* Metabolomic profiles of hepatocellular carcinoma in a European prospective cohort. *BMC Med*. 2015;13:242. doi: 10.1186/s12916-015-0462-9

6 Schmidt JA, Fensom GK, Rinaldi S, *et al.* Pre-diagnostic metabolite concentrations and prostate cancer risk in 1077 cases and 1077 matched controls in the European Prospective Investigation into Cancer and Nutrition. *BMC Med*. 2017;15:122. doi: 10.1186/s12916-017-0885-6

7 Mondul AM, Moore SC, Weinstein SJ, *et al.* 1-Stearoylglycerol is associated with risk of prostate cancer: results from a serum metabolomic profiling analysis. *Metabolomics*. 2014;10:1036–41. doi: 10.1007/s11306-014-0643-0

8 Huang J, Mondul AM, Weinstein SJ, *et al.* Prospective serum metabolomic profile of prostate cancer by size and extent of primary tumor. *Oncotarget*. 2017;8:45190–9. doi: 10.18632/oncotarget.16775

9 Huang J, Mondul AM, Weinstein SJ, *et al.* Serum metabolomic profiling of prostate cancer risk in the prostate, lung, colorectal, and ovarian cancer screening trial. *Br J Cancer*. 2016;115:1087–95. doi: 10.1038/bjc.2016.305

10 Röhnisch HE, Kyrø C, Olsen A, *et al.* Identification of metabolites associated with prostate cancer risk: a nested case-control study with long follow-up in the Northern Sweden Health and Disease Study. *BMC Med*. 2020;18:187. doi: 10.1186/s12916-020-01655-1

11 Kühn T, Floegel A, Sookthai D, *et al.* Higher plasma levels of lysophosphatidylcholine 18:0 are related to a lower risk of common cancers in a prospective metabolomics study. *BMC Med*. 2016;14:13. doi: 10.1186/s12916-016-0552-3

12 Mondul AM, Moore SC, Weinstein SJ, *et al.* Metabolomic analysis of prostate cancer risk in a prospective cohort: The alpha-tocopherol, beta-carotene cancer prevention (ATBC) study. *Int J Cancer*. 2015;137:2124–32. doi: 10.1002/ijc.29576

13 Östman JR, Pinto RC, Ebbels TMD, *et al.* Identification of prediagnostic metabolites associated with prostate cancer risk by untargeted mass spectrometry-based metabolomics: A case-control

study nested in the Northern Sweden Health and Disease Study. *Intl Journal of Cancer*. 2022;151:2115–27. doi: 10.1002/ijc.34223

14 Huang J, Mondul AM, Weinstein SJ, *et al.* Prospective serum metabolomic profiling of lethal prostate cancer. *Int J Cancer*. 2019;145:3231–43. doi: 10.1002/ijc.32218

15 Lécuyer L, Victor Bala A, Demidem A, *et al.* NMR metabolomic profiles associated with long-term risk of prostate cancer. *Metabolomics*. 2021;17:32. doi: 10.1007/s11306-021-01780-9

16 Baker SA, Rutter J. Metabolites as signalling molecules. *Nat Rev Mol Cell Biol*. 2023;24:355–74. doi: 10.1038/s41580-022-00572-w

17 Elia I, Haigis MC. Metabolites and the tumour microenvironment: from cellular mechanisms to systemic metabolism. *Nat Metab*. 2021;3:21–32. doi: 10.1038/s42255-020-00317-z

18 Metallo CM, Vander Heiden MG. Understanding metabolic regulation and its influence on cell physiology. *Mol Cell*. 2013;49:388–98. doi: 10.1016/j.molcel.2013.01.018

19 Chen L, Xu Y-X, Wang Y-S, *et al.* Lipid metabolism, amino acid metabolism, and prostate cancer: a crucial metabolic journey. *Asian Journal of Andrology*. 2024;26:123–34. doi: 10.4103/aja202363

20 Wanjari UR, Mukherjee AG, Gopalakrishnan AV, *et al.* Role of Metabolism and Metabolic Pathways in Prostate Cancer. *Metabolites*. 2023;13:183. doi: 10.3390/metabo13020183

21 Grenville ZS, Noor U, His M, *et al.* Diet and BMI Correlate with Metabolite Patterns Associated with Aggressive Prostate Cancer. *Nutrients*. 2022;14:3306. doi: 10.3390/nu14163306

22 Grenville ZS, Noor U, Rinaldi S, *et al.* Perturbations in the blood metabolome up to a decade before prostate cancer diagnosis in 4387 matched case–control sets from the European Prospective Investigation into Cancer and Nutrition. *Intl Journal of Cancer*. 2025;156:943–52. doi: 10.1002/ijc.35208

23 StataCorp. Stata Statistical Software: Release 17. College Station, TX: StataCorp LLC; 2021.

24 StataCorp. Stata Statistical Software: Release 18. College Station, TX: StataCorp LLC; 2023.

25 StataCorp. Stata Statistical Software: Release 19. College Station, TX: StataCorp LLC; 2025.

26 R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2020. Available at: https://www.R-project.org/.

27 Bray F, Laversanne M, Sung H, *et al.* Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA A Cancer J Clinicians*. 2024;74:229–63. doi: 10.3322/caac.21834

28 Kamangar F, Dores GM, Anderson WF. Patterns of Cancer Incidence, Mortality, and Prevalence Across Five Continents: Defining Priorities to Reduce Cancer Disparities in Different Geographic Regions of the World. *JCO*. 2006;24:2137–50. doi: 10.1200/JCO.2005.05.2308

29 Rawla P. Epidemiology of Prostate Cancer. *World J Oncol*. 2019;10:63–89. doi: 10.14740/wjon1191

30 Fleshner K, Carlsson SV, Roobol MJ. The effect of the USPSTF PSA screening recommendation on prostate cancer incidence patterns in the USA. *Nat Rev Urol*. 2017;14:26–37. doi: 10.1038/nrurol.2016.251

31 Graham NJ, Souter LH, Salami SS. A systematic review of family history, race/ethnicity, and genetic risk on prostate cancer detection and outcomes: Considerations in PSA-based screening. *Urologic Oncology: Seminars and Original Investigations*. 2025;43:29–40. doi: 10.1016/j.urolonc.2024.06.002

32 Kingdom R, Wright CF. Incomplete Penetrance and Variable Expressivity: From Clinical Studies to Population Cohorts. *Front Genet*. 2022;13:920390. doi: 10.3389/fgene.2022.920390

33 Hall R, Bancroft E, Pashayan N, *et al.* Genetics of prostate cancer: a review of latest evidence. *J Med Genet*. 2024;61:915–26. doi: 10.1136/jmg-2024-109845

34 Marino F, Totaro A, Gandi C, *et al.* Germline mutations in prostate cancer: a systematic review of the evidence for personalized medicine. *Prostate Cancer Prostatic Dis*. 2023;26:655–64. doi: 10.1038/s41391-022-00609-3

35 Dias A, Kote-Jarai Z, Mikropoulos C, *et al.* Prostate Cancer Germline Variations and Implications for Screening and Treatment. *Cold Spring Harb Perspect Med*. 2018;8:a030379. doi: 10.1101/cshperspect.a030379

36 Brandão A, Paulo P, Teixeira MR. Hereditary Predisposition to Prostate Cancer: From Genetics to Clinical Implications. *IJMS*. 2020;21:5036. doi: 10.3390/ijms21145036

37 Conti DV, Darst BF, Moss LC, *et al.* Trans-ancestry genome-wide association meta-analysis of prostate cancer identifies new susceptibility loci and informs genetic risk prediction. *Nat Genet*. 2021;53:65–75. doi: 10.1038/s41588-020-00748-0

38 Raghallaigh HN, Bott SR. The Role of Family History and Germline Genetics in Prostate Cancer Disease Profile and Screening. In: Frimley Park Hospital, Frimley Health NHS Foundation Trust, Camberley, UK, Barber N, eds. *Urologic Cancers*. Exon Publications 2022:199–213.

39 Nyberg T, Tischkowitz M, Antoniou AC. BRCA1 and BRCA2 pathogenic variants and prostate cancer risk: systematic review and meta-analysis. *Br J Cancer*. 2022;126:1067–81. doi: 10.1038/s41416-021-01675-5

40 Graham NJ, Souter LH, Salami SS. A Systematic Review of Family History, Race/Ethnicity, and Genetic Risk on Prostate Cancer Detection and Outcomes: Considerations in PSA-based Screening. *Urologic Oncology: Seminars and Original Investigations*. 2024;S1078143924005040. doi: 10.1016/j.urolonc.2024.06.002

41 Dee EC, Todd R, Ng K, *et al.* Racial disparities in prostate cancer in the UK and the USA: similarities, differences and steps forwards. *Nat Rev Urol*. 2025;22:223–34. doi: 10.1038/s41585-024-00948-x

42 Down L, Barlow M, Bailey SER, *et al.* Association between patient ethnicity and prostate cancer diagnosis following a prostate-specific antigen test: a cohort study of 730,000 men in primary care in the UK. *BMC Med*. 2024;22:82. doi: 10.1186/s12916-024-03283-5

43 Lloyd T, Hounsome L, Mehay A, *et al.* Lifetime risk of being diagnosed with, or dying from, prostate cancer by major ethnic group in England 2008–2010. *BMC Med*. 2015;13:171. doi: 10.1186/s12916-015-0405-5

44 Gu J, Chery L, González GMN, *et al.* A west African ancestry-associated SNP on 8q24 predicts a positive biopsy in African American men with suspected prostate cancer following PSA screening. *The Prostate*. 2024;84:694–705. doi: 10.1002/pros.24686

45 Vickers AJ, Elfiky A, Freeman VL, *et al.* Race, Biology, Disparities, and Prostate Cancer. *European Urology*. 2022;81:463–5. doi: 10.1016/j.eururo.2022.02.007

46 Watts EL, Perez-Cornago A, Fensom GK, *et al.* Circulating free testosterone and risk of aggressive prostate cancer: Prospective and Mendelian randomisation analyses in international consortia. *Intl Journal of Cancer*. 2022;151:1033–46. doi: 10.1002/ijc.34116

47 Shea JL, Wong P-Y, Chen Y. Free Testosterone. *Advances in Clinical Chemistry*. Elsevier 2014:59–84.

48 Travis RC, Appleby PN, Martin RM, *et al.* A Meta-analysis of Individual Participant Data Reveals an Association between Circulating Levels of IGF-I and Prostate Cancer Risk. *Cancer Res*. 2016;76:2288–300. doi: 10.1158/0008-5472.CAN-15-1551

49 Tzenios N, Tazanios ME, Chahine M. The impact of body mass index on prostate cancer: An updated systematic review and meta-analysis. *Medicine*. 2022;101:e30191. doi: 10.1097/MD.0000000000030191

50 Perez-Cornago A, Appleby PN, Pischon T, *et al.* Tall height and obesity are associated with an increased risk of aggressive prostate cancer: results from the EPIC cohort study. *BMC Med*. 2017;15:115. doi: 10.1186/s12916-017-0876-7

51 Perez-Cornago A, Dunneram Y, Watts EL, *et al.* Adiposity and risk of prostate cancer death: a prospective analysis in UK Biobank and meta-analysis of published studies. *BMC Med*. 2022;20:143. doi: 10.1186/s12916-022-02336-x

52 Clish CB. Metabolomics: an emerging but powerful tool for precision medicine. *Cold Spring Harb Mol Case Stud*. 2015;1:a000588. doi: 10.1101/mcs.a000588

53 Klassen A, Faccio AT, Canuto GAB, *et al.* Metabolomics: Definitions and Significance in Systems Biology. In: Sussulini A, ed. *Metabolomics: From Fundamentals to Clinical Applications*. Cham: Springer International Publishing 2017:3–17.

54 Silva AMN, Vitorino R, Domingues MRM, *et al.* Post-translational Modifications and Mass Spectrometry Detection. *Free Radical Biology and Medicine*. 2013;65:925–41. doi: 10.1016/j.freeradbiomed.2013.08.184

55 Emwas A-HM. The Strengths and Weaknesses of NMR Spectroscopy and Mass Spectrometry with Particular Focus on Metabolomics Research. In: Bjerrum JT, ed. *Metabonomics*. New York, NY: Springer New York 2015:161–93.

56 Alves S, Paris A, Rathahao-Paris E. Mass spectrometry-based metabolomics for an in-depth questioning of human health. *Advances in Clinical Chemistry*. Elsevier 2020:147–91.

57 Siskos AP, Jain P, Römisch-Margl W, *et al.* Interlaboratory Reproducibility of a Targeted Metabolomics Platform for Analysis of Human Serum and Plasma. *Anal Chem*. 2017;89:656–65. doi: 10.1021/acs.analchem.6b02930

58 Quell JD, Römisch-Margl W, Haid M, *et al.* Characterization of Bulk Phosphatidylcholine Compositions in Human Plasma Using Side-Chain Resolving Lipidomics. *Metabolites*. 2019;9:109. doi: 10.3390/metabo9060109

59 Floegel A, Drogan D, Wang-Sattler R, *et al.* Reliability of Serum Metabolite Concentrations over a 4-Month Period Using a Targeted Metabolomic Approach. *PLoS ONE*. 2011;6:e21103. doi: 10.1371/journal.pone.0021103

60 Wang Y, Jacobs EJ, Carter BD, *et al.* Plasma Metabolomic Profiles and Risk of Advanced and Fatal Prostate Cancer. *European Urology Oncology*. 2021;4:56–65. doi: 10.1016/j.euo.2019.07.005

61 Nagana Gowda GA, Raftery D. NMR-Based Metabolomics. In: Hu S, ed. *Cancer Metabolomics*. Cham: Springer International Publishing 2021:19–37.

62 Trock BJ. Application of metabolomics to prostate cancer. *Urologic Oncology: Seminars and Original Investigations*. 2011;29:572–81. doi: 10.1016/j.urolonc.2011.08.002

63 Julkunen H, Cichońska A, Tiainen M, *et al.* Atlas of plasma NMR biomarkers for health and disease in 118,461 individuals from the UK Biobank. *Nat Commun*. 2023;14:604. doi: 10.1038/s41467-023-36231-7

64 Uk Biobank. UK Biobank adds the first tranche of data from a study into circulating metabolomic biomarkers to its biomedical database. 2021. https://www.ukbiobank.ac.uk/learn-more-about-uk-biobank/news/uk-biobank-adds-the-first-tranche-of-data-from-a-study-into-circulating-metabolomic-biomarkers-to-its-biomedical-database (accessed 3 September 2024)

65 Beecher CWW. The Human Metabolome. In: Harrigan GG, Goodacre R, eds. *Metabolic Profiling: Its Role in Biomarker Discovery and Gene Function Analysis*. Boston, MA: Springer US 2003:311–9.

66 Djukovic D, Raftery D, Gowda N. Mass spectrometry and NMR spectroscopy based quantitative metabolomics. *Proteomic and Metabolomic Approaches to Biomarker Discovery*. Elsevier 2020:289–311.

67 Hu C, Fan Y, Lin Z, *et al.* Metabolomic landscape of overall and common cancers in the UK Biobank: A prospective cohort study. *Intl Journal of Cancer*. 2024;155:27–39. doi: 10.1002/ijc.34884

68 Lin X, Lécuyer L, Liu X, *et al.* Plasma Metabolomics for Discovery of Early Metabolic Markers of Prostate Cancer Based on Ultra-High-Performance Liquid Chromatography-High Resolution Mass Spectrometry. *Cancers (Basel)*. 2021;13:3140. doi: 10.3390/cancers13133140

69 Crowe FL, Appleby PN, Travis RC, *et al.* Circulating fatty acids and prostate cancer risk: individual participant meta-analysis of prospective studies. *J Natl Cancer Inst*. 2014;106:dju240. doi: 10.1093/jnci/dju240

70 Brasky TM, Till C, White E, *et al.* Serum Phospholipid Fatty Acids and Prostate Cancer Risk: Results From the Prostate Cancer Prevention Trial. *American Journal of Epidemiology*. 2011;173:1429–39. doi: 10.1093/aje/kwr027

71 Brasky TM, Darke AK, Song X, *et al.* Plasma Phospholipid Fatty Acids and Prostate Cancer Risk in the SELECT Trial. *JNCI: Journal of the National Cancer Institute*. 2013;105:1132–41. doi: 10.1093/jnci/djt174

72 Saccenti E, Hoefsloot HCJ, Smilde AK, *et al.* Reflections on univariate and multivariate analysis of metabolomics data. *Metabolomics*. 2014;10:361–74. doi: 10.1007/s11306-013-0598-6

73 Gorst-Rasmussen A. tt: Treelet Transform with Stata. *The Stata Journal*. 2012;12:130–46. doi: 10.1177/1536867X1201200108

74 Foerster J, Hyötyläinen T, Oresic M, *et al.* Serum Lipid and Serum Metabolite Components in relation to anthropometric parameters in EPIC-Potsdam participants. *Metabolism*. 2015;64:1348–58. doi: 10.1016/j.metabol.2015.07.004

75 Fuller H, Agasaro OP, Darst BF. Pre-diagnostic circulating metabolomics and prostate cancer risk: A systematic review and meta-analysis. 2025.

76 Breeur M, Ferrari P, Dossus L, *et al.* Pan-cancer analysis of pre-diagnostic blood metabolite concentrations in the European Prospective Investigation into Cancer and Nutrition. *BMC Med*. 2022;20:351. doi: 10.1186/s12916-022-02553-4

77 Scalbert A, Brennan L, Manach C, *et al.* The food metabolome: a window over dietary exposure. *The American Journal of Clinical Nutrition*. 2014;99:1286–308. doi: 10.3945/ajcn.113.076133

78 Scalbert A, Brennan L, Fiehn O, *et al.* Mass-spectrometry-based metabolomics: limitations and recommendations for future progress with particular focus on nutrition research. *Metabolomics*. 2009;5:435–58. doi: 10.1007/s11306-009-0168-0

79 His M, Viallon V, Dossus L, *et al.* Lifestyle correlates of eight breast cancer-related metabolites: a cross-sectional study within the EPIC cohort. *BMC Med*. 2021;19:312. doi: 10.1186/s12916-021-02183-2

80 Tong TYN, Koulman A, Griffin JL, *et al.* A Combination of Metabolites Predicts Adherence to the Mediterranean Diet Pattern and Its Associations with Insulin Sensitivity and Lipid Homeostasis in the General Population: The Fenland Study, United Kingdom. *The Journal of Nutrition*. 2020;150:568–78. doi: 10.1093/jn/nxz263

81 Riboli E, Hunt K, Slimani N, *et al.* European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutr*. 2002;5:1113–24. doi: 10.1079/PHN2002394

82 Kliemann N, Viallon V, Murphy N, *et al.* Metabolic signatures of greater body size and their associations with risk of colorectal and endometrial cancers in the European Prospective Investigation into Cancer and Nutrition. *BMC Med*. 2021;19:101. doi: 10.1186/s12916-021-01970-1

83 Guida F, Tan VY, Corbin LJ, *et al.* The blood metabolome of incident kidney cancer: A case-control study nested within the MetKid consortium. *PLoS Med*. 2021;18:e1003786. doi: 10.1371/journal.pmed.1003786

84 Stepien M, Duarte-Salles T, Fedirko V, *et al.* Alteration of amino acid and biogenic amine metabolism in hepatobiliary cancers: Findings from a prospective cohort study: Circulating biomarkers of hepatobiliary cancer risks. *Int J Cancer*. 2016;138:348–60. doi: 10.1002/ijc.29718

85 Guida F, Severi G, Giles GG, *et al.* Metabolomics and risk of kidney cancer. *Revue d'Épidémiologie et de Santé Publique*. 2018;66:S291. doi: 10.1016/j.respe.2018.05.144

86 Viallon V, His M, Rinaldi S, *et al.* A New Pipeline for the Normalization and Pooling of Metabolomics Data. *Metabolites*. 2021;11:631. doi: 10.3390/metabo11090631

87 Van Puyvelde H, Perez-Cornago A, Casagrande C, *et al.* Comparing Calculated Nutrient Intakes Using Different Food Composition Databases: Results from the European Prospective Investigation into Cancer and Nutrition (EPIC) Cohort. *Nutrients*. 2020;12:2906. doi: 10.3390/nu12102906

88 *Great Britain Ministry of Agriculture, Fisheries, and Food. Food Portion Sizes*. 2nd ed. London: HMSO 1993.

89 Murphy N, Norat T, Ferrari P, *et al.* Consumption of Dairy Products and Colorectal Cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC). *PLoS ONE*. 2013;8:e72715. doi: 10.1371/journal.pone.0072715

90 Pala V, Krogh V, Berrino F, *et al.* Meat, eggs, dairy products, and risk of breast cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. *The American Journal of Clinical Nutrition*. 2009;90:602–12. doi: 10.3945/ajcn.2008.27173

91 Zamora-Ros R, Castañeda J, Rinaldi S, *et al.* Consumption of Fish Is Not Associated with Risk of Differentiated Thyroid Carcinoma in the European Prospective Investigation into Cancer and Nutrition (EPIC) Study. *J Nutr*. 2017;147:1366–73. doi: 10.3945/jn.117.247874

92 Linseisen J, Bergström E, Gafá L, *et al.* Consumption of added fats and oils in the European Prospective Investigation into Cancer and Nutrition (EPIC) centres across 10 European countries as assessed by 24-hour dietary recalls. *Public Health Nutr*. 2002;5:1227–42. doi: 10.1079/PHN2002401

93 Zamora-Ros R, Béraud V, Franceschi S, *et al.* Consumption of fruits, vegetables and fruit juices and differentiated thyroid carcinoma risk in the European Prospective Investigation into Cancer and Nutrition (EPIC) study: Fruits, vegetables, and fruit juices and thyroid cancer in EPIC. *Int J Cancer*. 2018;142:449–59. doi: 10.1002/ijc.30880

94 Gorst-Rasmussen A, Dahm CC, Dethlefsen C, *et al.* Exploring Dietary Patterns By Using the Treelet Transform. *American Journal of Epidemiology*. 2011;173:1097–104. doi: 10.1093/aje/kwr060

95 Wareham NJ, Jakes RW, Rennie KL, *et al.* Validity and repeatability of a simple index derived from the short physical activity questionnaire used in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. *Public Health Nutr*. 2003;6:407–13. doi: 10.1079/PHN2002439

96 Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1995;57:289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

97 Pallister T, Jennings A, Mohney RP, *et al.* Characterizing Blood Metabolomics Profiles Associated with Self-Reported Food Intakes in Female Twins. *PLoS ONE*. 2016;11:e0158568. doi: 10.1371/journal.pone.0158568

98 Jaremek M, Yu Z, Mangino M, *et al.* Alcohol-induced metabolomic differences in humans. *Transl Psychiatry*. 2013;3:e276. doi: 10.1038/tp.2013.55

99 Lacruz ME, Kluttig A, Tiller D, *et al.* Cardiovascular Risk Factors Associated With Blood Metabolite Concentrations and Their Alterations During a 4-Year Period in a Population-Based Cohort. *Circ Cardiovasc Genet*. 2016;9:487–94. doi: 10.1161/CIRCGENETICS.116.001444

100 Hodson L, Crowe FL, McLachlan KJ, *et al.* Effect of supplementation with flaxseed oil and different doses of fish oil for 2 weeks on plasma phosphatidylcholine fatty acids in young women. *Eur J Clin Nutr*. 2018;72:832–40. doi: 10.1038/s41430-018-0174-2

101 Floegel A, von Ruesten A, Drogan D, *et al.* Variation of serum metabolites related to habitual diet: a targeted metabolomic approach in EPIC-Potsdam. *Eur J Clin Nutr*. 2013;67:1100–8. doi: 10.1038/ejcn.2013.147

102 Cheng S, Rhee EP, Larson MG, *et al.* Metabolite Profiling Identifies Pathways Associated With Metabolic Risk in Humans. *Circulation*. 2012;125:2222–31. doi: 10.1161/CIRCULATIONAHA.111.067827

103 Carayol M, Leitzmann MF, Ferrari P, *et al.* Blood Metabolic Signatures of Body Mass Index: A Targeted Metabolomics Study in the EPIC Cohort. *J Proteome Res*. 2017;16:3137–46. doi: 10.1021/acs.jproteome.6b01062

104 Abdolahi A, Georas SN, Thomas Brenna J, *et al.* The effects of aspirin and fish oil consumption on lysophosphatidylcholines and lysophosphatidic acids and their correlates with platelet aggregation in adults with diabetes mellitus. *Prostaglandins, Leukotrienes and Essential Fatty Acids*. 2014;90:61–8. doi: 10.1016/j.plefa.2013.12.004

105 Terry P, Lichtenstein P, Feychting M, *et al.* Fatty fish consumption and risk of prostate cancer. *The Lancet*. 2001;357:1764–6. doi: 10.1016/S0140-6736(00)04889-3

106 Augustsson K, Michaud DS, Rimm EB, *et al.* A prospective study of intake of fish and marine fatty acids and prostate cancer. *Cancer Epidemiol Biomarkers Prev*. 2003;12:64–7.

107 Allen NE, Sauvaget C, Roddam AW, *et al.* A prospective study of diet and prostate cancer in Japanese men. *Cancer Causes Control*. 2004;15:911–20. doi: 10.1007/s10552-004-1683-y

108 Mills PK, Beeson WL, Phillips RL, *et al.* Cohort study of diet, lifestyle, and prostate cancer in adventist men. *Cancer*. 1989;64:598–604. doi: 10.1002/1097-0142(19890801)64:3<598::AID-CNCR2820640306>3.0.CO;2-6

109 Severson RK, Nomura AM, Grove JS, *et al.* A prospective study of demographics, diet, and prostate cancer among men of Japanese ancestry in Hawaii. *Cancer Res*. 1989;49:1857–60.

110 Hsing AW, McLaughlin JK, Schuman LM, *et al.* Diet, tobacco use, and fatal prostate cancer: results from the Lutheran Brotherhood Cohort Study. *Cancer Res*. 1990;50:6836–40.

111 Le Marchand L, Kolonel LN, Wilkens LR, *et al.* Animal fat consumption and prostate cancer: a prospective study in Hawaii. *Epidemiology*. 1994;5:276–82. doi: 10.1097/00001648-199405000-00004

112 Grönberg H, Damber L, Damber JE. Total food consumption and body mass index in relation to prostate cancer risk: a case-control study in Sweden with prospectively collected exposure data. *J Urol*. 1996;155:969–74.

113 Schuurman AG, Brandt PA van den, Dorant E, *et al.* Animal products, calcium and protein and prostate cancer risk in the Netherlands Cohort Study. *Br J Cancer*. 1999;80:1107–13. doi: 10.1038/sj.bjc.6690472

114 World Cancer Research Fund/ American Institute for Cancer Research. Diet, nutrition, physical activity and prostate cancer. Continuous Update Project Expert Report. 2018. Available from: https://www.dietandcancerreport.com (accessed 30 June 2022).

115 Rohrmann S, Linseisen J, Key TJ, *et al.* Alcohol Consumption and the Risk for Prostate Cancer in the European Prospective Investigation into Cancer and Nutrition. *Cancer Epidemiol Biomarkers Prev*. 2008;17:1282–7. doi: 10.1158/1055-9965.EPI-07-2888

116 Schuurman AG, Goldbohm RA, van den Brandt PA. A prospective cohort study on consumption of alcoholic beverages in relation to prostate cancer incidence (The Netherlands). *Cancer Causes and Control*. 1999;10:597–605. doi: 10.1023/A:1008925103542

117 Breslow RA, Wideroff L, Graubard BI, *et al.* Alcohol and Prostate Cancer in the NHANES I Epidemiologic Follow-Up Study. *Annals of Epidemiology*. 1999;9:254–61. doi: 10.1016/S1047-2797(98)00071-4

118 Nilsen TIL, Vatten LJ. Anthropometry and prostate cancer risk: a prospective study of 22,248 Norwegian men. *Cancer Causes and Control*. 1999;10:269–75. doi: 10.1023/A:1008967330619

119 Baglietto L, Severi G, English DR, *et al.* Alcohol consumption and prostate cancer risk: Results from the Melbourne collaborative cohort study. *Int J Cancer*. 2006;119:1501–4. doi: 10.1002/ijc.21983

120 Weinstein SJ, Stolzenberg-Solomon R, Pietinen P, *et al.* Dietary factors of one-carbon metabolism and prostate cancer risk. *The American Journal of Clinical Nutrition*. 2006;84:929–35. doi: 10.1093/ajcn/84.4.929

121 Sesso HD, Paffenbarger RS, Lee I-M. Alcohol consumption and risk of prostate cancer: The Harvard Alumni Health Study. *International Journal of Epidemiology*. 2001;30:749–55. doi: 10.1093/ije/30.4.749

122 Platz EA. Alcohol Intake, Drinking Patterns, and Risk of Prostate Cancer in a Large Prospective Cohort Study. *American Journal of Epidemiology*. 2004;159:444–53. doi: 10.1093/aje/kwh062

123 Velicer CM, Kristal A, White E. Alcohol Use and the Risk of Prostate Cancer: Results From the VITAL Cohort Study. *Nutrition and Cancer*. 2006;56:50–6. doi: 10.1207/s15327914nc5601_7

124 Fujita K, Hayashi T, Matsushita M, *et al.* Obesity, Inflammation, and Prostate Cancer. *JCM*. 2019;8:201. doi: 10.3390/jcm8020201

125 Littman AJ, White E, Kristal AR. Anthropometrics and Prostate Cancer Risk. *American Journal of Epidemiology*. 2007;165:1271–9. doi: 10.1093/aje/kwm013

126 Rodriguez C, Freedland SJ, Deka A, *et al.* Body Mass Index, Weight Change, and Risk of Prostate Cancer in the Cancer Prevention Study II Nutrition Cohort. *Cancer Epidemiol Biomarkers Prev*. 2007;16:63–9. doi: 10.1158/1055-9965.EPI-06-0754

127 Cao Y, Ma J. Body Mass Index, Prostate Cancer–Specific Mortality, and Biochemical Recurrence: a Systematic Review and Meta-analysis. *Cancer Prev Res*. 2011;4:486–501. doi: 10.1158/1940-6207.CAPR-10-0229

128 Möller E, Wilson KM, Batista JL, *et al.* Body size across the life course and prostate cancer in the Health Professionals Follow-up Study: Body size and prostate cancer. *Int J Cancer*. 2016;138:853–65. doi: 10.1002/ijc.29842

129 Perez-Cornago A, Fensom GK, Andrews C, *et al.* Examination of potential novel biochemical factors in relation to prostate cancer incidence and mortality in UK Biobank. *Br J Cancer*. 2020;123:1808–17. doi: 10.1038/s41416-020-01081-3

130 Margetts B, Pietinen P. European Prospective Investigation into Cancer and Nutrition: validity studies on dietary assessment methods. *International Journal of Epidemiology*. 1997;26:1S – 5. doi: 10.1093/ije/26.suppl_1.S1

131 Fontana L, Adelaiye RM, Rastelli AL, *et al.* Dietary protein restriction inhibits tumor growth in human xenograft models of prostate and breast cancer. *Oncotarget*. 2013;4:2451–61. doi: 10.18632/oncotarget.1586

132 Melnik BC, John S, Carrera-Bastos P, *et al.* The impact of cow's milk-mediated mTORC1-signaling in the initiation and progression of prostate cancer. *Nutr Metab (Lond)*. 2012;9:74. doi: 10.1186/1743-7075-9-74

133 Labuschagne CF, van den Broek NJF, Mackay GM, *et al.* Serine, but Not Glycine, Supports One-Carbon Metabolism and Proliferation of Cancer Cells. *Cell Reports*. 2014;7:1248–58. doi: 10.1016/j.celrep.2014.04.045

134 Falegan OS, Jarvi K, Vogel HJ, *et al.* Seminal plasma metabolomics reveals lysine and serine dysregulation as unique features distinguishing between prostate cancer tumors of Gleason grades 6 and 7. *The Prostate*. 2021;81:713–20. doi: 10.1002/pros.24145

135 Pearce N. What Does the Odds Ratio Estimate in a Case-Control Study? *Int J Epidemiol*. 1993;22:1189–92. doi: 10.1093/ije/22.6.1189

136 Kelly RS, Vander Heiden MG, Giovannucci E, *et al.* Metabolomic Biomarkers of Prostate Cancer: Prediction, Diagnosis, Progression, Prognosis, and Recurrence. *Cancer Epidemiology, Biomarkers & Prevention*. 2016;25:887–906. doi: 10.1158/1055-9965.EPI-15-1223

137 Lima AR, Bastos MDL, Carvalho M, *et al.* Biomarker Discovery in Human Prostate Cancer: an Update in Metabolomics Studies. *Translational Oncology*. 2016;9:357–70. doi: 10.1016/j.tranon.2016.05.004

138 Dudka I, Thysell E, Lundquist K, *et al.* Comprehensive metabolomics analysis of prostate cancer tissue in relation to tumor aggressiveness and TMPRSS2-ERG fusion status. *BMC Cancer*. 2020;20:437. doi: 10.1186/s12885-020-06908-z

139 Lokhov PG, Dashtiev MI, Moshkovskii SA, *et al.* Metabolite profiling of blood plasma of patients with prostate cancer. *Metabolomics*. 2010;6:156–63. doi: 10.1007/s11306-009-0187-x

140 Strmiska V, Michalek P, Eckschlager T, *et al.* Prostate cancer-specific hallmarks of amino acids metabolism: Towards a paradigm of precision medicine. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*. 2019;1871:248–58. doi: 10.1016/j.bbcan.2019.01.001

141 Thapar R, Titus M. Recent Advances in Metabolic Profiling and Imaging of Prostate Cancer. *CMB*. 2014;2:53–69. doi: 10.2174/2213235X02666140301002510

142 Lloyd SM, Arnold J, Sreekumar A. Metabolomic profiling of hormone-dependent cancers: a bird's eye view. *Trends in Endocrinology & Metabolism*. 2015;26:477–85. doi: 10.1016/j.tem.2015.07.001

143 Miyagi Y, Higashiyama M, Gochi A, *et al.* Plasma Free Amino Acid Profiling of Five Types of Cancer Patients and Its Application for Early Detection. *PLoS ONE*. 2011;6:e24143. doi: 10.1371/journal.pone.0024143

144 Fages A, Ferrari P, Monni S, *et al.* Investigating sources of variability in metabolomic data in the EPIC study: the Principal Component Partial R-square (PC-PR2) method. *Metabolomics*. 2014;10:1074–83. doi: 10.1007/s11306-014-0647-9

145 Townsend MK, Bao Y, Poole EM, *et al.* Impact of Pre-analytic Blood Sample Collection Factors on Metabolomics. *Cancer Epidemiology, Biomarkers & Prevention*. 2016;25:823–9. doi: 10.1158/1055-9965.EPI-15-1206

146 Sampson JN, Boca SM, Shu XO, *et al.* Metabolomics in Epidemiology: Sources of Variability in Metabolite Measurements and Implications. *Cancer Epidemiology, Biomarkers & Prevention*. 2013;22:631–40. doi: 10.1158/1055-9965.EPI-12-1109

147 Kühn T, Sookthai D, Rolle-Kampczyk U, *et al.* Mid- and long-term correlations of plasma metabolite concentrations measured by a targeted metabolomics approach. *Metabolomics*. 2016;12:184. doi: 10.1007/s11306-016-1133-3

148 Carayol M, Licaj I, Achaintre D, *et al.* Reliability of Serum Metabolites over a Two-Year Period: A Targeted Metabolomic Approach in Fasting and Non-Fasting Samples from EPIC. *PLoS ONE*. 2015;10:e0135437. doi: 10.1371/journal.pone.0135437

149 Wishart DS. Metabolomics for Investigating Physiological and Pathophysiological Processes. *Physiological Reviews*. 2019;99:1819–75. doi: 10.1152/physrev.00035.2018

150 Schmidt JA, Huybrechts I, Overvad K, *et al.* Protein and amino acid intakes in relation to prostate cancer risk and mortality—A prospective study in the European Prospective Investigation into Cancer and Nutrition. *Cancer Medicine*. 2023;12:4725–38. doi: 10.1002/cam4.5289

151 Dereziński P, Klupczynska A, Sawicki W, *et al.* Amino Acid Profiles of Serum and Urine in Search for Prostate Cancer Biomarkers: a Pilot Study. *Int J Med Sci*. 2017;14:1–12. doi: 10.7150/ijms.15783

152 Watling CZ, Kelly RK, Tong TYN, *et al.* Associations between food group intakes and circulating insulin-like growth factor-I in the UK Biobank: a cross-sectional analysis. *Eur J Nutr*. 2023;62:115–24. doi: 10.1007/s00394-022-02954-4

153 UK Biobank. Nightingale Health completes the analysis of all 500,000 UK Biobank samples creating the world's largest blood biomarker database for chronic disease research and prevention | Nightingale Health. 2024. Available from: https://nightingalehealth.com/news/nightingale-health-completes-the-analysis-of-all-500-000-uk-biobank-samples-creating-the-worlds-largest-blood-biomarker-database-for-chronic-disease-research-and-prevention/ (accessed 11 October 2024).

154 Tsilidis KK, Allen NE, Appleby PN, *et al.* Diabetes mellitus and risk of prostate cancer in the EuropeanProspectiveInvestigation into Cancer and Nutrition: Diabetes and prostate cancer. *Int J Cancer*. 2015;136:372–81. doi: 10.1002/ijc.28989

155 Feng X, Song M, Preston MA, *et al.* The association of diabetes with risk of prostate cancer defined by clinical and molecular features. *Br J Cancer*. 2020;123:657–65. doi: 10.1038/s41416-020-0910-y

156 Ye G-C, Yang Y-X, Luo K-D, *et al.* The association between diabetes mellitus and prostate cancer: a meta-analysis and Mendelian randomization. *Aging (Albany NY)*. 2024;16:9584–98. doi: 10.18632/aging.205886

157 Hackett G. Type 2 Diabetes and Testosterone Therapy. *World J Mens Health*. 2019;37:31–44. doi: 10.5534/wjmh.180027

158 Watts EL, Appleby PN, Perez-Cornago A, *et al.* Low Free Testosterone and Prostate Cancer Risk: A Collaborative Analysis of 20 Prospective Studies. *European Urology*. 2018;74:585–94. doi: 10.1016/j.eururo.2018.07.024

159 Clarke R, Shipley M, Lewington S, *et al.* Underestimation of Risk Associations Due to Regression Dilution in Long-term Follow-up of Prospective Studies. *American Journal of Epidemiology*. 1999;150:341–53. doi: 10.1093/oxfordjournals.aje.a010013

160 Aucoin M, Cooley K, Knee C, *et al.* Fish-Derived Omega-3 Fatty Acids and Prostate Cancer: A Systematic Review. *Integr Cancer Ther*. 2017;16:32–62. doi: 10.1177/1534735416656052

161 Van Dael P. Role of n-3 long-chain polyunsaturated fatty acids in human nutrition and health: review of recent studies and recommendations. *Nutr Res Pract*. 2021;15:137–59. doi: 10.4162/nrp.2021.15.2.137

162 Ahmad F, Cherukuri MK, Choyke PL. Metabolic reprogramming in prostate cancer. *Br J Cancer*. 2021;125:1185–96. doi: 10.1038/s41416-021-01435-5

163 Calder PC. Omega-3 fatty acids and inflammatory processes. *Nutrients*. 2010;2:355–74. doi: 10.3390/nu2030355

164 Jayedi A, Djafarian K, Rezagholizadeh F, *et al.* Fasting blood glucose and risk of prostate cancer: A systematic review and meta-analysis of dose-response. *Diabetes & Metabolism*. 2018;44:320–7. doi: 10.1016/j.diabet.2017.09.004

165 Kumar M, Dutta D, Anne B, *et al.* Testosterone levels and type 2 diabetes in men: current knowledge and clinical implications. *DMSO*. 2014;481. doi: 10.2147/DMSO.S50777

166 Michaud JE, Billups KL, Partin AW. Testosterone and prostate cancer: an evidence-based review of pathogenesis and oncologic risk. *Ther Adv Urol*. 2015;7:378–87. doi: 10.1177/1756287215597633

167 Chiesa ST, Charakida M, Georgiopoulos G, *et al.* Glycoprotein Acetyls: A Novel Inflammatory Biomarker of Early Cardiovascular Risk in the Young. *Journal of the American Heart Association*. 2022;11:e024380. doi: 10.1161/JAHA.121.024380

168 Connelly MA, Otvos JD, Shalaurova I, *et al.* GlycA, a novel biomarker of systemic inflammation and cardiovascular disease risk. *J Transl Med*. 2017;15:219. doi: 10.1186/s12967-017-1321-6

169 Suri GS, Kaur G, Carbone GM, *et al.* Metabolomics in oncology. *Cancer Reports*. 2023;6:e1795. doi: 10.1002/cnr2.1795

170 Cacciatore S, Wium M, Licari C, *et al.* Inflammatory metabolic profile of South African patients with prostate cancer. *Cancer Metab*. 2021;9:29. doi: 10.1186/s40170-021-00265-6

171 Mondul AM, Moore SC, Weinstein SJ, *et al.* Metabolomic analysis of prostate cancer risk in a prospective cohort: The alpha-tocopherol, beta-carotene cancer prevention (ATBC) study. *Intl Journal of Cancer*. 2015;137:2124–32. doi: 10.1002/ijc.29576

172 Giskeødegård GF, Hansen AF, Bertilsson H, *et al.* Metabolic markers in blood can separate prostate cancer from benign prostatic hyperplasia. *Br J Cancer*. 2015;113:1712–9. doi: 10.1038/bjc.2015.411

173 Haferkamp S, Drexler K, Federlin M, *et al.* Extracellular Citrate Fuels Cancer Cell Metabolism and Growth. *Front Cell Dev Biol*. 2020;8:602476. doi: 10.3389/fcell.2020.602476

174 Costello LC, Franklin RB, Narayan P. Citrate in the diagnosis of prostate cancer. *Prostate*. 1999;38:237–45. doi: 10.1002/(sici)1097-0045(19990215)38:3<237::aid-pros8>3.0.co;2-o

175 Drexler K, Schmidt KM, Jordan K, *et al.* Cancer-associated cells release citrate to support tumour metastatic progression. *Life Sci Alliance*. 2021;4:e202000903. doi: 10.26508/lsa.202000903

176 Kühn T, Sookthai D, Graf ME, *et al.* Albumin, bilirubin, uric acid and cancer risk: results from a prospective population-based study. *Br J Cancer*. 2017;117:1572–9. doi: 10.1038/bjc.2017.313

177 Sejima T, Iwamoto H, Masago T, *et al.* Low pre-operative levels of serum albumin predict lymph node metastases and ultimately correlate with a biochemical recurrence of prostate cancer in radical prostatectomy patients. *Cent European J Urol*. 2013;66:126–32. doi: 10.5173/ceju.2013.02.art3

178 Xu K, Yan Y, Cheng C, *et al.* The relationship between serum albumin and prostate-specific antigen: A analysis of the National Health and Nutrition Examination Survey, 2003-2010. *Front Public Health*. 2023;11:1078280. doi: 10.3389/fpubh.2023.1078280

179 Tang Q, Li X, Sun C-R. Predictive value of serum albumin levels on cancer survival: a prospective cohort study. *Front Oncol*. 2024;14:1323192. doi: 10.3389/fonc.2024.1323192

180 Ghasemi Fard S, Wang F, Sinclair AJ, *et al.* How does high DHA fish oil affect health? A systematic review of evidence. *Critical Reviews in Food Science and Nutrition*. 2019;59:1684–727. doi: 10.1080/10408398.2018.1425978

181 Watts EL, Perez-Cornago A, Fensom GK, *et al.* Circulating insulin-like growth factors and risks of overall, aggressive and early-onset prostate cancer: a collaborative analysis of 20 prospective studies and Mendelian randomization analysis. *International Journal of Epidemiology*. 2023;52:71–86. doi: 10.1093/ije/dyac124

182 Eshaghian N, Heidarzadeh-Esfahani N, Akbari H, *et al.* Fish consumption and risk of prostate cancer or its mortality: an updated systematic review and dose-response meta-analysis of prospective cohort studies. *Front Nutr*. 2023;10:1221029. doi: 10.3389/fnut.2023.1221029

183 Capili B, Anastasi JK. Cohort Studies. *AJN, American Journal of Nursing*. 2021;121:45–8. doi: 10.1097/01.NAJ.0000803196.49507.08

184 Elliott P, Peakman TC, on behalf of UK Biobank. The UK Biobank sample handling and storage protocol for the collection, processing and archiving of human blood and urine. *International Journal of Epidemiology*. 2008;37:234–44. doi: 10.1093/ije/dym276

185 Navratilova H, Lanham-New S, Whetton A, *et al.* Associations of Diet with Health Outcomes in the UK Biobank: A Systematic Review. *Nutrients*. 2024;16:523. doi: 10.3390/nu16040523

186 Bradbury KE, Young HJ, Guo W, *et al.* Dietary assessment in UK Biobank: an evaluation of the performance of the touchscreen dietary questionnaire. *J Nutr Sci*. 2018;7:e6. doi: 10.1017/jns.2017.66

187 Fry A, Littlejohns TJ, Sudlow C, *et al.* Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *American Journal of Epidemiology*. 2017;186:1026–34. doi: 10.1093/aje/kwx246

188 Küllenberg D, Taylor LA, Schneider M, *et al.* Health effects of dietary phospholipids. *Lipids Health Dis*. 2012;11:3. doi: 10.1186/1476-511X-11-3

189 McMullan JE, Yeates AJ, Allsopp PJ, *et al.* Fish consumption and its lipid modifying effects - A review of intervention studies. *Neurotoxicology*. 2023;99:82–96. doi: 10.1016/j.neuro.2023.10.003

190 Larsson SC, Wolk K, Brismar K, *et al.* Association of diet with serum insulin-like growth factor I in middle-aged and elderly men. *The American Journal of Clinical Nutrition*. 2005;81:1163–7. doi: 10.1093/ajcn/81.5.1163

191 van Roekel E, Trijsburg L, Assi N, *et al.* Circulating Metabolites Associated with Alcohol Intake in the European Prospective Investigation into Cancer and Nutrition Cohort. *Nutrients*. 2018;10:654. doi: 10.3390/nu10050654

192 Langenau J, Boeing H, Bergmann MM, *et al.* The Association between Alcohol Consumption and Serum Metabolites and the Modifying Effect of Smoking. *Nutrients*. 2019;11:2331. doi: 10.3390/nu11102331

193 Li Y, Wang M, Liu X, *et al.* Circulating metabolites may illustrate relationship of alcohol consumption with cardiovascular disease. *BMC Med*. 2023;21:443. doi: 10.1186/s12916-023-03149-2

194 Ottosson F, Brunkwall L, Ericson U, *et al.* Connection Between BMI-Related Plasma Metabolite Profile and Gut Microbiota. *The Journal of Clinical Endocrinology & Metabolism*. 2018;103:1491–501. doi: 10.1210/jc.2017-02114

195 Maltais-Payette I, Bourgault J, Arsenault B, *et al.* Obesity, Body Fat Distribution, and Circulating Glutamate Concentrations, A BI-Directional Mendelian Randomization Study. *Journal of the Endocrine Society*. 2021;5:A23–4. doi: 10.1210/jendso/bvab048.045

196 Ramadani FG, Perdana NR, Ringoringo DRL. Body mass index, obesity and risk of prostate cancer: a systematic review and meta-analysis. *Cent European J Urol*. 2024;77:176–88. doi: 10.5173/ceju.2023.162

197 Lin Y, Yang Z, Li J, *et al.* Effects of glutamate and aspartate on prostate cancer and breast cancer: a Mendelian randomization study. *BMC Genomics*. 2022;23:213. doi: 10.1186/s12864-022-08442-7

198 López-Hernández Y, Andres-Lacueva C, Wishart DS, *et al.* Prostate cancer risk biomarkers from large cohort and prospective metabolomics studies: A systematic review. *Translational Oncology*. 2025;51:102196. doi: 10.1016/j.tranon.2024.102196

199 Gu D, Tang M, Wang Y, *et al.* The Causal Relationships Between Extrinsic Exposures and Risk of Prostate Cancer: A Phenome-Wide Mendelian Randomization Study. *Front Oncol*. 2022;12:829248. doi: 10.3389/fonc.2022.829248

200 Sundaresan VM, Smani S, Rajwa P, *et al.* Prostate-specific antigen screening for prostate cancer: Diagnostic performance, clinical thresholds, and strategies for refinement. *Urologic Oncology: Seminars and Original Investigations*. 2025;43:41–8. doi: 10.1016/j.urolonc.2024.06.003

201 Sen A, Papadimitriou N, Lagiou P, *et al.* Coffee and tea consumption and risk of prostate cancer in the European Prospective Investigation into Cancer and Nutrition. *Intl Journal of Cancer*. 2019;144:240–50. doi: 10.1002/ijc.31634

202 Kheirandish P, Chinegwundoh F. Ethnic differences in prostate cancer. *Br J Cancer*. 2011;105:481–5. doi: 10.1038/bjc.2011.273

203 Hinata N, Fujisawa M. Racial Differences in Prostate Cancer Characteristics and Cancer-Specific Mortality: An Overview. *World J Mens Health*. 2022;40:217. doi: 10.5534/wjmh.210070

204 Office for National Statistics. Population by ethnicity, England and Wales: Census 2001-2021. 2022. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates (accessed 30 Dec 2024).

205 Perez-Cornago A, Dunneram Y, Watts EL, *et al.* Adiposity and risk of prostate cancer death: a prospective analysis in UK Biobank and meta-analysis of published studies. *BMC Med*. 2022;20:143. doi: 10.1186/s12916-022-02336-x

206 Szymanski KM, Wheeler DC, Mucci LA. Fish consumption and prostate cancer risk: a review and meta-analysis. *The American Journal of Clinical Nutrition*. 2010;92:1223–33. doi: 10.3945/ajcn.2010.29530

207 Roberts LD, Souza AL, Gerszten RE, *et al.* Targeted Metabolomics. *CP Molecular Biology*. 2012;98. doi: 10.1002/0471142727.mb3002s98

208 Ribbenstedt A, Ziarrusta H, Benskin JP. Development, characterization and comparisons of targeted and non-targeted metabolomics methods. *PLoS ONE*. 2018;13:e0207082. doi: 10.1371/journal.pone.0207082

209 Vinayavekhin N, Saghatelian A. Untargeted Metabolomics. *CP Molecular Biology*. 2010;90. doi: 10.1002/0471142727.mb3001s90

210 Gertsman I, Barshop BA. Promises and pitfalls of untargeted metabolomics. *J of Inher Metab Disea*. 2018;41:355–66. doi: 10.1007/s10545-017-0130-7

211 Schrimpe-Rutledge AC, Codreanu SG, Sherrod SD, *et al.* Untargeted Metabolomics Strategies—Challenges and Emerging Directions. *J Am Soc Mass Spectrom*. 2016;27:1897–905. doi: 10.1007/s13361-016-1469-y

212 Yin X, Prendiville O, McNamara AE, *et al.* Targeted Metabolomic Approach to Assess the Reproducibility of Plasma Metabolites over a Four Month Period in a Free-Living Population. *J Proteome Res*. 2022;21:683–90. doi: 10.1021/acs.jproteome.1c00440

213 Floegel A, Stefan N, Yu Z, *et al.* Identification of Serum Metabolites Associated With Risk of Type 2 Diabetes Using a Targeted Metabolomic Approach. *Diabetes*. 2013;62:639–48. doi: 10.2337/db12-0495

214 Stevens VL, Hoover E, Wang Y, *et al.* Pre-Analytical Factors that Affect Metabolite Stability in Human Urine, Plasma, and Serum: A Review. *Metabolites*. 2019;9:156. doi: 10.3390/metabo9080156

215 Lozier BK, Kim RN, Zuromski LM, *et al.* Effect of fasting status and other pre-analytical variables on quantitation of long-chain fatty acids in red blood cells. *Prostaglandins, Leukotrienes and Essential Fatty Acids*. 2020;163:102211. doi: 10.1016/j.plefa.2020.102211

216 Taylor H, Lewins M, Foody MGB, *et al.* UK Biobank—A Unique Resource for Discovery and Translation Research on Genetics and Neurologic Disease. *Neurol Genet*. 2025;11:e200226. doi: 10.1212/NXG.0000000000200226

217 Salmon C, Song L, Muir K, *et al.* Marital status and prostate cancer incidence: a pooled analysis of 12 case–control studies from the PRACTICAL consortium. *Eur J Epidemiol*. 2021;36:913–25. doi: 10.1007/s10654-021-00781-1

218 Coughlin SS. A review of social determinants of prostate cancer risk, stage, and survival. *Prostate Int*. 2020;8:49–54. doi: 10.1016/j.prnil.2019.08.001

219 McTiernan A, Friedenreich CM, Katzmarzyk PT, *et al.* Physical Activity in Cancer Prevention and Survival: A Systematic Review. *Med Sci Sports Exerc*. 2019;51:1252–61. doi: 10.1249/MSS.0000000000001937

220 Cui H, Zhang W, Zhang L, *et al.* Risk factors for prostate cancer: An umbrella review of prospective observational studies and mendelian randomization analyses. *PLoS Med*. 2024;21:e1004362. doi: 10.1371/journal.pmed.1004362

221 Pietzner M, Stewart ID, Raffler J, *et al.* Plasma metabolites to profile pathways in noncommunicable disease multimorbidity. *Nat Med*. 2021;27:471–9. doi: 10.1038/s41591-021-01266-0

222 Al-Fayez S, El-Metwally A. Cigarette smoking and prostate cancer: A systematic review and meta-analysis of prospective cohort studies. *Tob Induc Dis*. 2023;21:19. doi: 10.18332/tid/157231

223 Hong S, Khil H, Lee DH, *et al.* Alcohol Consumption and the Risk of Prostate Cancer: A Dose-Response Meta-Analysis. *Nutrients*. 2020;12:2188. doi: 10.3390/nu12082188

224 Robinson O, Carter AR, Ala-Korpela M, *et al.* Metabolic profiles of socio-economic position: a multi-cohort analysis. *International Journal of Epidemiology*. 2021;50:768–82. doi: 10.1093/ije/dyaa188

225 Littlejohns TJ, Travis RC, Key TJ, *et al.* Lifestyle factors and prostate-specific antigen (PSA) testing in UK Biobank: Implications for epidemiological research. *Cancer Epidemiology*. 2016;45:40–6. doi: 10.1016/j.canep.2016.09.010

226 Ilic D, Djulbegovic M, Jung JH, *et al.* Prostate cancer screening with prostate-specific antigen (PSA) test: a systematic review and meta-analysis. *BMJ*. 2018;k3519. doi: 10.1136/bmj.k3519

227 Giunchi F, Fiorentino M, Loda M. The Metabolic Landscape of Prostate Cancer. *European Urology Oncology*. 2019;2:28–36. doi: 10.1016/j.euo.2018.06.010

228 Andersen MK, Giskeødegård GF, Tessem M-B. Metabolic alterations in tissues and biofluids of patients with prostate cancer. *Current Opinion in Endocrine and Metabolic Research*. 2020;10:23–8. doi: 10.1016/j.coemr.2020.02.003

229 Penney KL, Tyekucheva S, Rosenthal J, *et al.* Metabolomics of Prostate Cancer Gleason Score in Tumor Tissue and Serum. *Molecular Cancer Research*. 2021;19:475–84. doi: 10.1158/1541-7786.MCR-20-0548

230 Saoi M, Britz-McKibbin P. New Advances in Tissue Metabolomics: A Review. *Metabolites*. 2021;11:672. doi: 10.3390/metabo11100672

231 Richmond RC, Davey Smith G. Mendelian Randomization: Concepts and Scope. *Cold Spring Harb Perspect Med*. 2022;12:a040501. doi: 10.1101/cshperspect.a040501

232 Feng C, Li H, Zhang C, *et al.* Exploring the causal role of plasma metabolites and metabolite ratios in prostate cancer: a two-sample Mendelian randomization study. *Front Mol Biosci*. 2025;11:1406055. doi: 10.3389/fmolb.2024.1406055

233 Lotta LA, Pietzner M, Stewart ID, *et al.* A cross-platform approach identifies genetic regulators of human metabolism and health. *Nat Genet*. 2021;53:54–64. doi: 10.1038/s41588-020-00751-5

234 Yu B, Zanetti KA, Temprosa M, *et al.* The Consortium of Metabolomics Studies (COMETS): Metabolomics in 47 Prospective Cohort Studies. *American Journal of Epidemiology*. 2019;188:991–1012. doi: 10.1093/aje/kwz028

235 Desai TA, Hedman ÅK, Dimitriou M, *et al.* Identifying proteomic risk factors for overall, aggressive, and early onset prostate cancer using Mendelian Randomisation and tumour spatial transcriptomics. *eBioMedicine*. 2024;105:105168. doi: 10.1016/j.ebiom.2024.105168

236 Sadeesh N, Scaravilli M, Latonen L. Proteomic Landscape of Prostate Cancer: The View Provided by Quantitative Proteomics, Integrative Analyses, and Protein Interactomes. *Cancers*. 2021;13:4829. doi: 10.3390/cancers13194829

Appendix

List of Appendices

Appendix 3.1 Individual metabolites loading on each metabolite pattern

Metabolite Pattern	Contributing Metabolites
1	**Diacyl-alkyl-phosphatidylcholines:** PC aa C28:1 PC aa C30:0 PC aa C32:0 PC aa C32:1 PC aa C32:3 PC aa C34:1 PC aa C34:2 PC aa C34:3 PC aa C34:4 PC aa C36:0 PC aa C36:1 PC aa C36:2 PC aa C36:3 PC aa C36:4 PC aa C36:5 PC aa C36:6 PC aa C38:0 PC aa C38:3 PC aa C38:4 PC aa C38:5 PC aa C38:6 PC aa C40:2 PC aa C40:3 PC aa C40:4 PC aa C40:5 PC aa C40:6 PC aa C42:0 PC aa C42:1 PC aa C42:2 PC aa C42:4 PC aa C42:5 **Acyl-alkyl-phosphatidylcholines:** PC ae C30:0 PC ae C30:2 PC ae C32:1 PC ae C32:2 PC ae C34:0 PC ae C34:1 PC ae C34:2 PC ae C34:3 PC ae C36:0 PC ae C36:1 PC ae C36:2 PC ae C36:3 PC ae C36:4 PC ae C36:5 PC ae C38:2 PC ae C38:3

	PC ae C38:4 PC ae C38:5 PC ae C38:6 PC ae C40:1 PC ae C40:2 PC ae C40:3 PC ae C40:4 PC ae C40:5 PC ae C40:6 PC ae C42:1 PC ae C42:2 PC ae C42:3 PC ae C42:4 PC ae C42:5 PC ae C44:4 PC ae C44:5 PC ae C44:6 **Hydroxysphingomyelins:** SM (OH) C14:1 SM (OH) C16:1 SM (OH) C22:2
2	**Acylcarnitines:** C18:1 C18:2 **Amino acids:** Glutamate Ornithine **Biogenic Amine:** Taurine
3	**Lysophosphatidylcholines:** Lyso PC a C16:0 Lyso PC a C16:1 Lyso PC a C17:0 Lyso PC a C18:0 Lyso PC a C18:1 Lyso PC a C18:2 Lyso PC a C20:3 Lyso PC a C20:4
Individual metabolites loading on each metabolite pattern.	

Appendix 3.2 Instruments used to quantify metabolites in the study population

Method	Instrument	Study
Triple quadrupole mass spectrometry (MS)	Triple Quad 4500, AB Sciex, Framingham, MA	Prostate Colorectal
	Q-Trap 5500, AB Sciex, Framingham, MA	Liver Kidney
Liquid chromatography (LC)	Agilent 1290, Agilent Technologies, Santa Clara, CA	Prostate Colorectal Liver Kidney

Appendix 3.3 Associations of dietary exposures and BMI with metabolite patterns in the discovery and validation sets						
Metabolite pattern	**Exposure**	**β**	**P-value**	**Padj**	**β**	**P-value**
1	**Alcohol**	**0.061**	**8.00E-11**	**1.90E-09**	**0.044**	**3.50E-02**
	Butter	0.047	2.40E-06	2.90E-05	0.026	2.60E-01
	Total fish	**0.076**	**4.20E-06**	**3.30E-05**	**0.12**	**3.80E-04**
	Total fish products	**0.067**	**1.00E-05**	**4.80E-05**	**0.12**	**3.60E-04**
	Margarine	-0.033	1.00E-05	4.80E-05	-0.0097	5.80E-01
	Fatty fish	0.043	4.50E-04	1.80E-03	0.049	7.90E-02
	Lean fish	**0.026**	**1.00E-03**	**3.50E-03**	**0.044**	**8.70E-03**
	Cereals	-0.11	3.20E-03	9.50E-03	-0.1	1.20E-01
	BMI	-0.015	1.50E-02	4.10E-02	0.0088	5.20E-01
	Dairy	0.041	3.40E-02	8.10E-02		
	Cheese	0.035	6.90E-02	1.50E-01		
	Red meat	0.04	8.40E-02	1.70E-01		
	Milk	0.034	1.10E-01	2.00E-01		
	Vegetables (fruiting)	0.046	2.10E-01	3.70E-01		
	Yogurt	0.011	2.30E-01	3.70E-01		
	Fruit	0.011	3.10E-01	4.70E-01		
	Eggs	-0.0055	5.10E-01	7.20E-01		
	Vegetable oils	0.0044	5.70E-01	7.20E-01		
	Vegetables	0.0092	5.60E-01	7.20E-01		
	Vegetables (leafy)	-0.0055	6.20E-01	7.40E-01		
	Fats	-0.0042	7.60E-01	8.60E-01		
	Vegetables (root)	-0.006	7.90E-01	8.60E-01		
	Poultry	-0.0038	8.50E-01	8.80E-01		
	Processed meat	0.0022	9.10E-01	9.10E-01		
2	**BMI**	**0.039**	**5.90E-11**	**1.40E-09**	**0.046**	**1.00E-03**
	Butter	-0.049	5.40E-07	6.40E-06	-0.046	5.50E-02
	Dairy	-0.078	5.00E-05	3.60E-04	-0.057	1.70E-01
	Margarine	0.03	6.00E-05	3.60E-04	0.007	6.90E-01
	Milk	-0.073	4.80E-04	2.30E-03	-0.051	2.60E-01
	Alcohol	0.029	1.80E-03	7.30E-03	-0.015	4.80E-01
	Cereals	0.09	1.30E-02	4.30E-02	0.11	1.00E-01
	Processed meat	0.028	1.60E-01	4.30E-01		
	Fats	0.019	1.50E-01	4.30E-01		
	Yogurt	-0.01	2.40E-01	5.00E-01		
	Poultry	0.024	2.20E-01	5.00E-01		
	Vegetable oils	0.0086	2.50E-01	5.00E-01		
	Vegetables (leafy)	-0.012	2.80E-01	5.20E-01		
	Cheese	0.016	4.10E-01	6.10E-01		
	Total fish products	-0.013	3.90E-01	6.10E-01		
	Total fish	-0.013	4.40E-01	6.10E-01		
	Lean fish	-0.006	4.50E-01	6.10E-01		

Appendix 3.3 Associations of dietary exposures and BMI with metabolite patterns in the discovery and validation sets

Metabolite pattern	Exposure	β	P-value	Padj	β	P-value
	Vegetables	-0.013	4.00E-01	6.10E-01		
	Vegetables (fruiting)	-0.023	5.40E-01	6.40E-01		
	Fruit	0.0066	5.30E-01	6.40E-01		
	Fatty fish	-0.0067	5.90E-01	6.70E-01		
	Vegetables (root)	0.008	7.20E-01	7.80E-01		
	Red meat	0.0052	8.20E-01	8.60E-01		
	Eggs	-0.0011	8.90E-01	8.90E-01		
3	**BMI**	**-0.068**	**5.70E-32**	**1.40E-30**	**-0.036**	**7.40E-03**
	Alcohol	0.044	6.90E-07	8.30E-06	0.025	2.20E-01
	Fatty fish	**-0.041**	**4.10E-04**	**3.30E-03**	**-0.071**	**9.70E-03**
	Total fish products	-0.041	4.10E-03	2.40E-02	-0.026	4.20E-01
	Eggs	0.021	8.90E-03	3.80E-02	0.015	4.40E-01
	Total fish	-0.041	9.60E-03	3.80E-02	-0.028	4.30E-01
	Butter	0.016	8.10E-02	2.80E-01		
	Margarine	-0.011	1.20E-01	3.20E-01		
	Vegetables (fruiting)	0.055	1.20E-01	3.20E-01		
	Cereals	-0.05	1.50E-01	3.60E-01		
	Milk	0.026	1.90E-01	3.90E-01		
	Poultry	-0.025	1.80E-01	3.90E-01		
	Dairy	0.021	2.50E-01	4.30E-01		
	Lean fish	-0.0085	2.40E-01	4.30E-01		
	Cheese	0.015	4.10E-01	5.40E-01		
	Processed meat	0.016	4.10E-01	5.40E-01		
	Vegetables	0.013	3.90E-01	5.40E-01		
	Vegetables (leafy)	0.0098	3.50E-01	5.40E-01		
	Vegetables (root)	-0.016	4.40E-01	5.60E-01		
	Red meat	0.0084	7.10E-01	8.20E-01		
	Vegetable oils	0.0026	7.20E-01	8.20E-01		
	Yogurt	-0.0022	8.00E-01	8.70E-01		
	Fats	-0.0014	9.10E-01	9.10E-01		
	Fruit	0.0016	8.80E-01	9.10E-01		

Abbreviations: BMI, body mass index

All models were adjusted for age at blood collection, time of day of blood collection, energy intake (kcal/day), fasting status, education level, physical activity, smoking status, baseline alcohol consumption, and body mass index (BMI). Models that examined alcohol intake and BMI as main exposures were not adjusted for alcohol intake and BMI, respectively. Variables are ordered by lowest Padj value by metabolite pattern.

To account for multiple testing, only dietary correlates of metabolite patterns that passed the false-discovery rate of 0.05 (Padj < 0.05), using the Benjamini-Hochberg method, were further analysed in the validation set. Associations with a P-value < 0.05 in the validation set are bolded.

Appendix 3.3 Associations of dietary exposures and BMI with metabolite patterns in the discovery and validation sets

Metabolite pattern	Exposure	β	P-value	Padj	β	P-value

β is per one unit described in Table 3.1 (excluding BMI, in which β is per one unit kg/m^2).

Appendix 3.4 Associations of validated dietary exposures and BMI with individual metabolites loading on each metabolite pattern

Metabolite pattern	Exposure	Metabolite	β	P-value	Padj
1	Alcohol	PC aa C32:1	0.082	5.80E-91	1.50E-88
	Alcohol	PC aa C34:1	0.038	4.10E-68	5.50E-66
	Alcohol	PC aa C36:4	0.029	3.50E-53	3.10E-51
	Total fish products	PC aa C42:2	0.048	1.10E-47	7.00E-46
	Total fish	PC aa C42:2	0.052	1.50E-46	8.30E-45
	Alcohol	PC ae C36:0	0.03	2.60E-46	1.20E-44
	Total fish products	PC aa C38:0	0.049	3.10E-46	1.20E-44
	Alcohol	PC aa C32:0	0.023	9.90E-46	3.30E-44
	Total fish	PC aa C38:0	0.052	8.60E-44	2.60E-42
	Total fish products	PC aa C38:6	0.054	3.70E-43	9.90E-42
	Alcohol	PC ae C36:2	-0.026	1.90E-41	4.70E-40
	Total fish	PC aa C38:6	0.057	2.20E-39	4.80E-38
	Total fish products	PC aa C36:0	0.05	2.10E-38	4.40E-37
	Total fish	PC aa C36:0	0.054	4.00E-36	7.60E-35
	Alcohol	PC aa C36:5	0.049	2.60E-32	4.70E-31
	Total fish products	PC aa C40:6	0.049	1.10E-31	1.90E-30
	Alcohol	SM(OH) C16:1	-0.023	7.80E-29	1.20E-27
	Total fish products	PC aa C36:6	0.057	9.70E-29	1.40E-27
	Total fish	PC aa C40:6	0.052	9.50E-29	1.40E-27
	Alcohol	PC aa C30:0	0.029	1.40E-28	1.70E-27
	Alcohol	PC aa C34:4	0.03	1.30E-28	1.70E-27
	Total fish products	PC aa C36:5	0.071	4.70E-28	5.70E-27
	Total fish products	PC ae C40:6	0.034	7.10E-28	8.20E-27
	Total fish	PC aa C36:6	0.061	1.80E-27	2.00E-26
	Total fish	PC ae C40:6	0.037	3.30E-27	3.50E-26
	Total fish	PC aa C36:5	0.076	4.80E-27	5.00E-26
	Alcohol	PC ae C36:5	0.022	6.10E-27	6.10E-26
	Lean fish	PC aa C42:2	0.018	1.90E-26	1.90E-25
	Alcohol	SM(OH) C14:1	-0.022	1.10E-25	1.00E-24
	Alcohol	PC aa C36:1	0.024	1.40E-25	1.20E-24
	Total fish products	PC ae C38:6	0.034	1.80E-25	1.60E-24
	Alcohol	SM(OH) C22:2	-0.019	5.50E-25	4.60E-24
	Alcohol	PC aa C38:5	0.022	7.00E-25	5.70E-24
	Total fish	PC ae C42:3	0.032	9.50E-25	7.50E-24

	Total fish products	**PC ae C36:0**	**0.033**	**3.10E-24**	**2.40E-23**
	Total fish products	**PC ae C42:3**	**0.029**	**3.30E-24**	**2.50E-23**
	Total fish	**PC ae C38:6**	**0.036**	**5.10E-24**	**3.70E-23**
	Total fish	**PC aa C42:0**	**0.036**	**2.60E-23**	**1.80E-22**
	Total fish	**PC aa C42:1**	**0.036**	**4.00E-23**	**2.80E-22**
	Total fish products	**PC aa C42:0**	**0.033**	**1.90E-22**	**1.20E-21**
	Total fish	**PC ae C36:0**	**0.035**	**2.90E-22**	**1.90E-21**
	Total fish products	**PC aa C42:1**	**0.032**	**3.60E-22**	**2.30E-21**
	Alcohol	**PC aa C34:3**	**0.021**	**4.20E-21**	**2.60E-20**
	Alcohol	**PC aa C40:4**	**0.02**	**3.10E-20**	**1.90E-19**
	Alcohol	**PC ae C30:2**	**-0.017**	**4.90E-19**	**2.90E-18**
	Alcohol	**PC ae C38:3**	**-0.016**	**1.00E-18**	**5.80E-18**
	Alcohol	**PC aa C38:6**	**0.021**	**3.00E-17**	**1.70E-16**
	Lean fish	**PC aa C38:0**	**0.015**	**5.70E-17**	**3.20E-16**
	Total fish products	**PC aa C40:4**	**-0.029**	**6.30E-17**	**3.40E-16**
	Alcohol	**PC aa C40:5**	**0.019**	**3.10E-16**	**1.60E-15**
	Total fish	**PC aa C40:4**	**-0.03**	**2.40E-15**	**1.30E-14**
	Total fish products	**PC ae C40:1**	**0.024**	**3.00E-15**	**1.50E-14**
	Total fish	**PC ae C40:1**	**0.027**	**3.00E-15**	**1.50E-14**
	Total fish products	**PC ae C42:2**	**0.024**	**3.40E-15**	**1.70E-14**
	Total fish	**PC ae C42:2**	**0.026**	**6.10E-15**	**3.00E-14**
	Total fish	**PC ae C32:2**	**0.024**	**2.90E-14**	**1.40E-13**
	Total fish products	**PC ae C32:2**	**0.021**	**4.90E-14**	**2.30E-13**
	Total fish products	**PC aa C40:2**	**0.027**	**9.20E-14**	**4.20E-13**
	Alcohol	**PC aa C36:6**	**0.024**	**9.70E-14**	**4.40E-13**
	Lean fish	**PC aa C36:0**	**0.015**	**2.40E-13**	**1.10E-12**
	Lean fish	**PC ae C36:0**	**0.012**	**1.20E-12**	**5.30E-12**
	Alcohol	**PC aa C38:4**	**0.015**	**2.30E-12**	**1.00E-11**
	Total fish	**PC aa C40:2**	**0.028**	**2.50E-12**	**1.10E-11**
	Alcohol	**PC aa C34:2**	**0.009**	**4.00E-12**	**1.70E-11**
	Total fish products	**PC ae C40:2**	**0.022**	**4.10E-12**	**1.70E-11**
	Lean fish	**PC ae C42:3**	**0.01**	**4.90E-12**	**2.00E-11**
	Total fish	**PC ae C40:2**	**0.024**	**7.70E-12**	**3.10E-11**
	Total fish products	**PC ae C38:4**	**-0.019**	**8.20E-12**	**3.20E-11**
	Lean fish	**PC aa C38:6**	**0.014**	**1.20E-11**	**4.80E-11**
	Alcohol	**PC ae C40:6**	**-0.013**	**1.70E-11**	**6.50E-11**
	Lean fish	**PC ae C40:6**	**0.011**	**7.70E-11**	**2.90E-10**
	Lean fish	**PC aa C42:1**	**0.011**	**9.20E-11**	**3.40E-10**

	Alcohol	PC ae C40:1	0.012	2.80E-10	1.00E-09
	Total fish	PC ae C38:4	-0.019	4.10E-10	1.50E-09
	Alcohol	PC ae C32:1	0.01	9.00E-10	3.20E-09
	Alcohol	PC ae C34:2	-0.011	1.00E-09	3.60E-09
	Alcohol	PC aa C36:3	0.01	1.10E-09	3.80E-09
	Lean fish	PC ae C32:2	0.009	1.50E-09	5.10E-09
	Alcohol	PC ae C40:3	-0.009	1.80E-09	6.00E-09
	Lean fish	PC ae C38:6	0.01	3.10E-09	1.10E-08
	Alcohol	PC aa C32:3	-0.011	4.60E-09	1.50E-08
	Lean fish	PC aa C40:6	0.013	5.00E-09	1.60E-08
	Alcohol	PC ae C38:2	-0.01	1.90E-08	6.10E-08
	Alcohol	PC ae C42:1	0.009	1.90E-08	6.10E-08
	Lean fish	PC aa C36:6	0.015	2.40E-08	7.50E-08
	Lean fish	PC ae C40:1	0.009	2.40E-08	7.50E-08
	Lean fish	PC aa C42:0	0.01	2.70E-08	8.30E-08
	Lean fish	PC aa C36:5	0.019	3.60E-08	1.10E-07
	Total fish products	PC aa C40:3	0.019	6.30E-08	1.90E-07
	Total fish products	PC aa C38:4	-0.018	8.70E-08	2.60E-07
	Lean fish	PC ae C42:2	0.008	9.60E-08	2.80E-07
	Alcohol	PC aa C42:2	0.011	1.40E-07	4.00E-07
	Total fish	PC aa C38:4	-0.019	1.70E-07	4.90E-07
	Total fish	PC aa C40:3	0.02	2.70E-07	7.70E-07
	Total fish	PC ae C34:0	0.02	5.50E-07	1.50E-06
	Alcohol	PC ae C36:4	0.01	6.20E-07	1.70E-06
	Alcohol	PC ae C32:2	0.009	1.40E-06	3.80E-06
	Total fish products	PC aa C36:4	-0.014	1.50E-06	4.10E-06
	Lean fish	PC ae C40:2	0.008	2.40E-06	6.50E-06
	Total fish products	PC ae C34:0	0.017	4.20E-06	1.10E-05
	Total fish	PC aa C36:4	-0.015	4.60E-06	1.20E-05
	Total fish products	PC aa C42:5	0.017	1.00E-05	2.50E-05
	Total fish products	PC ae C40:4	-0.011	1.00E-05	2.50E-05
	Lean fish	PC aa C40:2	0.0081	1.00E-05	2.50E-05
	Lean fish	PC ae C34:0	0.0086	1.00E-05	2.50E-05
	Alcohol	PC ae C38:6	0.0091	1.00E-05	2.50E-05
	Alcohol	PC ae C38:5	0.007	2.00E-05	5.00E-05
	Alcohol	PC ae C40:4	-0.0068	2.00E-05	5.00E-05
	Total fish	PC aa C42:5	0.017	3.00E-05	7.40E-05
	Total fish	PC ae C30:2	0.013	4.00E-05	9.70E-05
	Alcohol	PC aa C38:3	0.0083	4.00E-05	9.70E-05
	Alcohol	PC aa C42:4	0.0072	7.00E-05	1.70E-04
	Lean fish	PC ae C30:2	0.006	9.00E-05	2.10E-04

	Alcohol	**PC aa C42:0**	**-0.0082**	**1.00E-04**	**2.40E-04**
	Total fish products	**PC ae C36:4**	**-0.012**	**1.20E-04**	**2.80E-04**
	Total fish products	**PC ae C30:2**	**0.011**	**1.30E-04**	**3.00E-04**
	Lean fish	**PC ae C40:5**	**0.005**	**1.40E-04**	**3.20E-04**
	Alcohol	**PC ae C38:4**	**-0.0066**	**1.60E-04**	**3.60E-04**
	Total fish	**PC ae C40:4**	**-0.01**	**2.00E-04**	**4.50E-04**
	Alcohol	**PC ae C42:4**	**-0.0081**	**2.50E-04**	**5.60E-04**
	Total fish	**PC aa C38:5**	**0.014**	**2.70E-04**	**5.90E-04**
	Alcohol	**PC ae C42:5**	**-0.0058**	**2.70E-04**	**5.90E-04**
	Alcohol	**PC aa C40:6**	**0.0097**	**2.80E-04**	**6.10E-04**
	Total fish products	**PC aa C38:5**	**0.012**	**3.80E-04**	**8.20E-04**
	Total fish	**PC ae C36:4**	**-0.012**	**4.00E-04**	**8.60E-04**
	Lean fish	**PC aa C40:4**	**-0.0062**	**4.70E-04**	**1.00E-03**
	Total fish products	**PC aa C36:3**	**-0.0094**	**5.10E-04**	**1.10E-03**
	Lean fish	**PC aa C40:3**	**0.0061**	**5.60E-04**	**1.20E-03**
	Alcohol	**PC ae C34:3**	**0.0073**	**8.50E-04**	**1.80E-03**
	Alcohol	**PC ae C36:3**	**-0.0058**	**1.00E-03**	**2.10E-03**
	Total fish	**PC aa C36:3**	**-0.0094**	**1.40E-03**	**2.90E-03**
	Total fish	**SM(OH) C22:2**	**0.0099**	**1.80E-03**	**3.70E-03**
	Alcohol	**PC aa C42:5**	**0.0073**	**1.80E-03**	**3.70E-03**
	Lean fish	**SM(OH) C22:2**	**0.0047**	**1.80E-03**	**3.70E-03**
	Alcohol	**PC aa C36:0**	**0.0076**	**2.10E-03**	**4.10E-03**
	Alcohol	**PC ae C36:1**	**-0.0061**	**2.50E-03**	**4.80E-03**
	Alcohol	**PC ae C42:2**	**0.0058**	**2.50E-03**	**4.90E-03**
	Lean fish	**PC ae C30:0**	**0.0058**	**2.60E-03**	**5.00E-03**
	Total fish products	**SM(OH) C22:2**	**0.0087**	**2.70E-03**	**5.30E-03**
	Alcohol	**PC ae C30:0**	**-0.0069**	**3.10E-03**	**6.00E-03**
	Total fish	**PC ae C30:0**	**0.012**	**3.60E-03**	**6.90E-03**
	Total fish	**PC ae C38:2**	**0.009**	**3.70E-03**	**6.90E-03**
	Total fish products	**PC ae C42:4**	**-0.0102**	**4.00E-03**	**7.60E-03**
	Total fish products	**PC aa C34:3**	**-0.01**	**4.50E-03**	**8.40E-03**
	Lean fish	**PC ae C36:1**	**0.0048**	**4.90E-03**	**9.10E-03**
	Lean fish	**PC aa C32:3**	**0.0042**	**5.20E-03**	**9.50E-03**
	Lean fish	**PC ae C40:3**	**0.0035**	**5.70E-03**	**1.00E-02**
	Total fish	**PC ae C44:6**	**0.01**	**5.80E-03**	**1.10E-02**
	Total fish	**PC ae C40:5**	**0.0076**	**6.10E-03**	**1.10E-02**
	Lean fish	**PC aa C36:4**	**-0.0041**	**6.10E-03**	**1.10E-02**
	Alcohol	**PC ae C40:5**	**-0.0043**	**6.20E-03**	**1.10E-02**
	Total fish products	**PC aa C40:5**	**-0.01**	**7.20E-03**	**1.30E-02**
	Total fish	**PC ae C32:1**	**0.0078**	**7.60E-03**	**1.30E-02**

	Total fish products	**PC ae C38:2**	**0.0075**	**7.90E-03**	**1.40E-02**
	Lean fish	**PC ae C38:2**	**0.0039**	**8.00E-03**	**1.40E-02**
	Lean fish	**PC aa C38:4**	**-0.0045**	**8.10E-03**	**1.40E-02**
	Total fish	**PC aa C34:3**	**-0.0097**	**1.20E-02**	**2.10E-02**
	Total fish	**PC aa C40:5**	**-0.01**	**1.20E-02**	**2.10E-02**
	Lean fish	**PC ae C32:1**	**0.0035**	**1.30E-02**	**2.10E-02**
	Total fish products	**PC aa C38:3**	**-0.0079**	**1.40E-02**	**2.40E-02**
	Lean fish	**PC aa C28:1**	**0.00412**	**1.50E-02**	**2.40E-02**
	Total fish products	**PC ae C30:0**	**0.0091**	**1.60E-02**	**2.60E-02**
	Total fish	**PC ae C36:1**	**0.0085**	**1.60E-02**	**2.60E-02**
	Total fish	**PC aa C32:0**	**0.0068**	**1.70E-02**	**2.80E-02**
	Total fish	**PC aa C28:1**	**0.0084**	**1.70E-02**	**2.80E-02**
	Total fish products	**PC ae C40:5**	**0.006**	**1.80E-02**	**2.80E-02**
	Alcohol	**PC ae C40:2**	**-0.0047**	**1.80E-02**	**3.00E-02**
	Lean fish	**PC ae C38:4**	**-0.0034**	**1.90E-02**	**3.00E-02**
	Lean fish	**PC aa C42:5**	**0.0043**	**2.00E-02**	**3.10E-02**
	Total fish products	**PC aa C28:1**	**0.0075**	**2.10E-02**	**3.30E-02**
	Lean Total fish	**SM(OH) C14:1**	**0.0039**	**2.20E-02**	**3.40E-02**
	Total fish	**PC aa C38:3**	**-0.0081**	**2.30E-02**	**3.60E-02**
	Total fish products	**PC ae C32:1**	**0.0061**	**2.30E-02**	**3.60E-02**
	Alcohol	**PC ae C42:3**	**0.0041**	**2.40E-02**	**3.70E-02**
	Total fish products	**PC ae C44:6**	**0.0077**	**2.40E-02**	**3.70E-02**
	Lean fish	**PC ae C42:1**	**0.003**	**2.60E-02**	**3.90E-02**
	Total fish products	**PC aa C34:4**	**-0.0097**	**2.60E-02**	**4.00E-02**
	Total fish products	**PC ae C36:1**	**0.0071**	**2.70E-02**	**4.10E-02**
	Lean fish	**PC aa C38:5**	**0.0038**	**3.30E-02**	**4.90E-02**
	Total fish products	PC aa C32:0	0.0054	3.70E-02	5.50E-02
	Total fish	PC aa C32:3	0.0066	4.20E-02	6.20E-02
	Total fish	PC ae C40:3	0.0054	4.50E-02	6.70E-02
	Lean fish	SM(OH) C16:1	0.0033	4.60E-02	6.80E-02
	Total fish	PC aa C34:4	-0.0094	4.90E-02	7.10E-02
	Total fish	PC ae C42:4	-0.0076	5.10E-02	7.40E-02
	Total fish products	PC aa C32:3	0.0057	5.30E-02	7.70E-02
	Lean fish	PC ae C34:2	0.0029	5.60E-02	8.10E-02
	Alcohol	PC aa C40:3	0.0041	6.30E-02	8.90E-02
	Lean fish	PC ae C34:1	0.0027	6.80E-02	9.60E-02
	Alcohol	PC aa C42:1	-0.0038	7.00E-02	9.90E-02

	Total fish products	PC aa C42:4	-0.0052	7.40E-02	1.00E-01
	Total fish products	PC ae C38:3	-0.0051	7.90E-02	1.10E-01
	Lean fish	PC aa C36:1	0.0033	8.00E-02	1.10E-01
	Total fish	SM(OH) C14:1	0.0064	8.10E-02	1.10E-01
	Total fish products	PC aa C32:1	-0.011	8.50E-02	1.20E-01
	Total fish products	PC ae C38:5	-0.0046	8.80E-02	1.20E-01
	Lean fish	PC aa C34:1	0.0031	9.00E-02	1.20E-01
	Total fish	SM(OH) C16:1	0.0059	9.90E-02	1.30E-01
	Total fish	PC aa C42:4	-0.0051	1.10E-01	1.50E-01
	Total fish products	PC ae C40:3	0.0039	1.10E-01	1.50E-01
	Lean fish	PC ae C42:5	0.0021	1.20E-01	1.60E-01
	Lean fish	PC ae C44:6	0.0027	1.20E-01	1.60E-01
	Total fish	PC ae C38:5	-0.0045	1.20E-01	1.60E-01
	Alcohol	PC ae C44:6	-0.0033	1.30E-01	1.70E-01
	Alcohol	PC aa C36:2	0.0022	1.30E-01	1.70E-01
	Lean fish	PC aa C34:3	-0.0027	1.30E-01	1.70E-01
	Total fish	PC aa C32:1	-0.0102	1.40E-01	1.80E-01
	Total fish	PC aa C34:1	0.0054	1.40E-01	1.80E-01
	Total fish products	SM(OH) C16:1	0.0047	1.50E-01	1.90E-01
	Lean fish	PC aa C34:4	-0.0032	1.50E-01	1.90E-01
	Lean fish	PC ae C36:2	0.0022	1.50E-01	1.90E-01
	Total fish products	PC aa C34:1	0.0048	1.50E-01	1.90E-01
	Total fish	PC ae C42:1	0.0042	1.50E-01	1.90E-01
	Total fish products	SM(OH) C14:1	0.0048	1.50E-01	1.90E-01
	Lean fish	PC aa C32:0	0.0019	1.60E-01	2.00E-01
	Alcohol	PC ae C44:5	0.003	1.60E-01	2.00E-01
	Total fish	PC ae C38:3	-0.0042	1.90E-01	2.30E-01
	Lean fish	PC aa C40:5	-0.0025	1.90E-01	2.30E-01
	Lean fish	PC ae C36:3	0.0018	2.10E-01	2.60E-01
	Total fish products	PC aa C36:1	0.0044	2.20E-01	2.70E-01
	Total fish	PC aa C36:1	0.0047	2.30E-01	2.80E-01
	Total fish	PC ae C34:2	0.0038	2.30E-01	2.80E-01
	Lean fish	PC ae C36:4	-0.0018	2.40E-01	2.80E-01
	Total fish	PC ae C36:2	0.0038	2.50E-01	3.00E-01
	Lean fish	PC ae C44:5	0.002	2.70E-01	3.20E-01
	Total fish	PC ae C36:5	0.004	2.70E-01	3.20E-01
	Lean fish	PC aa C32:1	-0.0034	3.00E-01	3.50E-01
	Lean fish	PC ae C34:3	0.0018	3.10E-01	3.70E-01

	Total fish products	PC ae C36:5	0.0033	3.30E-01	3.80E-01
	Total fish products	PC ae C42:1	0.0026	3.30E-01	3.90E-01
	Total fish products	PC ae C36:3	-0.0026	3.50E-01	4.10E-01
	Alcohol	PC ae C34:0	0.0021	3.60E-01	4.10E-01
	Lean fish	PC aa C42:4	-0.0013	3.80E-01	4.40E-01
	Total fish products	PC ae C36:2	0.0026	3.90E-01	4.50E-01
	Lean fish	PC aa C38:3	-0.0014	4.00E-01	4.60E-01
	Lean fish	PC ae C36:5	0.0014	4.00E-01	4.60E-01
	Lean fish	PC aa C36:3	-0.0012	4.10E-01	4.60E-01
	Total fish products	PC ae C34:2	0.0024	4.20E-01	4.70E-01
	Total fish	PC ae C34:3	0.003	4.40E-01	5.00E-01
	Alcohol	PC aa C38:0	-0.0016	4.50E-01	5.10E-01
	Lean fish	PC ae C44:4	0.0012	4.70E-01	5.20E-01
	Total fish products	PC aa C34:2	-0.0015	4.80E-01	5.30E-01
	Lean fish	PC aa C34:2	-0.0008	4.80E-01	5.30E-01
	Total fish	PC aa C34:2	-0.0016	4.90E-01	5.40E-01
	Lean fish	PC ae C38:3	0.001	4.90E-01	5.40E-01
	Alcohol	PC ae C34:1	0.0012	5.10E-01	5.50E-01
	Total fish	PC aa C30:0	0.0029	5.20E-01	5.70E-01
	Alcohol	PC aa C40:2	-0.0014	5.50E-01	5.90E-01
	Total fish	PC ae C34:1	0.0017	5.90E-01	6.40E-01
	Total fish	PC ae C42:5	0.0015	6.00E-01	6.40E-01
	Total fish	PC ae C36:3	-0.0015	6.30E-01	6.70E-01
	Alcohol	PC ae C44:4	0.0009	6.30E-01	6.70E-01
	Total fish	PC aa C36:2	-0.0012	6.50E-01	6.90E-01
	Total fish products	PC aa C36:2	-0.001	6.80E-01	7.20E-01
	Total fish products	PC ae C44:4	-0.0012	6.90E-01	7.20E-01
	Total fish products	PC aa C30:0	0.0012	7.80E-01	8.10E-01
	Alcohol	PC aa C28:1	0.0006	7.80E-01	8.10E-01
	Total fish products	PC ae C34:3	0.0009	8.00E-01	8.40E-01
	Total fish	PC ae C44:5	0.0009	8.10E-01	8.40E-01
	Total fish products	PC ae C44:5	-0.0008	8.20E-01	8.50E-01
	Lean fish	PC ae C40:4	-0.0003	8.20E-01	8.50E-01
	Lean fish	PC ae C38:5	0.0002	8.70E-01	8.90E-01
	Lean fish	PC ae C42:4	0.0003	8.70E-01	8.90E-01
	Total fish products	PC ae C34:1	0.0004	9.00E-01	9.10E-01
	Total fish	PC ae C44:4	0.0004	9.00E-01	9.10E-01

	Lean fish	PC aa C36:2	-0.0001	9.30E-01	9.40E-01
	Total fish products	PC ae C42:5	0.0002	9.50E-01	9.60E-01
	Lean fish	PC aa C30:0	0.0001	9.70E-01	9.70E-01
2	**BMI**	**Glutamate**	**0.03**	**1.10E-52**	**5.30E-52**
	BMI	C18:1	0.0027	9.30E-02	2.30E-01
	BMI	Ornithine	0.0017	2.30E-01	3.80E-01
	BMI	C18:2	-0.0015	4.60E-01	5.70E-01
	BMI	Taurine	0.0004	7.30E-01	7.30E-01
3	**BMI**	**Lyso PC a C18:1**	**-0.024**	**2.60E-58**	**4.20E-57**
	BMI	**Lyso PC a C18:2**	**-0.027**	**3.20E-57**	**2.50E-56**
	BMI	**Lyso PC a C17:0**	**-0.019**	**8.20E-35**	**4.40E-34**
	BMI	**Lyso PC a C18:0**	**-0.01**	**1.40E-17**	**5.80E-17**
	BMI	**Lyso PC a C20:4**	**-0.013**	**6.00E-17**	**1.90E-16**
	BMI	**Lyso PC a C16:0**	**-0.0081**	**8.60E-14**	**2.30E-13**
	Fatty fish	**Lyso PC a C20:4**	**-0.017**	**4.40E-08**	**1.00E-07**
	Fatty fish	**Lyso PC a C20:3**	**-0.017**	**1.60E-07**	**3.10E-07**
	Fatty fish	**Lyso PC a C16:1**	**-0.012**	**1.40E-04**	**2.50E-04**
	Fatty fish	**Lyso PC a C18:0**	**-0.0054**	**2.90E-02**	**3.50E-02**
	Fatty fish	**Lyso PC a C18:1**	**-0.0067**	**2.60E-02**	**3.50E-02**
	BMI	**Lyso PC a C16:1**	**-0.0034**	**2.70E-02**	**3.50E-02**
	BMI	**Lyso PC a C20:3**	**-0.0037**	**2.40E-02**	**3.50E-02**
	Fatty fish	Lyso PC a C16:0	-0.0044	4.80E-02	5.30E-02
	Fatty fish	Lyso PC a C17:0	-0.0061	5.30E-02	5.30E-02
	Fatty fish	Lyso PC a C18:2	-0.0064	5.30E-02	5.30E-02

Abbreviations: BMI, body mass index; PC aa, diacyl-phosphatidylcholine; PC ae, acyl-alkyl phosphatidylcholine; SM, sphingomyelin; SM(OH), hydroxysphingomyelin; lyso PC a, lysophosphatidylcholine.

All models were adjusted for age at blood collection, time of day of blood collection, energy intake (kcal/day), fasting status, education level, physical activity, smoking status, baseline alcohol consumption, and body mass index (BMI). Models that examined alcohol intake and BMI as main exposures were not adjusted for alcohol intake and BMI, respectively. Variables are ordered by lowest Padj value by metabolite pattern.

To account for multiple testing, only dietary correlates of treelets that passed the false-discovery rate of 0.05 (Padj < 0.05), using the Benjamini-Hochberg method, were further analysed in the validation set. Associations with a Padj < 0.05 are bolded.

β is per one unit described in Table 3.1 (excluding BMI, in which β is per one unit kg/m^2).

Appendix 4.1 Full dataset: Risk of overall prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by ≤ 10 years of follow-up time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
PC aa C36:3	0.84	0.76	0.94	0.001	0.21	1809
PC aa C36:2	0.85	0.77	0.95	0.003	0.24	1809
PC ae C38:2	0.89	0.82	0.97	0.006	0.26	1809
PC aa C38:3	0.89	0.81	0.97	0.007	0.26	1809
PC aa C36:1	0.91	0.84	0.98	0.013	0.31	1809
PC ae C36:3	0.89	0.80	0.98	0.018	0.31	1809
PC aa C40:2	0.90	0.82	0.98	0.018	0.31	1809
PC aa C40:3	0.91	0.83	0.99	0.024	0.31	1809
Threonine	1.08	1.01	1.16	0.027	0.31	1809
Aspartic acid	0.81	0.67	0.98	0.028	0.31	244
Serotonin	1.25	1.02	1.53	0.028	0.31	244
PC ae C44:3	0.77	0.61	0.97	0.028	0.31	244
PC ae C40:3	0.89	0.80	0.99	0.030	0.31	1809
PC ae C38:3	0.91	0.84	0.99	0.031	0.31	1809
PC aa C42:4	0.89	0.81	0.99	0.032	0.31	1807
PC aa C34:3	0.91	0.83	0.99	0.038	0.35	1809
C14:1	0.93	0.87	1.00	0.043	0.37	1809
PC aa C40:4	0.93	0.86	1.00	0.051	0.41	1809
Lyso PC a C20:3	0.92	0.85	1.00	0.053	0.41	1809
PC aa C28:1	1.09	0.99	1.19	0.068	0.48	1809
Lyso PC a C16:1	0.93	0.86	1.01	0.070	0.48	1809
C2	0.94	0.87	1.01	0.071	0.48	1809
PC ae C42:2	0.93	0.86	1.01	0.075	0.48	1809
PC ae C42:1	0.93	0.85	1.01	0.082	0.48	1809
Glutamine	0.91	0.82	1.01	0.085	0.48	1809
PC aa C38:4	0.93	0.85	1.01	0.086	0.48	1809
cit_orn	0.93	0.85	1.01	0.088	0.48	1565
PC aa C42:2	0.93	0.85	1.01	0.094	0.49	1809
C18:1	0.93	0.85	1.01	0.096	0.49	1809
PC ae C42:3	0.93	0.86	1.01	0.099	0.49	1809
Lyso PC a C18:1	0.93	0.86	1.01	0.10	0.49	1809
PC ae C34:3	0.93	0.86	1.01	0.11	0.49	1809
PC aa C42:5	0.93	0.85	1.02	0.12	0.53	1809
PC ae C30:0	1.07	0.98	1.17	0.12	0.53	1808
SM(OH) C14:1	1.07	0.98	1.16	0.13	0.55	1809
PC ae C42:4	0.94	0.87	1.02	0.14	0.57	1809
C18	0.95	0.89	1.02	0.15	0.61	1809
PC ae C44:4	0.95	0.88	1.02	0.16	0.61	1809
SM C20:2	0.94	0.86	1.03	0.16	0.61	1809
PC aa C32:1	0.94	0.87	1.02	0.17	0.61	1809

Appendix 4.1 Full dataset: Risk of overall prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by ≤ 10 years of follow-up time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
Hexose	0.94	0.86	1.03	0.17	0.61	1809
PC ae C36:2	0.94	0.85	1.03	0.17	0.61	1809
SM C24:1	1.09	0.96	1.25	0.18	0.63	1809
PC ae C36:1	0.95	0.88	1.03	0.19	0.64	1808
PC aa C34:1	0.94	0.87	1.03	0.20	0.64	1809
Glutamate	1.09	0.96	1.23	0.20	0.64	1809
isoleucine	1.05	0.98	1.13	0.20	0.64	1809
PC ae C40:4	0.95	0.88	1.03	0.21	0.64	1807
Alanine	1.05	0.97	1.13	0.23	0.70	1809
SDMA	1.14	0.92	1.41	0.24	0.71	244
ketoaa	1.05	0.97	1.13	0.24	0.71	1565
PC ae C34:2	0.95	0.86	1.04	0.26	0.72	1809
PC aa C32:3	0.95	0.87	1.04	0.26	0.72	1809
PC aa C34:2	0.93	0.82	1.06	0.27	0.72	1809
PC ae C40:5	0.96	0.88	1.04	0.28	0.72	1809
fisher_ratio	1.04	0.97	1.12	0.28	0.72	1565
PC aa C40:5	0.96	0.89	1.04	0.28	0.72	1809
tyr_phe	0.96	0.89	1.03	0.28	0.72	1565
eaa	1.04	0.97	1.12	0.29	0.73	1565
PC ae C38:4	0.95	0.87	1.04	0.31	0.76	1809
C5	1.11	0.91	1.36	0.31	0.76	244
PC ae C32:2	0.96	0.88	1.05	0.33	0.80	1809
Ornithine	1.05	0.95	1.15	0.35	0.83	1809
Lyso PC a C18:0	0.96	0.88	1.05	0.36	0.83	1809
SM(OH) C22:2	1.05	0.95	1.15	0.36	0.83	1809
C16	0.97	0.91	1.04	0.37	0.83	1809
PC ae C36:4	0.96	0.87	1.05	0.38	0.83	1809
PC aa C34:4	0.96	0.89	1.05	0.39	0.83	1809
Leucine	1.03	0.96	1.11	0.40	0.83	1809
bcaa	1.03	0.96	1.11	0.41	0.83	1565
SM C24:0	0.96	0.86	1.06	0.42	0.83	1809
PC ae C36:0	0.97	0.90	1.04	0.42	0.83	1807
PC aa C36:5	1.03	0.96	1.11	0.43	0.83	1809
PC ae C38:5	0.96	0.87	1.06	0.43	0.83	1809
C18:2	0.97	0.89	1.05	0.43	0.83	1809
C4	0.91	0.73	1.15	0.44	0.83	244
PC aa C36:4	0.96	0.87	1.06	0.44	0.83	1809
PC ae C34:1	0.96	0.88	1.06	0.44	0.83	1809
Citrulline	0.97	0.91	1.04	0.44	0.83	1809
PC aa C38:6	1.03	0.96	1.11	0.45	0.83	1809
C3	1.03	0.96	1.10	0.45	0.83	1809
PC aa C30:0	1.03	0.95	1.12	0.46	0.83	1809

Appendix 4.1 Full dataset: Risk of overall prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by ≤ 10 years of follow-up time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
Asparagine	1.03	0.95	1.13	0.47	0.84	1809
glucoaa	1.03	0.95	1.11	0.48	0.84	1565
SM(OH) C16:1	1.03	0.95	1.12	0.49	0.84	1809
cit_arg	0.96	0.87	1.07	0.49	0.84	1565
PC ae C30:2	1.03	0.94	1.14	0.50	0.84	1800
PC ae C40:1	0.97	0.90	1.05	0.50	0.84	1808
C0	0.98	0.91	1.05	0.51	0.84	1809
Proline	1.02	0.96	1.10	0.51	0.84	1809
PC aa C42:1	1.03	0.95	1.11	0.51	0.84	1809
PC aa C42:0	1.03	0.95	1.11	0.53	0.85	1809
Kynurenine	0.98	0.91	1.05	0.53	0.85	1809
Phenylalanine	1.02	0.95	1.10	0.54	0.85	1809
PC ae C40:6	1.02	0.95	1.10	0.55	0.85	1809
Lyso PC a C18:2	0.98	0.90	1.06	0.55	0.85	1809
PC aa C42:6	1.06	0.86	1.30	0.57	0.87	244
orn_arg	1.03	0.92	1.16	0.59	0.89	1565
PC ae C44:6	1.02	0.95	1.10	0.60	0.90	1809
Lyso PC a C16:0	1.02	0.94	1.11	0.60	0.90	1809
PC ae C42:5	0.98	0.91	1.06	0.61	0.90	1809
PC ae C36:5	0.98	0.89	1.07	0.63	0.91	1809
Lyso PC a C28:1	1.06	0.83	1.34	0.64	0.92	238
ureaaa	1.02	0.94	1.10	0.65	0.93	1565
Glycine	0.98	0.91	1.06	0.66	0.93	1809
SM(OH) C22:1	0.98	0.90	1.07	0.70	0.93	1809
Lyso PC a C17:0	1.02	0.94	1.10	0.70	0.93	1809
C14	0.96	0.79	1.17	0.70	0.93	244
Arginine	0.98	0.89	1.08	0.71	0.93	1809
C12	0.96	0.79	1.17	0.72	0.93	244
PC aa C32:0	1.02	0.92	1.12	0.72	0.93	1809
PC aa C38:5	0.99	0.91	1.07	0.72	0.93	1809
nsaa	1.02	0.93	1.12	0.73	0.93	1565
PC aa C36:6	1.01	0.94	1.09	0.73	0.93	1809
PC ae C40:2	0.99	0.91	1.07	0.73	0.93	1809
C12:1	0.97	0.79	1.18	0.73	0.93	244
PC aa C38:0	0.99	0.91	1.07	0.74	0.93	1809
PC aa C32:2	0.97	0.80	1.17	0.75	0.93	244
SM C26:0	1.04	0.83	1.29	0.75	0.93	243
Lysine	1.01	0.94	1.09	0.76	0.93	1809
neaa	1.01	0.93	1.10	0.77	0.93	1565
PC ae C34:0	1.01	0.94	1.09	0.77	0.93	1809
Serine	1.01	0.94	1.09	0.77	0.93	1809
SM C16:0	1.01	0.93	1.11	0.78	0.93	1809

Appendix 4.1 Full dataset: Risk of overall prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by ≤ 10 years of follow-up time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
Methionine	1.01	0.94	1.08	0.79	0.93	1809
SM C18:1	0.99	0.91	1.08	0.79	0.93	1809
PC ae C38:0	0.98	0.80	1.19	0.80	0.93	244
SM C18:0	1.01	0.92	1.11	0.81	0.93	1809
Taurine	0.99	0.89	1.09	0.81	0.93	1809
Valine	1.01	0.94	1.08	0.83	0.95	1809
SM C26:1	0.98	0.80	1.20	0.84	0.95	244
Sarcosine	1.01	0.94	1.08	0.85	0.95	1809
C10	0.98	0.78	1.23	0.86	0.95	244
Creatinine	1.01	0.94	1.08	0.86	0.95	1809
Histidine	0.99	0.92	1.07	0.87	0.96	1809
Tryptophan	1.01	0.94	1.08	0.88	0.96	1809
SM C16:1	1.01	0.92	1.10	0.89	0.96	1809
PC ae C38:6	1.00	0.92	1.08	0.92	0.98	1809
PC ae C44:5	1.00	0.93	1.07	0.92	0.98	1809
aromaa	1.00	0.93	1.08	0.94	0.99	1565
t4-hydroxyproline	1.00	0.94	1.07	0.95	0.99	1809
Tyrosine	1.00	0.93	1.07	0.95	0.99	1809
PC ae C32:1	1.00	0.91	1.09	0.96	0.99	1809
ADMA	1.00	0.93	1.07	0.96	0.99	1809
PC aa C36:0	1.00	0.91	1.10	0.97	0.99	1809
Lyso PC a C24:0	1.00	0.78	1.27	0.97	0.99	239
SM(OH) C24:1	1.00	0.91	1.11	0.97	0.99	1809
PC aa C40:6	1.00	0.93	1.08	0.99	0.99	1809
Lyso PC a C20:4	1.00	0.92	1.08	0.99	0.99	1809

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 Padj < 0.05 are statistically significant after allowing for multiple testing using a false discovery rate controlling procedure at α = 0.05 (Benjamini-Hochberg). P-values are sorted in ascending order.

Appendix 4.2. Full dataset: Risk of overall prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by > 10 years of follow-up time time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
PC aa C32:1	0.91	0.85	0.97	0.0063	0.7595	2477
Creatinine	1.08	1.01	1.15	0.0257	0.7595	2477
Histidine	1.08	1.01	1.15	0.0263	0.7595	2477
PC aa C32:0	0.91	0.84	0.99	0.0345	0.7595	2477
SM(OH) C22:2	1.09	1.00	1.19	0.0400	0.7595	2477
Lyso PC a C16:1	0.93	0.87	1.00	0.0461	0.7595	2477
Hexose	0.92	0.84	1.00	0.0486	0.7595	2477
SM(OH) C22:1	1.08	1.00	1.17	0.0593	0.7595	2477
PC aa C40:4	0.94	0.88	1.00	0.0652	0.7595	2477
PC aa C34:4	0.94	0.87	1.01	0.0686	0.7595	2477
cit_orn	1.10	0.99	1.21	0.0750	0.7595	1391
C4	1.09	0.99	1.21	0.0798	0.7595	1086
PC aa C42:2	1.07	0.99	1.15	0.0824	0.7595	2476
SM C26:0	1.08	0.99	1.17	0.0835	0.7595	1086
PC aa C36:4	0.93	0.85	1.01	0.0925	0.7595	2477
PC aa C40:5	0.95	0.89	1.01	0.0942	0.7595	2477
SM(OH) C14:1	1.06	0.99	1.15	0.1027	0.7595	2477
Glutamine	1.07	0.99	1.15	0.1042	0.7595	2477
Tryptophan	1.05	0.99	1.12	0.1043	0.7595	2477
Phenylalanine	1.05	0.99	1.13	0.1089	0.7595	2477
PC aa C34:1	0.94	0.87	1.01	0.1138	0.7595	2477
SM(OH) C24:1	1.07	0.98	1.17	0.1220	0.7595	2477
PC aa C38:4	0.94	0.87	1.02	0.1224	0.7595	2477
C5	1.09	0.98	1.21	0.1236	0.7595	1086
orn_arg	0.91	0.80	1.03	0.1327	0.7595	1391
PC aa C30:0	0.95	0.89	1.02	0.1376	0.7595	2477
SM(OH) C16:1	1.06	0.98	1.14	0.1376	0.7595	2477
Lyso PC a C18:2	1.05	0.98	1.12	0.1449	0.7713	2477
PC ae C44:3	1.07	0.98	1.16	0.1534	0.7750	1084
PC aa C36:1	0.96	0.90	1.02	0.1560	0.7750	2477
PC aa C42:5	0.95	0.88	1.02	0.1727	0.7772	2477
Asparagine	1.05	0.98	1.12	0.1736	0.7772	2477
Serine	1.05	0.98	1.11	0.1771	0.7772	2477
PC aa C38:3	0.95	0.88	1.02	0.1817	0.7772	2477
Proline	1.04	0.98	1.11	0.1923	0.7772	2477
Serotonin	1.06	0.97	1.17	0.1997	0.7772	1086
nsaa	0.93	0.84	1.04	0.2026	0.7772	1391
PC ae C38:0	1.06	0.97	1.16	0.2043	0.7772	1084
Lyso PC a C28:1	1.07	0.96	1.19	0.2072	0.7772	1073
Lyso PC a C17:0	1.05	0.98	1.12	0.2087	0.7772	2477
PC ae C36:2	1.05	0.97	1.14	0.2366	0.8591	2477
PC ae C40:2	1.04	0.97	1.12	0.2422	0.8591	2477

Appendix 4.2. Full dataset: Risk of overall prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by > 10 years of follow-up time time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
PC aa C28:1	1.05	0.97	1.13	0.2597	0.9000	2477
PC aa C34:3	0.96	0.89	1.03	0.2769	0.9289	2477
PC ae C36:0	1.04	0.97	1.10	0.2806	0.9289	2474
Ornithine	0.96	0.88	1.04	0.2976	0.9641	2477
PC ae C36:4	0.96	0.88	1.04	0.3175	0.9832	2477
C14	1.05	0.95	1.16	0.3321	0.9832	1086
PC aa C42:6	1.04	0.96	1.14	0.3370	0.9832	1086
PC ae C30:2	1.04	0.96	1.14	0.3419	0.9832	2462
PC aa C40:6	0.97	0.91	1.03	0.3421	0.9832	2477
PC aa C36:3	0.96	0.88	1.05	0.3587	0.9832	2477
tyr_phe	0.96	0.88	1.05	0.3597	0.9832	1391
PC aa C36:0	1.03	0.96	1.11	0.3679	0.9832	2476
PC ae C34:1	0.96	0.89	1.05	0.3733	0.9832	2477
PC ae C38:5	0.96	0.88	1.05	0.3777	0.9832	2477
C14:1	0.97	0.90	1.04	0.3852	0.9832	2477
Arginine	1.04	0.95	1.15	0.3920	0.9832	2477
C12	1.04	0.95	1.14	0.4012	0.9832	1086
ADMA	0.97	0.91	1.04	0.4222	0.9832	2477
Sarcosine	0.98	0.92	1.04	0.4254	0.9832	2477
Tyrosine	0.98	0.92	1.04	0.4419	0.9832	2477
SM C24:0	1.04	0.95	1.13	0.4443	0.9832	2477
PC ae C40:3	1.02	0.96	1.09	0.4485	0.9832	2476
PC ae C36:1	1.02	0.96	1.09	0.4714	0.9832	2476
PC ae C34:2	1.03	0.95	1.12	0.4762	0.9832	2477
C12:1	1.03	0.94	1.13	0.4780	0.9832	1086
fisher_ratio	0.97	0.89	1.06	0.4889	0.9832	1391
PC ae C38:2	1.02	0.96	1.08	0.5153	0.9832	2476
PC aa C38:5	0.98	0.92	1.04	0.5267	0.9832	2477
PC ae C42:1	0.98	0.92	1.04	0.5337	0.9832	2476
PC ae C34:3	1.02	0.95	1.10	0.5756	0.9832	2477
SM C16:0	1.02	0.94	1.11	0.5757	0.9832	2477
PC ae C38:4	0.98	0.90	1.06	0.5787	0.9832	2477
aromaa	1.02	0.94	1.12	0.6077	0.9832	1391
Lysine	1.02	0.95	1.09	0.6082	0.9832	2477
PC ae C40:4	1.02	0.96	1.08	0.6124	0.9832	2477
Lyso PC a C18:0	1.02	0.95	1.10	0.6159	0.9832	2477
SDMA	1.02	0.93	1.12	0.6191	0.9832	1086
PC aa C42:1	1.02	0.95	1.09	0.6208	0.9832	2476
cit_arg	0.97	0.87	1.09	0.6495	0.9832	1391
PC aa C38:0	1.02	0.95	1.09	0.6502	0.9832	2477
PC ae C30:0	1.02	0.94	1.10	0.6517	0.9832	2477
PC ae C34:0	1.02	0.95	1.09	0.6548	0.9832	2477

Appendix 4.2. Full dataset: Risk of overall prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by > 10 years of follow-up time time						
Metabolite	**OR**	**LL**	**UL**	**P-value[2]**	**Padj[2]**	**n_{case}**
PC aa C32:3	1.02	0.94	1.10	0.6584	0.9832	2477
Glycine	0.99	0.92	1.05	0.6689	0.9832	2477
C18:2	0.98	0.90	1.07	0.6764	0.9832	2477
Valine	1.01	0.95	1.08	0.6789	0.9832	2477
PC aa C42:0	1.01	0.95	1.09	0.6860	0.9832	2477
C18:1	0.99	0.91	1.06	0.7053	0.9832	2477
Lyso PC a C24:0	0.98	0.87	1.10	0.7072	0.9832	1048
Lyso PC a C20:4	0.99	0.92	1.06	0.7149	0.9832	2477
ureaaa	0.98	0.90	1.08	0.7154	0.9832	1391
Lyso PC a C20:3	0.99	0.92	1.06	0.7269	0.9832	2477
PC ae C38:3	1.01	0.95	1.07	0.7376	0.9832	2477
Leucine	1.01	0.95	1.08	0.7415	0.9832	2477
Citrulline	1.01	0.95	1.08	0.7454	0.9832	2477
PC aa C36:6	0.99	0.93	1.05	0.7473	0.9832	2477
Methionine	1.01	0.95	1.08	0.7511	0.9832	2477
PC ae C40:1	0.99	0.93	1.06	0.7550	0.9832	2477
C16	0.99	0.92	1.06	0.7570	0.9832	2477
Taurine	1.01	0.93	1.10	0.7651	0.9832	2477
SM C24:1	1.02	0.91	1.14	0.7665	0.9832	2477
Kynurenine	0.99	0.93	1.06	0.7728	0.9832	2477
isoleucine	1.01	0.95	1.08	0.7741	0.9832	2477
PC ae C36:3	1.01	0.93	1.10	0.7791	0.9832	2477
PC ae C32:2	1.01	0.94	1.09	0.7795	0.9832	2477
PC ae C36:5	0.99	0.92	1.07	0.7896	0.9832	2477
Lyso PC a C16:0	1.01	0.94	1.08	0.7905	0.9832	2477
Glutamate	0.99	0.88	1.10	0.7957	0.9832	2477
PC ae C42:3	1.01	0.94	1.08	0.7957	0.9832	2477
PC ae C42:5	0.99	0.93	1.06	0.8067	0.9832	2477
PC ae C40:5	1.01	0.95	1.07	0.8079	0.9832	2476
SM C20:2	1.01	0.93	1.10	0.8190	0.9832	2477
PC aa C32:2	0.99	0.91	1.08	0.8262	0.9832	1086
PC ae C42:4	0.99	0.93	1.06	0.8266	0.9832	2477
t4-hydroxyproline	1.01	0.95	1.07	0.8296	0.9832	2477
Aspartic acid	1.01	0.92	1.11	0.8348	0.9832	1086
PC aa C36:5	1.01	0.95	1.07	0.8381	0.9832	2477
Threonine	0.99	0.93	1.06	0.8384	0.9832	2477
PC ae C38:6	1.01	0.94	1.08	0.8404	0.9832	2476
C3	0.99	0.91	1.08	0.8425	0.9832	2477
C2	0.99	0.94	1.06	0.8570	0.9832	2477
PC ae C40:6	1.01	0.94	1.07	0.8585	0.9832	2477
PC aa C42:4	0.99	0.93	1.06	0.8593	0.9832	2475
PC ae C44:6	1.01	0.94	1.07	0.8781	0.9832	2477

Appendix 4.2. Full dataset: Risk of overall prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by > 10 years of follow-up time time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
PC ae C32:1	0.99	0.92	1.07	0.8800	0.9832	2477
Lyso PC a C18:1	0.99	0.93	1.06	0.8826	0.9832	2477
SM C18:0	1.01	0.93	1.09	0.8876	0.9832	2477
PC aa C36:2	0.99	0.91	1.09	0.8950	0.9832	2477
SM C18:1	1.00	0.94	1.08	0.8958	0.9832	2477
PC ae C44:4	1.00	0.93	1.08	0.8987	0.9832	2477
PC aa C38:6	1.00	0.94	1.06	0.9106	0.9832	2477
ketoaa	1.00	0.92	1.10	0.9234	0.9832	1391
PC ae C44:5	1.00	0.93	1.06	0.9245	0.9832	2477
neaa	1.00	0.92	1.10	0.9250	0.9832	1391
C10	1.00	0.92	1.10	0.9256	0.9832	1086
PC aa C40:3	1.00	0.94	1.07	0.9343	0.9832	2477
PC aa C34:2	1.00	0.90	1.12	0.9347	0.9832	2477
eaa	1.00	0.92	1.10	0.9390	0.9832	1391
C0	1.00	0.94	1.07	0.9397	0.9832	2477
bcaa	1.00	0.91	1.09	0.9419	0.9832	1391
SM C16:1	1.00	0.93	1.08	0.9436	0.9832	2477
glucoaa	1.00	0.92	1.09	0.9611	0.9898	1391
SM C26:1	1.00	0.91	1.09	0.9681	0.9898	1086
Alanine	1.00	0.94	1.07	0.9717	0.9898	2477
C18	1.00	0.92	1.08	0.9794	0.9898	2477
PC ae C42:2	1.00	0.93	1.07	0.9831	0.9898	2477
PC aa C40:2	1.00	0.94	1.07	0.9924	0.9924	2476

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 Padj < 0.05 are statistically significant after allowing for multiple testing using a false discovery rate controlling procedure at $\alpha = 0.05$ (Benjamini-Hochberg). P-values are sorted in ascending order.

Appendix 4.3. Full dataset: Risk of advanced prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1]

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
PC ae C42:3	0.83	0.74	0.93	0.0013	0.0989	943
C18:2	0.79	0.69	0.91	0.0013	0.0989	943
PC aa C40:2	0.85	0.76	0.95	0.0030	0.1258	943
PC aa C40:3	0.85	0.76	0.95	0.0034	0.1258	943
PC aa C32:1	0.84	0.75	0.95	0.0053	0.1430	943
PC aa C42:4	0.85	0.75	0.95	0.0058	0.1430	941
Serotonin	1.27	1.07	1.52	0.0067	0.1430	363
Lyso PC a C16:1	0.86	0.77	0.96	0.0078	0.1459	943
PC aa C42:5	0.85	0.76	0.96	0.0104	0.1652	943
Lyso PC a C28:1	1.27	1.06	1.53	0.0113	0.1652	356
PC ae C42:2	0.87	0.77	0.97	0.0129	0.1652	943
SM(OH) C14:1	1.17	1.03	1.32	0.0133	0.1652	943
PC aa C34:3	0.86	0.76	0.97	0.0185	0.2119	943
PC aa C42:2	0.87	0.78	0.98	0.0241	0.2342	943
SM(OH) C16:1	1.15	1.02	1.30	0.0248	0.2342	943
PC aa C40:6	0.89	0.81	0.99	0.0251	0.2342	943
PC ae C40:1	0.88	0.79	0.99	0.0291	0.2550	943
PC ae C38:2	0.90	0.81	0.99	0.0388	0.2957	943
nsaa	0.84	0.72	0.99	0.0397	0.2957	580
SM C26:0	1.18	1.01	1.37	0.0410	0.2957	362
SM(OH) C22:2	1.16	1.01	1.34	0.0417	0.2957	943
PC ae C32:2	0.88	0.77	1.00	0.0440	0.2982	943
PC ae C42:1	0.90	0.82	1.00	0.0567	0.3227	943
PC aa C38:0	0.90	0.80	1.00	0.0588	0.3227	943
PC aa C36:1	0.90	0.81	1.00	0.0609	0.3227	943
PC aa C36:6	0.91	0.82	1.00	0.0615	0.3227	943
PC aa C34:1	0.88	0.77	1.01	0.0617	0.3227	943
PC aa C28:1	1.13	0.99	1.28	0.0689	0.3227	943
PC aa C32:0	0.88	0.76	1.01	0.0704	0.3227	943
C12:1	1.17	0.99	1.39	0.0711	0.3227	363
C0	0.91	0.83	1.01	0.0715	0.3227	943
orn_arg	0.82	0.65	1.02	0.0742	0.3227	580
PC aa C42:1	0.90	0.80	1.01	0.0745	0.3227	943
PC ae C36:0	0.91	0.82	1.01	0.0800	0.3227	941
PC aa C40:4	0.91	0.81	1.01	0.0813	0.3227	943
cit_arg	0.84	0.70	1.02	0.0818	0.3227	580
PC ae C44:4	0.90	0.80	1.01	0.0819	0.3227	943
PC ae C40:3	0.91	0.81	1.01	0.0844	0.3227	943
PC aa C32:3	0.89	0.78	1.02	0.0910	0.3227	943
C18:1	0.88	0.77	1.02	0.0934	0.3227	943
Lyso PC a C24:0	0.84	0.69	1.03	0.0936	0.3227	345
C12	1.16	0.97	1.38	0.0952	0.3227	363

Appendix 4.3. Full dataset: Risk of advanced prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1]

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
PC aa C42:0	0.91	0.81	1.02	0.0980	0.3227	943
PC aa C34:4	0.90	0.80	1.02	0.1000	0.3227	943
PC aa C40:5	0.91	0.82	1.02	0.1041	0.3227	943
PC ae C42:5	0.92	0.82	1.02	0.1046	0.3227	943
PC aa C38:6	0.92	0.83	1.02	0.1049	0.3227	943
C14:1	0.92	0.83	1.02	0.1054	0.3227	943
PC aa C36:0	0.90	0.80	1.02	0.1061	0.3227	942
Asparagine	1.10	0.98	1.23	0.1165	0.3436	943
PC aa C38:3	0.91	0.80	1.02	0.1176	0.3436	943
Glycine	0.92	0.82	1.02	0.1236	0.3502	943
PC aa C36:5	0.92	0.83	1.02	0.1246	0.3502	943
PC ae C40:5	0.93	0.84	1.02	0.1272	0.3502	942
PC aa C36:2	0.89	0.77	1.03	0.1307	0.3502	943
PC aa C36:3	0.90	0.78	1.03	0.1316	0.3502	943
PC aa C38:5	0.92	0.82	1.03	0.1353	0.3538	943
ureaaa	0.90	0.79	1.03	0.1386	0.3561	580
PC ae C30:0	1.10	0.96	1.25	0.1644	0.4151	942
C14	1.13	0.95	1.35	0.1719	0.4213	363
Serine	1.07	0.97	1.19	0.1725	0.4213	943
C10	1.13	0.95	1.34	0.1790	0.4301	363
PC ae C42:4	0.93	0.84	1.04	0.1922	0.4545	943
ADMA	0.93	0.84	1.04	0.2125	0.4923	943
Lyso PC a C17:0	1.07	0.96	1.20	0.2147	0.4923	943
Lyso PC a C18:1	0.93	0.83	1.04	0.2212	0.4993	943
PC ae C38:6	0.93	0.83	1.05	0.2328	0.5178	942
Glutamine	1.08	0.95	1.23	0.2412	0.5275	943
SM(OH) C22:1	1.08	0.95	1.22	0.2443	0.5275	943
Lyso PC a C16:0	0.93	0.83	1.05	0.2650	0.5537	943
Lyso PC a C18:0	0.93	0.83	1.05	0.2709	0.5537	943
C18	0.94	0.85	1.05	0.2711	0.5537	943
PC ae C38:4	1.08	0.94	1.23	0.2713	0.5537	943
Glutamate	0.90	0.75	1.09	0.2750	0.5537	943
Arginine	1.10	0.93	1.29	0.2832	0.5625	943
SM C24:0	0.92	0.80	1.07	0.2883	0.5652	943
SM C18:0	1.07	0.94	1.21	0.3262	0.6313	943
t4-hydroxyproline	1.05	0.95	1.16	0.3462	0.6613	943
PC ae C32:1	0.94	0.83	1.07	0.3515	0.6630	943
PC aa C36:4	0.93	0.81	1.08	0.3638	0.6645	943
PC ae C44:6	0.95	0.86	1.06	0.3639	0.6645	943
PC aa C34:2	0.92	0.76	1.11	0.3657	0.6645	943
PC ae C34:3	0.95	0.84	1.07	0.3738	0.6710	943
PC ae C40:4	0.96	0.87	1.06	0.3983	0.7066	942

Appendix 4.3. Full dataset: Risk of advanced prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1]

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
Citrulline	0.96	0.87	1.06	0.4104	0.7122	943
neaa	0.95	0.83	1.08	0.4229	0.7122	580
PC ae C38:3	0.96	0.87	1.06	0.4232	0.7122	943
Aspartic acid	0.94	0.81	1.09	0.4250	0.7122	363
PC ae C36:2	1.06	0.92	1.21	0.4254	0.7122	943
PC ae C34:2	1.06	0.92	1.22	0.4333	0.7173	943
Proline	1.04	0.94	1.15	0.4411	0.7222	943
aromaa	0.95	0.84	1.08	0.4480	0.7255	580
glucoaa	0.95	0.84	1.08	0.4565	0.7313	580
fisher_ratio	1.05	0.92	1.19	0.4666	0.7396	580
PC aa C38:4	0.95	0.84	1.09	0.4799	0.7527	943
Lyso PC a C20:3	0.96	0.86	1.07	0.4925	0.7529	943
C3	0.96	0.87	1.07	0.4930	0.7529	943
C2	0.96	0.87	1.07	0.4952	0.7529	943
PC aa C30:0	0.96	0.86	1.08	0.5045	0.7593	943
SM C24:1	0.94	0.77	1.14	0.5186	0.7697	943
PC ae C36:5	0.96	0.84	1.09	0.5217	0.7697	943
PC ae C40:6	0.97	0.87	1.07	0.5309	0.7755	943
C4	1.06	0.89	1.26	0.5416	0.7835	363
C16	0.97	0.88	1.07	0.5717	0.8190	943
Ornithine	0.96	0.84	1.11	0.5846	0.8296	943
Histidine	1.03	0.93	1.14	0.5962	0.8379	943
Tyrosine	0.97	0.88	1.07	0.6017	0.8379	943
PC ae C38:0	1.04	0.89	1.23	0.6133	0.8390	362
PC ae C36:3	0.97	0.84	1.11	0.6146	0.8390	943
Lyso PC a C18:2	1.03	0.92	1.15	0.6243	0.8390	943
PC aa C42:6	1.04	0.88	1.23	0.6250	0.8390	363
SM(OH) C24:1	1.03	0.90	1.19	0.6435	0.8561	943
eaa	0.97	0.86	1.10	0.6518	0.8595	580
SM C26:1	0.97	0.83	1.13	0.6679	0.8650	363
Taurine	1.03	0.89	1.19	0.6767	0.8650	943
ketoaa	0.97	0.86	1.10	0.6802	0.8650	580
SM C20:2	0.97	0.85	1.11	0.6831	0.8650	943
Threonine	1.02	0.92	1.13	0.6851	0.8650	943
tyr_phe	0.98	0.86	1.11	0.6980	0.8740	580
Valine	0.98	0.89	1.08	0.7050	0.8754	943
Phenylalanine	0.98	0.89	1.08	0.7120	0.8756	943
SM C16:1	0.98	0.86	1.11	0.7170	0.8756	943
Kynurenine	0.98	0.88	1.09	0.7237	0.8767	943
Hexose	1.02	0.90	1.17	0.7400	0.8892	943
PC ae C34:0	1.02	0.91	1.14	0.7497	0.8937	943
Creatinine	1.02	0.91	1.13	0.7755	0.9163	943

Appendix 4.3. Full dataset: Risk of advanced prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1]

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
cit_orn	1.02	0.88	1.19	0.7810	0.9163	580
PC ae C38:5	0.98	0.85	1.13	0.7888	0.9182	943
isoleucine	0.99	0.89	1.09	0.8061	0.9311	943
SM C18:1	1.01	0.90	1.14	0.8183	0.9364	943
Lysine	0.99	0.89	1.10	0.8233	0.9364	943
Leucine	0.99	0.90	1.09	0.8428	0.9421	943
PC ae C36:4	1.01	0.88	1.16	0.8499	0.9421	943
Methionine	1.01	0.91	1.12	0.8522	0.9421	943
PC aa C32:2	1.02	0.87	1.19	0.8536	0.9421	363
Sarcosine	1.01	0.91	1.11	0.9008	0.9859	943
bcaa	0.99	0.88	1.12	0.9119	0.9859	580
PC ae C44:5	1.00	0.89	1.11	0.9357	0.9859	943
PC ae C30:2	1.01	0.87	1.16	0.9435	0.9859	935
SDMA	0.99	0.85	1.17	0.9445	0.9859	363
PC ae C44:3	1.01	0.86	1.18	0.9471	0.9859	362
C5	1.01	0.85	1.18	0.9491	0.9859	363
PC ae C36:1	1.00	0.90	1.12	0.9515	0.9859	941
Tryptophan	1.00	0.91	1.11	0.9528	0.9859	943
PC ae C40:2	1.00	0.89	1.12	0.9826	0.9970	943
Lyso PC a C20:4	1.00	0.89	1.13	0.9827	0.9970	943
Alanine	1.00	0.89	1.12	0.9837	0.9970	943
SM C16:0	1.00	0.88	1.14	0.9915	0.9976	943
PC ae C34:1	1.00	0.87	1.15	0.9976	0.9976	943

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 Padj < 0.05 are statistically significant after allowing for multiple testing using a false discovery rate controlling procedure at α = 0.05 (Benjamini-Hochberg). P-values are sorted in ascending order.

Appendix 4.4 Full dataset: Risk of advanced prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by ≤ 10 years of follow-up time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
PC aa C42:4	**0.64**	**0.50**	**0.82**	**0.0003**	**0.0237**	**393**
PC ae C38:2	**0.70**	**0.57**	**0.85**	**0.0003**	**0.0237**	**394**
PC aa C40:2	**0.69**	**0.56**	**0.85**	**0.0005**	**0.0237**	**394**
PC aa C40:3	**0.72**	**0.59**	**0.88**	**0.0010**	**0.0377**	**394**
PC ae C40:3	**0.67**	**0.53**	**0.86**	**0.0013**	**0.0377**	**394**
PC ae C42:3	0.73	0.60	0.89	0.0019	0.0465	394
PC aa C42:2	0.75	0.61	0.91	0.0034	0.0730	394
PC aa C42:5	0.73	0.59	0.90	0.0040	0.0742	394
PC aa C36:2	0.73	0.58	0.93	0.0119	0.1900	394
PC ae C40:5	0.79	0.66	0.95	0.0142	0.1900	394
PC ae C42:2	0.80	0.66	0.96	0.0168	0.1900	394
PC aa C36:3	0.75	0.59	0.95	0.0173	0.1900	394
PC aa C34:3	0.78	0.64	0.96	0.0178	0.1900	394
PC ae C32:2	0.78	0.64	0.96	0.0193	0.1900	394
PC aa C32:3	0.78	0.63	0.96	0.0204	0.1900	394
PC ae C42:5	0.81	0.68	0.97	0.0216	0.1900	394
C18:2	0.79	0.65	0.97	0.0217	0.1900	394
PC aa C42:0	0.81	0.67	0.98	0.0282	0.2337	394
nsaa	0.78	0.63	0.98	0.0364	0.2844	328
PC aa C36:1	0.84	0.71	0.99	0.0389	0.2844	394
PC aa C34:2	0.73	0.54	0.99	0.0415	0.2844	394
PC ae C40:1	0.83	0.69	0.99	0.0433	0.2844	394
PC ae C38:3	0.83	0.69	1.00	0.0453	0.2844	394
PC aa C38:0	0.82	0.68	1.00	0.0463	0.2844	394
PC ae C42:4	0.82	0.68	1.00	0.0514	0.2844	394
C18	0.87	0.76	1.00	0.0518	0.2844	394
PC aa C36:0	0.81	0.66	1.00	0.0529	0.2844	394
C14:1	0.87	0.76	1.00	0.0534	0.2844	394
PC aa C38:3	0.83	0.68	1.01	0.0567	0.2900	394
PC ae C44:4	0.82	0.67	1.01	0.0585	0.2900	394
PC ae C36:0	0.85	0.72	1.01	0.0603	0.2900	393
PC ae C40:4	0.84	0.70	1.02	0.0783	0.3590	393
PC ae C42:1	0.85	0.70	1.02	0.0795	0.3590	394
C18:1	0.82	0.66	1.03	0.0827	0.3597	394
Lyso PC a C16:1	0.86	0.72	1.02	0.0864	0.3597	394
PC aa C40:4	0.85	0.71	1.03	0.0904	0.3597	394
PC aa C34:1	0.84	0.69	1.03	0.0910	0.3597	394
Glutamine	0.82	0.65	1.03	0.0946	0.3597	394
Aspartic acid	0.73	0.50	1.06	0.0958	0.3597	66
SM C24:0	0.82	0.64	1.04	0.0966	0.3597	394
PC aa C40:5	0.86	0.72	1.03	0.1038	0.3613	394
PC ae C30:2	0.83	0.66	1.04	0.1044	0.3613	392

Appendix 4.4 Full dataset: Risk of advanced prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by ≤ 10 years of follow-up time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
PC aa C40:6	0.87	0.73	1.03	0.1047	0.3613	394
PC ae C32:1	0.84	0.68	1.04	0.1067	0.3613	394
PC aa C42:1	0.86	0.71	1.04	0.1111	0.3618	394
PC ae C36:3	0.83	0.66	1.05	0.1131	0.3618	394
ureaaa	0.86	0.71	1.04	0.1164	0.3618	328
PC aa C38:5	0.86	0.72	1.04	0.1182	0.3618	394
PC aa C36:6	0.88	0.74	1.04	0.1226	0.3618	394
PC aa C32:1	0.86	0.72	1.04	0.1248	0.3618	394
Glycine	0.87	0.73	1.04	0.1249	0.3618	394
PC ae C38:6	0.86	0.71	1.04	0.1263	0.3618	394
Ornithine	0.84	0.66	1.06	0.1334	0.3750	394
Serotonin	1.44	0.89	2.33	0.1401	0.3867	66
PC aa C32:0	0.85	0.68	1.06	0.1447	0.3920	394
Lyso PC a C18:1	0.87	0.72	1.05	0.1514	0.3975	394
PC ae C44:3	0.71	0.45	1.13	0.1521	0.3975	66
Lyso PC a C28:1	1.42	0.86	2.35	0.1704	0.4255	63
Histidine	0.89	0.76	1.05	0.1709	0.4255	394
Lysine	0.89	0.75	1.05	0.1718	0.4255	394
PC ae C36:1	0.89	0.75	1.05	0.1742	0.4255	393
Tryptophan	0.90	0.76	1.05	0.1837	0.4415	394
neaa	0.88	0.73	1.07	0.2027	0.4795	328
Lyso PC a C18:0	0.88	0.72	1.08	0.2127	0.4881	394
PC ae C34:3	0.88	0.73	1.07	0.2129	0.4881	394
glucoaa	0.89	0.74	1.07	0.2195	0.4956	328
Valine	0.90	0.76	1.07	0.2288	0.5059	394
cit_arg	0.86	0.66	1.10	0.2310	0.5059	328
C0	0.91	0.78	1.06	0.2351	0.5059	394
C4	1.42	0.79	2.55	0.2377	0.5059	66
SM C24:1	0.83	0.60	1.14	0.2420	0.5079	394
PC aa C38:6	0.91	0.77	1.08	0.2678	0.5473	394
SM C26:0	1.30	0.82	2.05	0.2682	0.5473	65
SM C16:1	0.89	0.73	1.09	0.2725	0.5486	394
Lyso PC a C24:0	0.73	0.41	1.30	0.2854	0.5625	65
SM C16:0	0.89	0.72	1.10	0.2869	0.5625	394
orn_arg	0.86	0.64	1.15	0.3112	0.6021	328
PC aa C36:5	0.92	0.78	1.09	0.3198	0.6033	394
Phenylalanine	0.92	0.79	1.08	0.3217	0.6033	394
aromaa	0.92	0.78	1.09	0.3250	0.6033	328
PC ae C40:2	0.91	0.76	1.10	0.3280	0.6033	394
C14	0.80	0.51	1.27	0.3525	0.6321	66
PC aa C34:4	0.91	0.75	1.11	0.3548	0.6321	394
PC ae C34:1	0.90	0.73	1.12	0.3564	0.6321	394

Appendix 4.4 Full dataset: Risk of advanced prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by ≤ 10 years of follow-up time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
Serine	0.93	0.79	1.09	0.3612	0.6331	394
eaa	0.93	0.78	1.10	0.3856	0.6680	328
PC aa C30:0	0.92	0.77	1.11	0.4084	0.6954	394
Citrulline	0.94	0.81	1.09	0.4148	0.6954	394
PC ae C36:2	0.92	0.74	1.13	0.4154	0.6954	394
SM(OH) C24:1	0.91	0.73	1.14	0.4221	0.6968	394
Kynurenine	1.07	0.90	1.28	0.4256	0.6968	394
PC ae C40:6	0.93	0.79	1.11	0.4339	0.7028	394
isoleucine	0.94	0.81	1.10	0.4492	0.7198	394
ketoaa	0.94	0.79	1.12	0.4775	0.7405	328
t4-hydroxyproline	0.95	0.81	1.10	0.4821	0.7405	394
C16	0.95	0.84	1.09	0.4846	0.7405	394
SM(OH) C22:1	0.93	0.76	1.14	0.4936	0.7405	394
SM C20:2	0.93	0.75	1.15	0.4963	0.7405	394
PC ae C36:5	0.93	0.75	1.15	0.4963	0.7405	394
Lyso PC a C20:4	1.07	0.88	1.29	0.5042	0.7405	394
PC aa C38:4	0.93	0.76	1.15	0.5055	0.7405	394
C2	0.94	0.80	1.12	0.5069	0.7405	394
tyr_phe	0.95	0.80	1.12	0.5210	0.7479	328
Proline	0.95	0.81	1.11	0.5266	0.7479	394
bcaa	0.95	0.80	1.12	0.5270	0.7479	328
Methionine	0.95	0.81	1.12	0.5341	0.7507	394
PC ae C44:6	0.95	0.80	1.13	0.5461	0.7578	394
PC ae C34:2	0.93	0.74	1.17	0.5493	0.7578	394
PC ae C30:0	1.06	0.87	1.30	0.5571	0.7616	393
Tyrosine	0.96	0.82	1.11	0.5743	0.7678	394
SM(OH) C16:1	1.06	0.87	1.29	0.5776	0.7678	394
Lyso PC a C16:0	0.94	0.77	1.16	0.5787	0.7678	394
C12	0.87	0.54	1.41	0.5832	0.7678	66
Alanine	0.95	0.79	1.14	0.5874	0.7678	394
ADMA	0.95	0.80	1.13	0.5960	0.7722	394
PC aa C36:4	0.94	0.74	1.19	0.6036	0.7753	394
Lyso PC a C20:3	0.96	0.80	1.14	0.6174	0.7863	394
Sarcosine	1.04	0.89	1.21	0.6433	0.8107	394
Lyso PC a C17:0	1.04	0.87	1.24	0.6475	0.8107	394
PC ae C34:0	0.96	0.81	1.14	0.6643	0.8249	394
Creatinine	0.97	0.83	1.13	0.6779	0.8264	394
PC ae C44:5	0.96	0.81	1.15	0.6801	0.8264	394
Leucine	0.97	0.83	1.13	0.6843	0.8264	394
PC aa C42:6	1.10	0.70	1.73	0.6877	0.8264	66
Taurine	0.96	0.76	1.21	0.7201	0.8519	394
PC ae C38:4	1.04	0.84	1.29	0.7231	0.8519	394

Appendix 4.4 Full dataset: Risk of advanced prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by ≤ 10 years of follow-up time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
SM(OH) C14:1	1.04	0.85	1.26	0.7261	0.8519	394
SM C18:1	0.97	0.80	1.18	0.7368	0.8577	394
PC ae C38:0	1.08	0.68	1.71	0.7604	0.8783	66
Asparagine	0.97	0.78	1.20	0.7684	0.8807	394
SM(OH) C22:2	1.03	0.82	1.29	0.7848	0.8927	394
PC ae C36:4	1.03	0.82	1.29	0.7953	0.8949	394
fisher_ratio	1.02	0.87	1.20	0.7988	0.8949	328
SM C18:0	1.03	0.83	1.27	0.8148	0.9060	394
C12:1	0.95	0.57	1.57	0.8300	0.9112	66
Lyso PC a C18:2	0.98	0.82	1.18	0.8346	0.9112	394
PC ae C38:5	0.98	0.77	1.23	0.8378	0.9112	394
Threonine	1.01	0.87	1.19	0.8577	0.9261	394
C10	1.05	0.55	1.99	0.8902	0.9471	66
PC aa C32:2	1.03	0.67	1.60	0.8939	0.9471	66
SDMA	1.03	0.65	1.63	0.8962	0.9471	66
Hexose	1.01	0.83	1.23	0.9111	0.9561	394
cit_orn	0.99	0.81	1.22	0.9231	0.9590	328
Glutamate	1.01	0.75	1.36	0.9268	0.9590	394
PC aa C28:1	0.99	0.81	1.22	0.9419	0.9678	394
C3	1.00	0.87	1.15	0.9539	0.9735	394
C5	0.99	0.68	1.45	0.9696	0.9828	66
SM C26:1	1.01	0.66	1.54	0.9766	0.9832	66
Arginine	1.00	0.79	1.27	0.9853	0.9853	394

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 Padj < 0.05 are statistically significant after allowing for multiple testing using a false discovery rate controlling procedure at α = 0.05 (Benjamini-Hochberg). P-values are sorted in ascending order. Associations with Padj < 0.05 are bolded.

Appendix 4.5 Full dataset: Risk of advanced prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by > 10 years of follow-up time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
SM(OH) C14:1	1.28	1.09	1.51	0.0033	0.3284	549
Glutamine	1.27	1.07	1.51	0.0073	0.3284	549
SM(OH) C22:2	1.30	1.07	1.57	0.0091	0.3284	549
Serotonin	1.29	1.06	1.57	0.0110	0.3284	297
SM(OH) C16:1	1.23	1.05	1.44	0.0110	0.3284	549
Lyso PC a C28:1	1.30	1.05	1.61	0.0152	0.3447	293
PC aa C28:1	1.23	1.04	1.46	0.0176	0.3447	549
Serine	1.18	1.03	1.36	0.0185	0.3447	549
PC aa C32:1	0.83	0.71	0.98	0.0249	0.4129	549
SM(OH) C22:1	1.21	1.02	1.43	0.0290	0.4160	549
Asparagine	1.17	1.01	1.36	0.0307	0.4160	549
Histidine	1.16	1.01	1.33	0.0410	0.5091	549
PC ae C36:2	1.21	1.00	1.46	0.0450	0.5157	549
C18:2	0.81	0.65	1.00	0.0514	0.5475	549
t4-hydroxyproline	1.14	0.99	1.30	0.0641	0.5697	549
SM C26:0	1.17	0.99	1.38	0.0701	0.5697	297
Lyso PC a C16:1	0.87	0.75	1.01	0.0737	0.5697	549
C12	1.19	0.98	1.45	0.0742	0.5697	297
C12:1	1.19	0.98	1.43	0.0752	0.5697	297
Lyso PC a C17:0	1.15	0.99	1.35	0.0765	0.5697	549
PC ae C34:2	1.18	0.98	1.42	0.0825	0.5856	549
PC ae C42:3	0.88	0.76	1.02	0.0946	0.6197	549
C14	1.19	0.97	1.45	0.0960	0.6197	297
PC ae C30:2	1.17	0.97	1.42	0.0998	0.6197	543
PC ae C30:0	1.15	0.96	1.38	0.1176	0.6683	549
Arginine	1.21	0.95	1.55	0.1198	0.6683	549
Glutamate	0.82	0.64	1.05	0.1243	0.6683	549
SM(OH) C24:1	1.15	0.96	1.39	0.1320	0.6683	549
orn_arg	0.76	0.53	1.09	0.1377	0.6683	252
PC aa C40:6	0.91	0.80	1.03	0.1389	0.6683	549
Proline	1.11	0.97	1.27	0.1390	0.6683	549
cit_arg	0.80	0.60	1.08	0.1491	0.6940	252
PC ae C38:4	1.13	0.95	1.35	0.1650	0.7271	549
PC ae C36:1	1.11	0.96	1.28	0.1738	0.7271	548
PC aa C40:3	0.91	0.80	1.04	0.1783	0.7271	549
PC aa C40:2	0.91	0.80	1.04	0.1790	0.7271	549
C10	1.13	0.94	1.36	0.1806	0.7271	297
C0	0.92	0.81	1.05	0.1993	0.7816	549
Tryptophan	1.09	0.95	1.25	0.2146	0.7984	549
SM C18:0	1.11	0.94	1.31	0.2242	0.7984	549
PC ae C40:1	0.91	0.79	1.06	0.2244	0.7984	549
PC aa C36:6	0.92	0.80	1.05	0.2280	0.7984	549

Appendix 4.5 Full dataset: Risk of advanced prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by > 10 years of follow-up time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
PC aa C36:5	0.92	0.80	1.06	0.2304	0.7984	549
PC aa C34:4	0.91	0.77	1.07	0.2394	0.8061	549
PC aa C38:6	0.92	0.80	1.06	0.2484	0.8061	549
PC ae C42:2	0.92	0.79	1.06	0.2545	0.8061	549
C18	1.12	0.92	1.35	0.2577	0.8061	549
PC aa C42:4	0.93	0.81	1.06	0.2790	0.8061	548
Creatinine	1.09	0.94	1.26	0.2800	0.8061	549
Lyso PC a C24:0	0.88	0.70	1.11	0.2823	0.8061	280
Lyso PC a C18:2	1.08	0.94	1.25	0.2835	0.8061	549
PC ae C34:1	1.11	0.92	1.34	0.2853	0.8061	549
PC ae C34:0	1.09	0.93	1.28	0.2867	0.8061	549
PC ae C42:1	0.94	0.82	1.06	0.2950	0.8139	549
PC aa C42:5	0.93	0.80	1.07	0.3100	0.8258	549
PC aa C32:0	0.91	0.76	1.09	0.3104	0.8258	549
PC aa C42:1	0.93	0.80	1.08	0.3307	0.8339	549
PC aa C34:3	0.92	0.78	1.09	0.3315	0.8339	549
PC ae C40:2	1.08	0.92	1.26	0.3325	0.8339	549
SM C16:0	1.09	0.92	1.28	0.3358	0.8339	549
Kynurenine	0.93	0.81	1.08	0.3623	0.8849	549
ADMA	0.94	0.81	1.08	0.3793	0.9115	549
PC aa C34:1	0.92	0.77	1.11	0.3881	0.9179	549
PC aa C38:0	0.94	0.81	1.09	0.4003	0.9225	549
SM C18:1	1.07	0.91	1.25	0.4027	0.9225	549
fisher_ratio	1.09	0.89	1.34	0.4131	0.9225	252
Taurine	1.08	0.90	1.29	0.4155	0.9225	549
PC ae C36:3	1.08	0.90	1.29	0.4210	0.9225	549
Lysine	1.06	0.92	1.21	0.4293	0.9271	549
PC ae C36:0	0.95	0.82	1.09	0.4534	0.9652	548
PC aa C40:4	0.95	0.82	1.10	0.4648	0.9754	549
Methionine	1.05	0.92	1.20	0.4892	0.9868	549
Citrulline	0.96	0.84	1.09	0.4922	0.9868	549
PC aa C34:2	1.09	0.85	1.40	0.4967	0.9868	549
Threonine	1.05	0.92	1.20	0.5004	0.9868	549
nsaa	0.92	0.72	1.18	0.5047	0.9868	252
C3	0.95	0.80	1.12	0.5140	0.9868	549
PC aa C42:2	0.95	0.81	1.11	0.5236	0.9868	549
Glycine	0.96	0.83	1.10	0.5437	0.9868	549
PC ae C32:2	0.95	0.80	1.12	0.5447	0.9868	549
PC aa C36:0	0.95	0.82	1.11	0.5500	0.9868	548
PC ae C44:6	0.96	0.84	1.10	0.5527	0.9868	549
SM C16:1	1.05	0.89	1.24	0.5545	0.9868	549
PC ae C44:4	0.95	0.82	1.11	0.5563	0.9868	549

Appendix 4.5 Full dataset: Risk of advanced prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by > 10 years of follow-up time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
PC aa C36:1	0.96	0.83	1.11	0.5694	0.9895	549
PC aa C38:5	0.96	0.83	1.11	0.5831	0.9895	549
C18:1	0.95	0.78	1.15	0.5862	0.9895	549
PC aa C40:5	0.96	0.83	1.11	0.6106	0.9895	549
PC aa C36:4	0.95	0.78	1.16	0.6250	0.9895	549
PC aa C42:6	1.05	0.87	1.26	0.6322	0.9895	297
PC ae C44:3	1.04	0.88	1.24	0.6388	0.9895	296
SM C26:1	0.96	0.81	1.14	0.6445	0.9895	297
Ornithine	1.04	0.87	1.25	0.6455	0.9895	549
Lyso PC a C16:0	0.96	0.82	1.13	0.6520	0.9895	549
Sarcosine	0.97	0.85	1.11	0.6631	0.9895	549
bcaa	1.04	0.86	1.26	0.6633	0.9895	252
PC aa C42:0	0.97	0.84	1.12	0.6852	0.9895	549
PC aa C38:3	0.97	0.82	1.14	0.7054	0.9895	549
PC ae C44:5	1.03	0.89	1.19	0.7075	0.9895	549
ureaaa	0.96	0.78	1.19	0.7167	0.9895	252
PC ae C38:0	1.03	0.86	1.24	0.7169	0.9895	296
Phenylalanine	1.02	0.90	1.17	0.7215	0.9895	549
PC ae C38:3	1.02	0.91	1.15	0.7268	0.9895	549
C4	1.03	0.85	1.26	0.7313	0.9895	297
Tyrosine	0.98	0.86	1.12	0.7757	0.9895	549
PC ae C36:4	1.03	0.86	1.22	0.7827	0.9895	549
PC ae C32:1	1.02	0.87	1.21	0.7877	0.9895	549
PC aa C36:2	1.03	0.84	1.25	0.7911	0.9895	549
Alanine	1.02	0.88	1.18	0.7921	0.9895	549
eaa	1.03	0.85	1.24	0.7940	0.9895	252
Aspartic acid	0.98	0.82	1.16	0.8115	0.9895	297
SDMA	0.98	0.82	1.17	0.8171	0.9895	297
PC ae C40:5	0.99	0.88	1.11	0.8184	0.9895	548
PC ae C38:6	0.98	0.85	1.14	0.8233	0.9895	548
PC ae C42:4	0.99	0.86	1.12	0.8281	0.9895	549
C5	1.02	0.84	1.24	0.8291	0.9895	297
SM C24:1	1.03	0.80	1.33	0.8305	0.9895	549
Valine	1.01	0.89	1.16	0.8319	0.9895	549
C2	0.99	0.86	1.13	0.8347	0.9895	549
Lyso PC a C20:3	0.98	0.85	1.14	0.8379	0.9895	549
isoleucine	1.01	0.89	1.16	0.8400	0.9895	549
PC ae C34:3	1.02	0.87	1.19	0.8409	0.9895	549
glucoaa	1.02	0.84	1.24	0.8493	0.9895	252
SM C24:0	1.02	0.84	1.24	0.8644	0.9895	549
cit_orn	1.02	0.80	1.30	0.8665	0.9895	252
PC ae C40:4	1.01	0.90	1.14	0.8679	0.9895	549

Appendix 4.5 Full dataset: Risk of advanced prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by > 10 years of follow-up time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
C14:1	0.99	0.84	1.15	0.8720	0.9895	549
PC ae C42:5	0.99	0.86	1.13	0.8809	0.9895	549
neaa	1.02	0.83	1.25	0.8818	0.9895	252
PC aa C38:4	0.99	0.83	1.17	0.8837	0.9895	549
PC aa C32:2	1.01	0.85	1.21	0.8877	0.9895	297
PC ae C40:3	0.99	0.87	1.13	0.8941	0.9895	549
ketoaa	1.01	0.83	1.23	0.8959	0.9895	252
PC ae C38:5	1.01	0.84	1.21	0.9083	0.9895	549
aromaa	0.99	0.81	1.20	0.9121	0.9895	252
Lyso PC a C18:1	0.99	0.85	1.16	0.9140	0.9895	549
Lyso PC a C20:4	0.99	0.85	1.16	0.9142	0.9895	549
PC ae C36:5	0.99	0.84	1.18	0.9270	0.9895	549
Hexose	1.01	0.83	1.22	0.9285	0.9895	549
tyr_phe	0.99	0.81	1.22	0.9524	0.9895	252
PC aa C36:3	1.01	0.83	1.21	0.9543	0.9895	549
Lyso PC a C18:0	1.00	0.86	1.18	0.9572	0.9895	549
PC ae C38:2	1.00	0.88	1.13	0.9605	0.9895	549
PC aa C32:3	1.00	0.83	1.19	0.9615	0.9895	549
C16	1.00	0.85	1.19	0.9647	0.9895	549
SM C20:2	1.00	0.83	1.19	0.9746	0.9895	549
PC aa C30:0	1.00	0.86	1.16	0.9762	0.9895	549
Leucine	1.00	0.88	1.14	0.9852	0.9919	549
PC ae C40:6	1.00	0.87	1.15	0.9925	0.9925	549

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 Padj < 0.05 are statistically significant after allowing for multiple testing using a false discovery rate controlling procedure at α = 0.05 (Benjamini-Hochberg). P-values are sorted in ascending order.

Appendix 4.6 Full dataset: Risk of aggressive prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1]

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
PC ae C42:3	0.89	0.81	0.97	0.0076	0.3963	1481
Serotonin	1.23	1.05	1.43	0.0089	0.3963	468
PC aa C28:1	1.14	1.03	1.27	0.0091	0.3963	1481
SM(OH) C14:1	1.14	1.03	1.25	0.0106	0.3963	1481
PC aa C40:3	0.90	0.82	0.98	0.0134	0.4002	1481
PC aa C40:2	0.90	0.82	0.98	0.0172	0.4267	1481
Serine	1.10	1.01	1.19	0.0237	0.4894	1481
C12:1	1.18	1.02	1.37	0.0272	0.4894	468
PC aa C42:4	0.90	0.82	0.99	0.0296	0.4894	1479
Lyso PC a C28:1	1.19	1.01	1.40	0.0377	0.5398	460
C12	1.17	1.01	1.36	0.0398	0.5398	468
PC aa C42:2	0.91	0.83	1.00	0.0447	0.5545	1481
PC ae C42:2	0.92	0.84	1.00	0.0504	0.5771	1481
PC ae C44:4	0.92	0.84	1.01	0.0721	0.7245	1481
C14	1.14	0.98	1.33	0.0830	0.7245	468
Proline	1.07	0.99	1.16	0.0840	0.7245	1481
PC aa C42:5	0.92	0.84	1.01	0.0879	0.7245	1481
Taurine	1.10	0.99	1.24	0.0893	0.7245	1481
C10	1.14	0.98	1.33	0.0970	0.7245	468
PC ae C38:2	0.93	0.86	1.01	0.0973	0.7245	1481
Threonine	1.07	0.99	1.15	0.1079	0.7654	1481
PC ae C40:1	0.93	0.86	1.02	0.1130	0.7654	1481
SM C26:0	1.11	0.97	1.27	0.1253	0.7752	467
SM(OH) C16:1	1.08	0.98	1.18	0.1275	0.7752	1481
C18:2	0.92	0.83	1.02	0.1327	0.7752	1481
PC ae C30:0	1.08	0.97	1.19	0.1450	0.7752	1480
Lysine	1.06	0.98	1.15	0.1535	0.7752	1481
t4-hydroxyproline	1.06	0.98	1.14	0.1550	0.7752	1481
Lyso PC a C16:1	0.94	0.86	1.02	0.1556	0.7752	1481
PC aa C36:6	0.94	0.87	1.02	0.1561	0.7752	1481
SM(OH) C22:2	1.08	0.97	1.21	0.1701	0.8126	1481
PC aa C38:0	0.94	0.86	1.03	0.1908	0.8126	1481
PC aa C36:0	0.94	0.85	1.03	0.1919	0.8126	1480
Lyso PC a C24:0	0.89	0.74	1.06	0.1919	0.8126	449
PC aa C38:3	0.94	0.85	1.03	0.1945	0.8126	1481
Asparagine	1.06	0.97	1.17	0.2019	0.8126	1481
Glutamine	1.07	0.96	1.19	0.2254	0.8126	1481
PC aa C42:1	0.95	0.87	1.04	0.2293	0.8126	1481
C14:1	0.95	0.88	1.03	0.2415	0.8126	1481
PC aa C42:0	0.95	0.87	1.04	0.2469	0.8126	1481
PC ae C38:4	1.06	0.96	1.18	0.2484	0.8126	1481
cit_arg	0.93	0.81	1.06	0.2503	0.8126	1013

Appendix 4.6 Full dataset: Risk of aggressive prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1]

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
PC ae C40:3	0.95	0.87	1.04	0.2513	0.8126	1481
PC aa C40:6	0.96	0.88	1.03	0.2566	0.8126	1481
Tryptophan	1.05	0.97	1.13	0.2658	0.8126	1481
PC aa C32:1	0.95	0.86	1.04	0.2665	0.8126	1481
PC ae C30:2	1.07	0.95	1.19	0.2702	0.8126	1472
Arginine	1.07	0.95	1.21	0.2764	0.8126	1481
PC ae C34:0	1.05	0.96	1.15	0.2814	0.8126	1481
PC aa C34:3	0.95	0.86	1.05	0.2830	0.8126	1481
PC aa C34:4	0.95	0.86	1.05	0.2970	0.8126	1481
Histidine	1.05	0.96	1.14	0.2991	0.8126	1481
PC ae C42:5	0.96	0.88	1.04	0.2997	0.8126	1481
PC ae C42:1	0.96	0.88	1.04	0.3029	0.8126	1481
Lyso PC a C18:1	0.95	0.87	1.04	0.3095	0.8126	1481
PC ae C32:2	0.95	0.86	1.05	0.3134	0.8126	1481
C0	0.96	0.89	1.04	0.3142	0.8126	1481
PC aa C36:1	0.96	0.88	1.04	0.3200	0.8126	1481
orn_arg	0.93	0.80	1.08	0.3301	0.8126	1013
C16	1.04	0.96	1.12	0.3364	0.8126	1481
SM C24:0	0.94	0.84	1.06	0.3379	0.8126	1481
neaa	1.05	0.95	1.16	0.3390	0.8126	1013
Lyso PC a C17:0	1.04	0.95	1.14	0.3600	0.8126	1481
PC aa C36:3	0.95	0.85	1.06	0.3643	0.8126	1481
Lyso PC a C20:3	0.96	0.88	1.05	0.3647	0.8126	1481
Methionine	1.04	0.96	1.12	0.3759	0.8126	1481
glucoaa	1.04	0.95	1.15	0.3764	0.8126	1013
SM C18:0	1.05	0.95	1.16	0.3822	0.8126	1481
SM C26:1	0.94	0.82	1.08	0.3872	0.8126	468
PC ae C34:3	0.96	0.87	1.05	0.3903	0.8126	1481
ketoaa	1.04	0.95	1.14	0.3978	0.8126	1013
Glycine	0.96	0.88	1.05	0.3980	0.8126	1481
PC ae C36:4	1.05	0.94	1.17	0.3981	0.8126	1481
PC ae C40:5	0.97	0.90	1.05	0.4092	0.8239	1480
C5	1.06	0.92	1.23	0.4226	0.8326	468
Glutamate	1.06	0.92	1.22	0.4247	0.8326	1481
C4	1.06	0.91	1.25	0.4490	0.8575	468
PC ae C38:6	0.97	0.88	1.06	0.4511	0.8575	1480
Alanine	1.03	0.95	1.12	0.4568	0.8575	1481
Citrulline	0.97	0.90	1.05	0.4629	0.8575	1481
PC aa C34:2	1.06	0.91	1.22	0.4661	0.8575	1481
PC aa C36:5	0.97	0.90	1.05	0.4744	0.8620	1481
PC ae C42:4	0.97	0.89	1.06	0.4970	0.8923	1481
PC aa C40:4	0.97	0.89	1.06	0.5092	0.9011	1481

Appendix 4.6 Full dataset: Risk of aggressive prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1]

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
PC ae C38:3	0.97	0.90	1.06	0.5141	0.9011	1481
nsaa	0.96	0.86	1.09	0.5374	0.9159	1013
PC aa C36:2	0.96	0.86	1.09	0.5466	0.9159	1481
PC ae C36:2	1.03	0.93	1.15	0.5500	0.9159	1481
PC ae C36:0	0.98	0.90	1.06	0.5517	0.9159	1477
PC ae C34:2	1.03	0.92	1.15	0.5659	0.9159	1481
PC ae C36:3	0.97	0.87	1.08	0.5777	0.9159	1481
eaa	1.03	0.94	1.13	0.5813	0.9159	1013
SM C24:1	0.96	0.82	1.12	0.5894	0.9159	1481
SM C18:1	1.03	0.93	1.13	0.5901	0.9159	1481
SM(OH) C22:1	1.03	0.93	1.14	0.6003	0.9159	1481
PC aa C36:4	1.03	0.92	1.15	0.6030	0.9159	1481
PC aa C38:6	0.98	0.90	1.06	0.6108	0.9159	1481
PC aa C32:3	0.97	0.88	1.08	0.6116	0.9159	1481
ADMA	1.02	0.94	1.11	0.6160	0.9159	1481
Kynurenine	0.98	0.90	1.06	0.6187	0.9159	1481
PC ae C44:6	0.98	0.90	1.06	0.6239	0.9159	1481
Lyso PC a C16:0	1.02	0.93	1.12	0.6270	0.9159	1481
Sarcosine	1.02	0.94	1.10	0.6436	0.9290	1481
C18	1.02	0.94	1.11	0.6485	0.9290	1481
PC ae C38:0	1.03	0.90	1.19	0.6582	0.9295	467
Ornithine	1.02	0.92	1.14	0.6631	0.9295	1481
PC ae C44:3	0.97	0.84	1.12	0.6677	0.9295	467
Lyso PC a C20:4	1.02	0.93	1.12	0.6757	0.9295	1481
isoleucine	1.02	0.94	1.10	0.6919	0.9295	1481
SM C16:1	1.02	0.92	1.13	0.6938	0.9295	1481
tyr_phe	0.98	0.89	1.08	0.6951	0.9295	1013
PC aa C38:5	0.98	0.90	1.07	0.7059	0.9295	1481
PC ae C34:1	1.02	0.92	1.14	0.7076	0.9295	1481
PC ae C40:6	0.98	0.91	1.07	0.7112	0.9295	1481
Leucine	1.01	0.94	1.10	0.7222	0.9319	1481
PC ae C38:5	1.02	0.91	1.14	0.7307	0.9319	1481
C18:1	0.98	0.89	1.09	0.7373	0.9319	1481
Tyrosine	0.99	0.91	1.07	0.7380	0.9319	1481
SM C20:2	1.02	0.92	1.13	0.7555	0.9418	1481
PC aa C40:5	0.99	0.90	1.08	0.7585	0.9418	1481
PC ae C36:1	1.01	0.93	1.10	0.7711	0.9428	1479
SM(OH) C24:1	1.02	0.91	1.14	0.7720	0.9428	1481
Lyso PC a C18:0	0.99	0.90	1.08	0.8022	0.9718	1481
C3	0.99	0.91	1.08	0.8160	0.9746	1481
Phenylalanine	1.01	0.93	1.09	0.8302	0.9746	1481
PC aa C34:1	0.99	0.90	1.09	0.8486	0.9746	1481

Appendix 4.6 Full dataset: Risk of aggressive prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1]

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
C2	1.01	0.93	1.09	0.8563	0.9746	1481
PC aa C42:6	0.99	0.86	1.14	0.8682	0.9746	468
SM C16:0	0.99	0.90	1.10	0.8754	0.9746	1481
SDMA	0.99	0.86	1.14	0.8825	0.9746	468
PC ae C40:2	1.01	0.92	1.10	0.8900	0.9746	1481
bcaa	1.01	0.92	1.10	0.8973	0.9746	1013
PC aa C32:0	1.01	0.90	1.12	0.9041	0.9746	1481
Lyso PC a C18:2	1.01	0.92	1.10	0.9089	0.9746	1481
Hexose	1.01	0.91	1.11	0.9102	0.9746	1481
PC aa C38:4	1.01	0.91	1.11	0.9176	0.9746	1481
PC ae C40:4	1.00	0.92	1.08	0.9202	0.9746	1480
aromaa	1.00	0.92	1.10	0.9231	0.9746	1013
PC ae C36:5	1.00	0.90	1.10	0.9333	0.9746	1481
Aspartic acid	0.99	0.87	1.14	0.9391	0.9746	468
cit_orn	1.00	0.89	1.12	0.9396	0.9746	1013
fisher_ratio	1.00	0.91	1.10	0.9409	0.9746	1013
PC ae C44:5	1.00	0.92	1.09	0.9430	0.9746	1481
PC aa C32:2	1.00	0.87	1.16	0.9461	0.9746	468
PC aa C30:0	1.00	0.91	1.10	0.9531	0.9746	1481
Creatinine	1.00	0.93	1.09	0.9550	0.9746	1481
ureaaa	1.00	0.90	1.10	0.9667	0.9799	1013
PC ae C32:1	1.00	0.91	1.11	0.9820	0.9886	1481
Valine	1.00	0.92	1.08	0.9989	0.9989	1481

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 Padj < 0.05 are statistically significant after allowing for multiple testing using a false discovery rate controlling procedure at $\alpha = 0.05$ (Benjamini-Hochberg). P-values are sorted in ascending order. Associations with Padj < 0.05 are bolded.

Appendix 4.7 Full dataset: Risk of aggressive prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by ≤ 10 years of follow-up time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
PC ae C38:2	0.82	0.71	0.95	0.0069	0.6589	674
Serotonin	1.59	1.12	2.26	0.0088	0.6589	111
PC aa C40:2	0.84	0.72	0.97	0.0199	0.7444	674
PC aa C42:4	0.82	0.69	0.98	0.0259	0.7444	673
PC ae C42:3	0.86	0.75	0.98	0.0284	0.7444	674
PC aa C40:3	0.86	0.75	0.99	0.0364	0.7444	674
PC ae C40:3	0.84	0.71	0.99	0.0377	0.7444	674
PC aa C36:3	0.83	0.70	1.00	0.0440	0.7444	674
PC aa C42:2	0.86	0.74	1.00	0.0450	0.7444	674
Lyso PC a C18:1	0.87	0.75	1.00	0.0521	0.7763	674
PC ae C44:3	0.73	0.52	1.03	0.0728	0.8711	111
PC ae C36:4	1.17	0.99	1.38	0.0742	0.8711	674
PC aa C36:2	0.85	0.71	1.02	0.0760	0.8711	674
Lyso PC a C16:1	0.89	0.78	1.02	0.0963	0.9054	674
PC ae C44:4	0.89	0.77	1.03	0.1152	0.9054	674
PC aa C34:3	0.88	0.76	1.03	0.1154	0.9054	674
SM C26:0	1.34	0.93	1.95	0.1202	0.9054	110
cit_orn	0.89	0.76	1.04	0.1368	0.9054	563
SM(OH) C14:1	1.11	0.96	1.29	0.1482	0.9054	674
PC ae C30:0	1.12	0.96	1.30	0.1492	0.9054	673
PC ae C38:5	1.13	0.95	1.35	0.1519	0.9054	674
PC aa C28:1	1.12	0.96	1.30	0.1525	0.9054	674
PC ae C42:2	0.91	0.80	1.04	0.1758	0.9054	674
Glycine	0.92	0.80	1.04	0.1758	0.9054	674
Lyso PC a C18:0	0.91	0.79	1.05	0.1798	0.9054	674
PC aa C36:1	0.92	0.81	1.04	0.1818	0.9054	674
Aspartic acid	0.84	0.64	1.09	0.1822	0.9054	111
PC ae C38:3	0.91	0.80	1.04	0.1822	0.9054	674
Threonine	1.08	0.96	1.21	0.1842	0.9054	674
Lyso PC a C28:1	1.29	0.88	1.88	0.1885	0.9054	107
PC ae C40:1	0.92	0.80	1.05	0.1940	0.9054	674
Sarcosine	1.08	0.96	1.21	0.1945	0.9054	674
PC ae C38:4	1.10	0.94	1.29	0.2140	0.9292	674
PC aa C38:3	0.91	0.79	1.06	0.2201	0.9292	674
C14:1	0.94	0.85	1.04	0.2288	0.9292	674
PC ae C42:4	0.92	0.80	1.06	0.2494	0.9292	674
Citrulline	0.94	0.84	1.05	0.2502	0.9292	674
Lyso PC a C20:3	0.92	0.81	1.06	0.2505	0.9292	674
PC aa C42:5	0.92	0.79	1.07	0.2596	0.9292	674
Kynurenine	1.07	0.95	1.20	0.2796	0.9292	674
C0	0.94	0.83	1.06	0.2839	0.9292	674
ADMA	1.07	0.95	1.21	0.2871	0.9292	674

Appendix 4.7 Full dataset: Risk of aggressive prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by ≤ 10 years of follow-up time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
PC ae C36:5	1.09	0.93	1.27	0.2890	0.9292	674
Lyso PC a C18:2	0.93	0.81	1.06	0.2904	0.9292	674
PC ae C34:0	1.07	0.94	1.22	0.2951	0.9292	674
PC aa C36:4	1.09	0.92	1.29	0.2972	0.9292	674
SM C24:0	0.91	0.76	1.09	0.2987	0.9292	674
Taurine	1.09	0.92	1.30	0.2993	0.9292	674
Histidine	0.94	0.83	1.06	0.3099	0.9423	674
PC aa C32:2	1.19	0.84	1.68	0.3238	0.9581	111
C16	1.05	0.95	1.17	0.3414	0.9581	674
SDMA	1.18	0.83	1.68	0.3452	0.9581	111
PC aa C32:0	1.08	0.92	1.28	0.3590	0.9581	674
PC ae C36:3	0.92	0.78	1.10	0.3657	0.9581	674
Lyso PC a C24:0	0.84	0.58	1.22	0.3696	0.9581	109
C18:2	0.94	0.81	1.08	0.3708	0.9581	674
SM(OH) C16:1	1.07	0.92	1.24	0.3736	0.9581	674
Glutamine	0.92	0.78	1.10	0.3744	0.9581	674
PC aa C42:0	0.94	0.82	1.08	0.3815	0.9581	674
Lyso PC a C20:4	1.06	0.93	1.22	0.3864	0.9581	674
Glutamate	1.10	0.89	1.36	0.3923	0.9581	674
PC ae C42:1	0.94	0.82	1.08	0.4071	0.9704	674
C5	1.12	0.85	1.47	0.4136	0.9704	111
SM C18:0	1.07	0.91	1.25	0.4168	0.9704	674
PC ae C42:5	0.95	0.84	1.08	0.4500	0.9838	674
SM(OH) C22:2	1.07	0.90	1.26	0.4549	0.9838	674
PC aa C42:1	0.95	0.83	1.09	0.4574	0.9838	674
Lysine	1.05	0.93	1.18	0.4592	0.9838	674
Alanine	1.05	0.92	1.20	0.4761	0.9838	674
PC ae C32:2	0.95	0.82	1.10	0.4903	0.9838	674
PC ae C36:1	0.96	0.84	1.09	0.4909	0.9838	673
PC aa C32:3	0.95	0.81	1.11	0.4924	0.9838	674
ketoaa	1.04	0.92	1.18	0.5042	0.9838	563
orn_arg	1.07	0.87	1.31	0.5242	0.9838	563
PC aa C38:4	1.05	0.91	1.21	0.5287	0.9838	674
PC ae C40:5	0.96	0.84	1.09	0.5355	0.9838	674
C18	0.97	0.87	1.07	0.5394	0.9838	674
PC ae C30:2	1.05	0.89	1.25	0.5395	0.9838	671
PC aa C32:1	0.96	0.83	1.10	0.5473	0.9838	674
PC ae C34:3	0.96	0.83	1.10	0.5518	0.9838	674
Tryptophan	1.03	0.92	1.16	0.5668	0.9838	674
PC aa C40:4	0.96	0.84	1.10	0.5690	0.9838	674
PC aa C30:0	1.04	0.90	1.20	0.5722	0.9838	674
C3	1.03	0.93	1.14	0.5758	0.9838	674

Appendix 4.7 Full dataset: Risk of aggressive prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by ≤ 10 years of follow-up time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
tyr_phe	0.96	0.85	1.10	0.5802	0.9838	563
PC aa C38:0	0.96	0.84	1.10	0.5820	0.9838	674
Leucine	1.03	0.92	1.16	0.5826	0.9838	674
SM C18:1	1.04	0.90	1.20	0.5901	0.9838	674
Arginine	0.96	0.80	1.14	0.6037	0.9838	674
PC aa C34:2	0.94	0.76	1.18	0.6080	0.9838	674
PC aa C36:0	0.96	0.83	1.12	0.6084	0.9838	674
Asparagine	1.04	0.89	1.21	0.6101	0.9838	674
PC aa C36:5	1.03	0.91	1.17	0.6141	0.9838	674
Creatinine	1.03	0.92	1.15	0.6255	0.9874	674
PC ae C36:2	0.96	0.82	1.13	0.6513	0.9874	674
PC aa C36:6	0.97	0.86	1.10	0.6602	0.9874	674
PC ae C44:5	1.03	0.91	1.16	0.6626	0.9874	674
Valine	0.97	0.86	1.10	0.6631	0.9874	674
PC ae C40:6	1.03	0.91	1.16	0.6763	0.9874	674
cit_arg	0.96	0.80	1.15	0.6767	0.9874	563
C2	0.97	0.86	1.10	0.6862	0.9874	674
PC aa C38:6	1.03	0.91	1.16	0.6896	0.9874	674
C10	1.08	0.73	1.60	0.6925	0.9874	111
isoleucine	1.02	0.91	1.15	0.6937	0.9874	674
C14	0.94	0.70	1.28	0.6993	0.9874	111
C18:1	0.97	0.83	1.13	0.7024	0.9874	674
Ornithine	1.03	0.87	1.22	0.7107	0.9897	674
eaa	1.02	0.90	1.15	0.7309	0.9966	563
Methionine	1.02	0.91	1.15	0.7416	0.9966	674
C4	1.06	0.72	1.57	0.7602	0.9966	111
PC aa C34:1	0.98	0.84	1.13	0.7603	0.9966	674
PC aa C38:5	1.02	0.89	1.17	0.7662	0.9966	674
SM C16:1	1.02	0.88	1.19	0.7768	0.9966	674
PC ae C40:4	0.98	0.86	1.12	0.7795	0.9966	673
PC ae C40:2	1.02	0.89	1.16	0.7959	0.9966	674
SM(OH) C22:1	0.98	0.84	1.15	0.8122	0.9966	674
PC ae C34:2	1.02	0.86	1.21	0.8163	0.9966	674
ureaaa	0.98	0.86	1.13	0.8210	0.9966	563
Proline	1.01	0.90	1.14	0.8276	0.9966	674
PC ae C36:0	0.99	0.88	1.11	0.8339	0.9966	672
PC ae C34:1	1.02	0.87	1.19	0.8404	0.9966	674
Serine	1.01	0.90	1.14	0.8442	0.9966	674
PC ae C38:6	1.01	0.88	1.16	0.8491	0.9966	674
SM C16:0	0.99	0.84	1.15	0.8523	0.9966	674
PC ae C32:1	1.01	0.87	1.18	0.8524	0.9966	674
t4-hydroxyproline	1.01	0.90	1.14	0.8583	0.9966	674

Appendix 4.7 Full dataset: Risk of aggressive prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by ≤ 10 years of follow-up time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
glucoaa	1.01	0.89	1.16	0.8605	0.9966	563
SM C24:1	1.02	0.81	1.27	0.8926	0.9966	674
Hexose	1.01	0.88	1.16	0.8964	0.9966	674
PC ae C38:0	1.02	0.76	1.37	0.9011	0.9966	111
nsaa	0.99	0.84	1.16	0.9063	0.9966	563
C12	0.98	0.72	1.34	0.9081	0.9966	111
SM C20:2	0.99	0.85	1.16	0.9127	0.9966	674
C12:1	1.02	0.74	1.40	0.9181	0.9966	111
PC aa C40:6	1.00	0.88	1.12	0.9378	0.9966	674
neaa	1.00	0.87	1.15	0.9446	0.9966	563
PC aa C34:4	0.99	0.86	1.15	0.9452	0.9966	674
SM C26:1	0.99	0.73	1.34	0.9523	0.9966	111
Phenylalanine	1.00	0.89	1.12	0.9550	0.9966	674
Lyso PC a C16:0	1.00	0.86	1.15	0.9652	0.9966	674
PC aa C42:6	0.99	0.73	1.36	0.9674	0.9966	111
fisher_ratio	1.00	0.89	1.13	0.9740	0.9966	563
bcaa	1.00	0.89	1.13	0.9773	0.9966	563
PC aa C40:5	1.00	0.88	1.15	0.9829	0.9966	674
Lyso PC a C17:0	1.00	0.87	1.15	0.9856	0.9966	674
Tyrosine	1.00	0.89	1.12	0.9856	0.9966	674
PC ae C44:6	1.00	0.88	1.13	0.9916	0.9966	674
aromaa	1.00	0.89	1.13	0.9943	0.9966	563
SM(OH) C24:1	1.00	0.85	1.18	0.9966	0.9966	674

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 Padj < 0.05 are statistically significant after allowing for multiple testing using a false discovery rate controlling procedure at α = 0.05 (Benjamini-Hochberg). P-values are sorted in ascending order.

Appendix 4.8 Full dataset: Risk of aggressive prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by > 10 years of follow-up time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
Serine	1.20	1.07	1.35	0.0024	0.3351	800
Histidine	1.17	1.04	1.32	0.0070	0.3351	800
C12:1	1.26	1.06	1.51	0.0102	0.3351	357
C12	1.27	1.06	1.52	0.0107	0.3351	357
Glutamine	1.21	1.04	1.40	0.0119	0.3351	800
PC aa C28:1	1.19	1.03	1.37	0.0151	0.3351	800
SM(OH) C14:1	1.18	1.03	1.35	0.0161	0.3351	800
C14	1.25	1.04	1.51	0.0180	0.3351	357
Proline	1.13	1.01	1.27	0.0266	0.4403	800
Arginine	1.20	1.00	1.44	0.0443	0.6089	800
Lyso PC a C28:1	1.22	1.00	1.47	0.0449	0.6089	353
C18	1.15	1.00	1.33	0.0549	0.6814	800
orn_arg	0.80	0.63	1.02	0.0766	0.8222	443
C10	1.16	0.98	1.38	0.0934	0.8222	357
Lysine	1.10	0.98	1.24	0.0977	0.8222	800
Lyso PC a C17:0	1.11	0.98	1.26	0.1004	0.8222	800
t4-hydroxyproline	1.10	0.98	1.22	0.1028	0.8222	800
Serotonin	1.16	0.97	1.39	0.1037	0.8222	357
PC aa C34:2	1.18	0.96	1.44	0.1107	0.8222	800
SM(OH) C16:1	1.11	0.97	1.26	0.1221	0.8222	800
Asparagine	1.10	0.97	1.24	0.1234	0.8222	800
PC ae C36:2	1.12	0.97	1.30	0.1277	0.8222	800
PC ae C42:3	0.91	0.81	1.03	0.1327	0.8222	800
SM(OH) C22:2	1.13	0.96	1.32	0.1349	0.8222	800
PC aa C36:6	0.92	0.83	1.03	0.1380	0.8222	800
Kynurenine	0.92	0.81	1.03	0.1486	0.8407	800
PC aa C40:3	0.92	0.82	1.03	0.1564	0.8407	800
PC aa C36:5	0.92	0.83	1.03	0.1596	0.8407	800
Lyso PC a C18:0	1.10	0.96	1.24	0.1636	0.8407	800
Threonine	1.08	0.97	1.20	0.1796	0.8557	800
PC ae C30:2	1.11	0.95	1.30	0.1802	0.8557	794
neaa	1.11	0.95	1.29	0.1841	0.8557	443
Tryptophan	1.08	0.96	1.21	0.2131	0.8557	800
PC ae C36:1	1.08	0.96	1.21	0.2142	0.8557	799
Lyso PC a C18:2	1.08	0.96	1.21	0.2148	0.8557	800
Taurine	1.10	0.94	1.28	0.2185	0.8557	800
SM(OH) C22:1	1.09	0.95	1.26	0.2192	0.8557	800
PC aa C40:6	0.94	0.84	1.04	0.2254	0.8557	800
cit_orn	1.12	0.93	1.34	0.2298	0.8557	443
PC ae C42:2	0.93	0.83	1.05	0.2393	0.8557	800
PC aa C40:2	0.94	0.84	1.05	0.2409	0.8557	800
glucoaa	1.09	0.94	1.27	0.2412	0.8557	443

Appendix 4.8 Full dataset: Risk of aggressive prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by > 10 years of follow-up time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
PC aa C34:4	0.93	0.81	1.06	0.2515	0.8715	800
Lyso PC a C16:0	1.07	0.95	1.22	0.2683	0.8891	800
PC aa C42:5	0.93	0.82	1.06	0.2685	0.8891	800
cit_arg	0.89	0.73	1.10	0.2753	0.8917	443
SM C26:0	1.08	0.93	1.26	0.2843	0.8946	357
PC aa C36:2	1.09	0.93	1.28	0.2939	0.8946	800
Methionine	1.06	0.95	1.18	0.2992	0.8946	800
PC aa C32:1	0.94	0.82	1.06	0.3057	0.8946	800
PC aa C38:0	0.94	0.84	1.06	0.3062	0.8946	800
PC aa C36:0	0.94	0.82	1.07	0.3292	0.9432	799
SM C26:1	0.93	0.79	1.08	0.3411	0.9588	357
PC aa C38:6	0.95	0.85	1.06	0.3553	0.9741	800
PC aa C42:4	0.95	0.84	1.07	0.3730	0.9741	799
SM(OH) C24:1	1.07	0.92	1.25	0.3763	0.9741	800
Sarcosine	0.95	0.86	1.06	0.3926	0.9741	800
C3	0.94	0.81	1.09	0.3976	0.9741	800
C4	1.08	0.90	1.30	0.4019	0.9741	357
PC ae C34:2	1.07	0.92	1.24	0.4019	0.9741	800
PC ae C38:6	0.95	0.85	1.07	0.4115	0.9741	799
C2	1.04	0.94	1.16	0.4197	0.9741	800
PC ae C40:1	0.95	0.85	1.07	0.4239	0.9741	800
PC ae C38:4	1.06	0.92	1.23	0.4258	0.9741	800
PC ae C34:0	1.05	0.93	1.19	0.4401	0.9741	800
SM C18:0	1.05	0.92	1.21	0.4460	0.9741	800
SDMA	0.94	0.80	1.11	0.4493	0.9741	357
PC ae C30:0	1.05	0.92	1.21	0.4619	0.9741	800
PC aa C36:3	1.06	0.91	1.24	0.4703	0.9741	800
C18:2	0.94	0.81	1.11	0.4725	0.9741	800
C16	1.04	0.93	1.17	0.4820	0.9741	800
PC ae C44:4	0.96	0.84	1.09	0.4904	0.9741	800
PC ae C34:1	1.05	0.91	1.23	0.4971	0.9741	800
ketoaa	1.05	0.91	1.22	0.5036	0.9741	443
PC aa C42:1	0.96	0.85	1.08	0.5069	0.9741	800
SM C16:1	1.05	0.91	1.21	0.5213	0.9741	800
Lyso PC a C24:0	0.93	0.76	1.15	0.5225	0.9741	340
PC aa C42:2	0.96	0.85	1.09	0.5250	0.9741	800
eaa	1.05	0.90	1.22	0.5332	0.9741	443
C5	1.06	0.88	1.28	0.5434	0.9741	357
Aspartic acid	1.05	0.89	1.24	0.5450	0.9741	357
SM C18:1	1.04	0.91	1.19	0.5468	0.9741	800
Lyso PC a C18:1	1.04	0.92	1.17	0.5474	0.9741	800
PC ae C36:5	0.96	0.83	1.10	0.5492	0.9741	800

Appendix 4.8 Full dataset: Risk of aggressive prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by > 10 years of follow-up time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
PC aa C42:0	0.96	0.85	1.09	0.5621	0.9794	800
SM C24:1	0.94	0.76	1.16	0.5768	0.9794	800
nsaa	0.95	0.79	1.14	0.5782	0.9794	443
PC ae C42:5	0.97	0.87	1.08	0.5784	0.9794	800
Valine	1.03	0.92	1.15	0.5853	0.9799	800
Phenylalanine	1.03	0.92	1.15	0.6015	0.9879	800
SM C20:2	1.04	0.89	1.21	0.6181	0.9879	800
PC ae C32:2	0.97	0.85	1.11	0.6203	0.9879	800
PC ae C40:6	0.97	0.87	1.09	0.6225	0.9879	800
Ornithine	1.04	0.89	1.21	0.6232	0.9879	800
PC aa C38:5	0.97	0.87	1.09	0.6329	0.9883	800
PC ae C38:0	1.04	0.88	1.23	0.6421	0.9883	356
PC ae C36:0	0.98	0.87	1.09	0.6675	0.9883	798
PC ae C42:1	0.98	0.88	1.09	0.6676	0.9883	800
ureaaa	1.03	0.88	1.21	0.6704	0.9883	443
C0	0.98	0.88	1.09	0.7020	0.9883	800
Tyrosine	0.98	0.88	1.09	0.7066	0.9883	800
PC aa C38:3	0.97	0.85	1.11	0.7091	0.9883	800
Lyso PC a C16:1	0.98	0.87	1.10	0.7136	0.9883	800
PC ae C38:5	0.97	0.83	1.13	0.7142	0.9883	800
SM C16:0	1.03	0.89	1.18	0.7238	0.9883	800
PC ae C44:6	0.98	0.88	1.10	0.7310	0.9883	800
PC ae C36:3	1.03	0.89	1.19	0.7340	0.9883	800
PC aa C32:0	0.98	0.84	1.13	0.7424	0.9883	800
PC ae C38:3	1.02	0.92	1.13	0.7454	0.9883	800
C18:1	1.02	0.88	1.19	0.7458	0.9883	800
PC ae C40:5	0.98	0.89	1.09	0.7516	0.9883	799
Alanine	1.02	0.91	1.14	0.7547	0.9883	800
PC ae C40:4	1.02	0.92	1.13	0.7557	0.9883	800
bcaa	1.02	0.88	1.18	0.7562	0.9883	443
PC aa C30:0	0.98	0.87	1.11	0.7660	0.9891	800
fisher_ratio	1.02	0.88	1.19	0.7826	0.9891	443
PC ae C32:1	1.02	0.89	1.16	0.7908	0.9891	800
isoleucine	1.01	0.91	1.13	0.7944	0.9891	800
PC aa C32:3	1.02	0.88	1.17	0.8226	0.9891	800
Glutamate	1.02	0.84	1.25	0.8308	0.9891	800
Hexose	0.98	0.85	1.13	0.8319	0.9891	800
ADMA	0.99	0.87	1.12	0.8496	0.9891	800
Leucine	1.01	0.91	1.13	0.8530	0.9891	800
PC ae C36:4	0.99	0.85	1.15	0.8637	0.9891	800
PC ae C40:2	1.01	0.89	1.15	0.8697	0.9891	800
aromaa	1.01	0.87	1.18	0.8712	0.9891	443

Appendix 4.8 Full dataset: Risk of aggressive prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1] by > 10 years of follow-up time

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
PC ae C44:5	0.99	0.88	1.11	0.8784	0.9891	800
PC aa C34:1	1.01	0.88	1.16	0.8801	0.9891	800
PC ae C42:4	1.01	0.90	1.12	0.8860	0.9891	800
PC ae C44:3	1.01	0.86	1.18	0.8889	0.9891	356
PC ae C38:2	1.01	0.91	1.12	0.8987	0.9891	800
tyr_phe	1.01	0.86	1.18	0.9041	0.9891	443
PC aa C36:1	1.01	0.90	1.13	0.9222	0.9891	800
Lyso PC a C20:4	1.01	0.89	1.14	0.9241	0.9891	800
PC aa C32:2	0.99	0.84	1.17	0.9338	0.9891	357
C14:1	0.99	0.88	1.12	0.9352	0.9891	800
PC ae C40:3	1.00	0.90	1.12	0.9482	0.9891	800
Glycine	1.00	0.89	1.13	0.9553	0.9891	800
PC aa C38:4	1.00	0.87	1.14	0.9581	0.9891	800
PC aa C34:3	1.00	0.87	1.14	0.9678	0.9891	800
Creatinine	1.00	0.89	1.12	0.9709	0.9891	800
PC ae C34:3	1.00	0.88	1.14	0.9716	0.9891	800
SM C24:0	1.00	0.85	1.18	0.9728	0.9891	800
Citrulline	1.00	0.90	1.12	0.9739	0.9891	800
PC aa C42:6	1.00	0.85	1.17	0.9758	0.9891	357
PC aa C36:4	1.00	0.85	1.17	0.9791	0.9891	800
PC aa C40:4	1.00	0.89	1.12	0.9792	0.9891	800
PC aa C40:5	1.00	0.89	1.12	0.9877	0.9891	800
Lyso PC a C20:3	1.00	0.89	1.13	0.9891	0.9891	800

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 Padj < 0.05 are statistically significant after allowing for multiple testing using a false discovery rate controlling procedure at α = 0.05 (Benjamini-Hochberg). P-values are sorted in ascending order.

Appendix 4.9 Full dataset: Risk of high grade prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1]

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
Lyso PC a C24:0	0.69	0.47	1.02	0.0644	0.9740	132
PC aa C28:1	1.19	0.98	1.44	0.0794	0.9740	462
Aspartic acid	0.80	0.62	1.03	0.0853	0.9740	137
PC aa C42:0	0.87	0.74	1.02	0.0863	0.9740	462
PC aa C40:3	0.87	0.74	1.02	0.0868	0.9740	462
C10	1.26	0.96	1.63	0.0908	0.9740	137
PC ae C42:3	0.87	0.75	1.02	0.0909	0.9740	462
PC aa C38:0	0.88	0.75	1.03	0.1022	0.9740	462
PC aa C36:6	0.89	0.77	1.03	0.1144	0.9740	462
PC aa C42:4	0.86	0.72	1.04	0.1179	0.9740	461
Lyso PC a C17:0	1.14	0.97	1.35	0.1194	0.9740	462
SM(OH) C14:1	1.15	0.96	1.39	0.1268	0.9740	462
PC aa C42:5	0.88	0.74	1.04	0.1314	0.9740	462
PC aa C40:2	0.89	0.76	1.04	0.1337	0.9740	462
C14:1	0.89	0.76	1.04	0.1454	0.9740	462
SM C24:1	0.82	0.62	1.08	0.1499	0.9740	462
PC aa C40:6	0.90	0.78	1.04	0.1502	0.9740	462
C12:1	1.21	0.93	1.55	0.1515	0.9740	137
C12	1.20	0.93	1.54	0.1614	0.9740	137
PC ae C38:4	1.15	0.94	1.40	0.1618	0.9740	462
SM(OH) C16:1	1.14	0.95	1.36	0.1715	0.9740	462
SM C26:0	1.19	0.93	1.53	0.1717	0.9740	137
Kynurenine	0.90	0.77	1.05	0.1882	0.9740	462
Taurine	1.15	0.93	1.42	0.1912	0.9740	462
SDMA	1.20	0.91	1.59	0.1933	0.9740	137
C14	1.18	0.92	1.51	0.1934	0.9740	137
PC aa C38:6	0.91	0.79	1.05	0.1939	0.9740	462
PC aa C42:1	0.90	0.76	1.06	0.1972	0.9740	462
PC aa C32:3	0.88	0.73	1.07	0.2025	0.9740	462
PC aa C42:2	0.90	0.76	1.06	0.2097	0.9740	462
PC ae C34:0	1.11	0.94	1.31	0.2265	0.9740	462
Serine	1.09	0.94	1.27	0.2318	0.9740	462
Serotonin	1.17	0.89	1.53	0.2531	0.9740	137
PC ae C38:6	0.91	0.78	1.07	0.2543	0.9740	462
PC ae C34:3	0.91	0.77	1.07	0.2578	0.9740	462
PC ae C40:3	0.91	0.77	1.08	0.2649	0.9740	462
C16	1.07	0.95	1.22	0.2732	0.9740	462
PC aa C36:0	0.91	0.77	1.08	0.2764	0.9740	462
PC ae C36:3	0.90	0.74	1.09	0.2777	0.9740	462
PC ae C42:5	0.92	0.79	1.07	0.2813	0.9740	462
SM(OH) C24:1	1.12	0.91	1.37	0.2888	0.9740	462
Lyso PC a C20:4	1.09	0.92	1.29	0.3114	0.9740	462

Appendix 4.9 Full dataset: Risk of high grade prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1]

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
C4	1.18	0.85	1.63	0.3257	0.9740	137
PC ae C42:2	0.93	0.80	1.08	0.3280	0.9740	462
PC ae C36:2	1.10	0.91	1.34	0.3291	0.9740	462
Valine	0.93	0.79	1.08	0.3359	0.9740	462
PC ae C30:0	1.09	0.91	1.31	0.3428	0.9740	461
PC aa C34:3	0.92	0.77	1.10	0.3497	0.9740	462
PC aa C36:4	1.10	0.90	1.35	0.3504	0.9740	462
PC ae C40:5	0.93	0.81	1.08	0.3549	0.9740	461
PC aa C36:5	0.93	0.80	1.08	0.3578	0.9740	462
Hexose	0.92	0.78	1.09	0.3602	0.9740	462
PC ae C42:4	0.93	0.80	1.08	0.3607	0.9740	462
PC ae C38:0	1.13	0.87	1.47	0.3681	0.9740	136
PC aa C40:5	0.93	0.80	1.09	0.3839	0.9740	462
PC aa C34:2	1.13	0.86	1.48	0.3899	0.9740	462
aromaa	0.92	0.77	1.11	0.3918	0.9740	325
Tyrosine	0.94	0.81	1.09	0.3980	0.9740	462
Leucine	0.94	0.81	1.09	0.4085	0.9740	462
PC ae C38:2	0.94	0.81	1.09	0.4371	0.9740	462
SM C18:0	1.08	0.89	1.29	0.4444	0.9740	462
PC ae C40:6	0.95	0.82	1.09	0.4550	0.9740	462
PC ae C44:6	0.95	0.82	1.10	0.4611	0.9740	462
PC aa C38:5	0.94	0.81	1.10	0.4627	0.9740	462
Threonine	1.06	0.91	1.22	0.4635	0.9740	462
PC aa C34:4	0.94	0.79	1.11	0.4641	0.9740	462
SM(OH) C22:2	1.08	0.88	1.32	0.4684	0.9740	462
isoleucine	0.95	0.82	1.10	0.4694	0.9740	462
bcaa	0.94	0.79	1.12	0.4843	0.9740	325
PC aa C32:2	0.90	0.68	1.20	0.4902	0.9740	137
PC ae C40:1	0.95	0.82	1.10	0.5031	0.9740	462
PC aa C32:1	0.94	0.80	1.12	0.5060	0.9740	462
Lyso PC a C16:0	1.06	0.89	1.26	0.5073	0.9740	462
Phenylalanine	0.95	0.82	1.10	0.5074	0.9740	462
Lyso PC a C28:1	1.13	0.79	1.60	0.5089	0.9740	132
SM C26:1	1.09	0.84	1.42	0.5147	0.9740	137
SM C16:0	0.94	0.79	1.13	0.5169	0.9740	462
PC ae C36:0	0.95	0.82	1.11	0.5329	0.9740	460
C3	0.94	0.78	1.14	0.5412	0.9740	462
PC ae C36:1	1.05	0.90	1.22	0.5442	0.9740	462
SM(OH) C22:1	1.06	0.88	1.28	0.5585	0.9740	462
Ornithine	1.06	0.87	1.28	0.5640	0.9740	462
SM C24:0	0.94	0.76	1.17	0.5751	0.9740	462
Lyso PC a C18:0	1.05	0.88	1.24	0.5974	0.9740	462

Appendix 4.9 Full dataset: Risk of high grade prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1]

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
PC aa C38:4	1.05	0.88	1.25	0.5978	0.9740	462
PC aa C32:0	0.95	0.78	1.16	0.6028	0.9740	462
SM C16:1	0.95	0.79	1.14	0.6040	0.9740	462
C5	1.07	0.83	1.38	0.6125	0.9740	137
eaa	0.96	0.80	1.14	0.6303	0.9740	325
PC ae C34:2	0.95	0.78	1.16	0.6371	0.9740	462
PC aa C30:0	0.96	0.81	1.14	0.6377	0.9740	462
Alanine	1.04	0.89	1.21	0.6486	0.9740	462
Lyso PC a C18:2	1.04	0.89	1.21	0.6550	0.9740	462
PC aa C38:3	0.96	0.81	1.14	0.6660	0.9740	462
PC ae C44:4	0.96	0.82	1.14	0.6663	0.9740	462
PC aa C36:1	0.97	0.83	1.13	0.6670	0.9740	462
PC ae C30:2	1.04	0.84	1.29	0.6885	0.9740	455
tyr_phe	0.96	0.80	1.16	0.6890	0.9740	325
PC ae C32:2	0.97	0.82	1.14	0.6988	0.9740	462
PC aa C40:4	0.97	0.83	1.13	0.7012	0.9740	462
PC ae C36:4	1.04	0.85	1.27	0.7182	0.9740	462
glucoaa	0.97	0.81	1.16	0.7242	0.9740	325
Lyso PC a C16:1	0.97	0.83	1.14	0.7288	0.9740	462
C2	1.02	0.89	1.18	0.7313	0.9740	462
PC ae C44:5	0.97	0.83	1.14	0.7421	0.9740	462
Asparagine	1.03	0.86	1.24	0.7501	0.9740	462
PC aa C42:6	0.96	0.75	1.23	0.7598	0.9740	137
Sarcosine	1.02	0.88	1.19	0.7627	0.9740	462
C0	0.98	0.86	1.12	0.7635	0.9740	462
Glycine	0.98	0.84	1.14	0.7777	0.9740	462
C18	1.02	0.88	1.18	0.7895	0.9740	462
PC ae C42:1	0.98	0.85	1.13	0.7952	0.9740	462
cit_orn	1.03	0.84	1.26	0.7958	0.9740	325
Lyso PC a C20:3	1.02	0.87	1.19	0.7966	0.9740	462
C18:1	0.98	0.81	1.18	0.7971	0.9740	462
ketoaa	0.98	0.82	1.17	0.7975	0.9740	325
neaa	0.98	0.81	1.18	0.7977	0.9740	325
Lysine	1.02	0.88	1.18	0.8042	0.9740	462
PC ae C34:1	1.02	0.84	1.25	0.8069	0.9740	462
t4-hydroxyproline	0.98	0.85	1.13	0.8071	0.9740	462
PC ae C38:5	0.97	0.79	1.20	0.8090	0.9740	462
Creatinine	1.02	0.88	1.18	0.8114	0.9740	462
ADMA	1.02	0.87	1.20	0.8121	0.9740	462
C18:2	1.02	0.84	1.25	0.8161	0.9740	462
cit_arg	1.03	0.82	1.29	0.8171	0.9740	325
Tryptophan	0.98	0.84	1.14	0.8242	0.9746	462

Appendix 4.9 Full dataset: Risk of high grade prostate cancer per 1 SD increase in log metabolite concentrations in EPIC[1]

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
SM C18:1	1.02	0.86	1.21	0.8387	0.9794	462
Lyso PC a C18:1	0.98	0.84	1.16	0.8413	0.9794	462
PC ae C38:3	0.99	0.85	1.14	0.8510	0.9829	462
Citrulline	1.01	0.87	1.17	0.8729	0.9897	462
PC ae C36:5	0.99	0.82	1.18	0.8799	0.9897	462
PC ae C40:4	0.99	0.85	1.15	0.8808	0.9897	461
nsaa	0.98	0.79	1.22	0.8843	0.9897	325
Histidine	0.99	0.85	1.15	0.8963	0.9897	462
PC aa C34:1	0.99	0.83	1.18	0.9029	0.9897	462
PC aa C36:2	1.01	0.82	1.25	0.9143	0.9897	462
PC aa C36:3	0.99	0.81	1.21	0.9166	0.9897	462
Methionine	0.99	0.87	1.14	0.9202	0.9897	462
PC ae C40:2	1.01	0.86	1.18	0.9233	0.9897	462
PC ae C32:1	1.01	0.85	1.20	0.9354	0.9909	462
Proline	0.99	0.86	1.15	0.9392	0.9909	462
Glutamate	1.01	0.78	1.30	0.9443	0.9909	462
Arginine	1.01	0.82	1.24	0.9555	0.9913	462
SM C20:2	1.00	0.83	1.20	0.9580	0.9913	462
orn_arg	1.01	0.78	1.30	0.9693	0.9934	325
ureaaa	1.00	0.82	1.20	0.9734	0.9934	325
fisher_ratio	1.00	0.84	1.19	0.9844	0.9951	325
PC ae C44:3	1.00	0.76	1.32	0.9943	0.9951	136
Glutamine	1.00	0.83	1.20	0.9951	0.9951	462

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 Padj < 0.05 are statistically significant after allowing for multiple testing using a false discovery rate controlling procedure at α = 0.05 (Benjamini-Hochberg). P-values are sorted in ascending order.

Appendix 4.10 Full dataset: Risk of prostate cancer death per 1 SD increase in log metabolite concentrations in EPIC[1]

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
Lyso PC a C18:1	0.80	0.67	0.96	0.0141	0.6474	417
Taurine	1.32	1.05	1.66	0.0168	0.6474	417
Lyso PC a C18:2	0.82	0.70	0.97	0.0234	0.6474	417
PC ae C42:3	0.82	0.69	0.98	0.0271	0.6474	417
SM C26:0	1.47	1.04	2.09	0.0312	0.6474	119
Serotonin	1.41	1.02	1.96	0.0379	0.6474	120
PC ae C34:3	0.82	0.68	1.00	0.0448	0.6474	417
C3	1.16	1.00	1.36	0.0527	0.6474	417
PC aa C40:3	0.85	0.72	1.01	0.0626	0.6474	417
cit_orn	0.81	0.65	1.01	0.0667	0.6474	297
PC ae C38:2	0.86	0.74	1.01	0.0676	0.6474	417
Lyso PC a C16:1	0.85	0.71	1.01	0.0677	0.6474	417
C10	1.34	0.98	1.84	0.0698	0.6474	120
Lyso PC a C17:0	0.86	0.74	1.01	0.0733	0.6474	417
isoleucine	1.14	0.99	1.32	0.0770	0.6474	417
cit_arg	0.79	0.61	1.03	0.0791	0.6474	297
C5	1.35	0.96	1.90	0.0847	0.6474	120
PC aa C38:4	1.19	0.98	1.46	0.0852	0.6474	417
PC aa C42:4	0.84	0.70	1.02	0.0864	0.6474	416
Lyso PC a C16:0	0.85	0.71	1.02	0.0890	0.6474	417
PC aa C38:3	1.18	0.97	1.44	0.1034	0.6474	417
tyr_phe	1.18	0.97	1.43	0.1076	0.6474	297
PC aa C40:2	0.87	0.73	1.03	0.1145	0.6474	417
C12:1	1.26	0.94	1.69	0.1153	0.6474	120
C12	1.26	0.94	1.68	0.1176	0.6474	120
Ornithine	1.18	0.96	1.45	0.1240	0.6474	417
Valine	1.13	0.97	1.32	0.1275	0.6474	417
PC ae C36:3	0.85	0.68	1.05	0.1286	0.6474	417
Lyso PC a C18:0	0.87	0.73	1.04	0.1299	0.6474	417
PC ae C34:2	0.85	0.68	1.05	0.1304	0.6474	417
Glycine	0.88	0.75	1.04	0.1350	0.6489	417
C14	1.24	0.93	1.65	0.1394	0.6492	120
Methionine	1.12	0.96	1.30	0.1457	0.6578	417
SM C26:1	0.82	0.61	1.09	0.1746	0.7097	120
C18	1.11	0.95	1.29	0.1770	0.7097	417
Hexose	1.12	0.95	1.33	0.1835	0.7097	417
t4-hydroxyproline	1.10	0.95	1.28	0.1938	0.7097	417
ADMA	1.12	0.94	1.32	0.1980	0.7097	417
PC ae C44:4	0.89	0.74	1.07	0.2038	0.7097	417
PC aa C32:0	1.15	0.93	1.43	0.2065	0.7097	417
ketoaa	1.12	0.94	1.34	0.2094	0.7097	297
C2	1.11	0.94	1.32	0.2160	0.7097	417

Appendix 4.10 Full dataset: Risk of prostate cancer death per 1 SD increase in log metabolite concentrations in EPIC[1]

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
Leucine	1.10	0.95	1.27	0.2174	0.7097	417
Tyrosine	1.10	0.94	1.29	0.2186	0.7097	417
PC aa C36:4	1.15	0.92	1.44	0.2201	0.7097	417
C4	1.28	0.86	1.89	0.2213	0.7097	120
Lysine	1.10	0.94	1.29	0.2239	0.7097	417
fisher_ratio	1.11	0.93	1.33	0.2437	0.7195	297
glucoaa	1.11	0.93	1.34	0.2482	0.7195	297
Lyso PC a C20:4	0.90	0.75	1.08	0.2510	0.7195	417
Proline	1.09	0.94	1.27	0.2620	0.7195	417
PC aa C42:2	0.90	0.75	1.08	0.2621	0.7195	417
eaa	1.11	0.93	1.32	0.2625	0.7195	297
ureaaa	1.12	0.92	1.35	0.2642	0.7195	297
bcaa	1.10	0.93	1.31	0.2656	0.7195	297
neaa	1.11	0.92	1.33	0.2945	0.7568	297
SM C18:0	1.11	0.91	1.35	0.2949	0.7568	417
SM C18:1	1.11	0.91	1.35	0.2964	0.7568	417
Lyso PC a C20:3	0.91	0.76	1.09	0.3018	0.7568	417
PC aa C40:5	1.10	0.92	1.32	0.3107	0.7568	417
PC aa C40:6	1.08	0.93	1.26	0.3118	0.7568	417
C16	1.10	0.92	1.31	0.3149	0.7568	417
PC ae C42:4	0.92	0.77	1.09	0.3224	0.7596	417
Alanine	1.09	0.92	1.28	0.3263	0.7596	417
Tryptophan	1.08	0.92	1.26	0.3431	0.7763	417
Arginine	1.12	0.89	1.41	0.3439	0.7763	417
PC aa C40:4	1.09	0.90	1.32	0.3576	0.7852	417
PC ae C40:5	0.93	0.80	1.09	0.3585	0.7852	417
PC ae C38:4	1.09	0.90	1.33	0.3677	0.7852	417
PC aa C38:5	1.08	0.91	1.28	0.3741	0.7852	417
PC ae C44:3	0.87	0.64	1.18	0.3742	0.7852	120
C0	1.07	0.92	1.25	0.3896	0.8064	417
Threonine	1.06	0.92	1.24	0.4119	0.8408	417
PC aa C34:2	1.12	0.85	1.48	0.4260	0.8577	417
PC aa C38:6	1.06	0.91	1.24	0.4627	0.9116	417
Asparagine	1.07	0.89	1.28	0.4687	0.9116	417
nsaa	1.08	0.87	1.35	0.4954	0.9116	297
PC ae C40:3	0.94	0.79	1.12	0.4999	0.9116	417
PC ae C32:2	0.94	0.78	1.13	0.5142	0.9116	417
PC aa C36:0	0.94	0.77	1.14	0.5205	0.9116	417
C18:1	1.07	0.88	1.30	0.5272	0.9116	417
Lyso PC a C28:1	1.13	0.78	1.63	0.5278	0.9116	118
PC ae C38:5	1.07	0.86	1.33	0.5309	0.9116	417
PC aa C30:0	1.06	0.89	1.26	0.5344	0.9116	417

Appendix 4.10 Full dataset: Risk of prostate cancer death per 1 SD increase in log metabolite concentrations in EPIC[1]

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
PC ae C40:1	0.95	0.81	1.12	0.5425	0.9116	417
PC aa C32:3	0.94	0.78	1.14	0.5496	0.9116	417
PC ae C40:2	0.95	0.81	1.12	0.5563	0.9116	417
PC ae C36:1	0.96	0.82	1.11	0.5571	0.9116	416
PC aa C34:3	0.94	0.78	1.15	0.5603	0.9116	417
PC ae C30:2	1.06	0.86	1.32	0.5604	0.9116	414
Citrulline	0.96	0.82	1.12	0.5726	0.9116	417
PC ae C42:2	0.95	0.81	1.13	0.5836	0.9116	417
PC ae C36:2	0.95	0.78	1.15	0.5844	0.9116	417
PC ae C34:0	1.04	0.89	1.23	0.5898	0.9116	417
Serine	1.04	0.89	1.22	0.5927	0.9116	417
PC aa C36:2	1.06	0.85	1.34	0.5947	0.9116	417
PC ae C36:4	1.06	0.86	1.30	0.6033	0.9116	417
PC aa C34:1	1.05	0.87	1.26	0.6047	0.9116	417
PC aa C36:6	0.96	0.82	1.12	0.6057	0.9116	417
PC ae C42:5	0.96	0.82	1.12	0.6214	0.9177	417
SDMA	0.92	0.67	1.27	0.6274	0.9177	120
PC aa C36:3	1.06	0.84	1.33	0.6358	0.9177	417
PC ae C44:5	1.04	0.89	1.21	0.6400	0.9177	417
PC aa C36:5	1.04	0.88	1.23	0.6485	0.9177	417
SM C16:1	1.05	0.86	1.28	0.6500	0.9177	417
PC aa C28:1	1.04	0.87	1.26	0.6549	0.9177	417
Sarcosine	1.03	0.89	1.21	0.6618	0.9177	417
Creatinine	0.97	0.83	1.12	0.6665	0.9177	417
Glutamate	1.06	0.81	1.39	0.6754	0.9177	417
aromaa	1.04	0.87	1.25	0.6775	0.9177	297
PC ae C44:6	1.03	0.89	1.20	0.6851	0.9196	417
PC aa C34:4	0.96	0.80	1.16	0.7033	0.9263	417
SM(OH) C16:1	1.04	0.86	1.25	0.7040	0.9263	417
Glutamine	1.04	0.85	1.27	0.7088	0.9263	417
PC ae C38:6	0.97	0.82	1.15	0.7195	0.9322	417
PC aa C32:1	1.03	0.86	1.25	0.7276	0.9346	417
orn_arg	0.95	0.71	1.27	0.7434	0.9467	297
SM(OH) C22:1	0.97	0.80	1.18	0.7522	0.9498	417
SM(OH) C14:1	1.03	0.86	1.23	0.7593	0.9508	417
Aspartic acid	0.96	0.72	1.29	0.7906	0.9816	120
SM C24:1	0.96	0.72	1.29	0.8078	0.9905	417
SM C24:0	0.97	0.78	1.21	0.8144	0.9905	417
PC aa C38:0	0.98	0.82	1.16	0.8176	0.9905	417
SM C16:0	0.98	0.80	1.19	0.8287	0.9929	417
PC aa C36:1	1.02	0.86	1.20	0.8468	0.9929	417
SM C20:2	0.98	0.80	1.21	0.8557	0.9929	417

Appendix 4.10 Full dataset: Risk of prostate cancer death per 1 SD increase in log metabolite concentrations in EPIC[1]

Metabolite	OR	LL	UL	P-value[2]	Padj[2]	n_{case}
Histidine	0.99	0.84	1.16	0.8575	0.9929	417
Lyso PC a C24:0	1.04	0.70	1.52	0.8594	0.9929	115
C14:1	1.01	0.87	1.17	0.8640	0.9929	417
PC ae C40:6	1.01	0.86	1.19	0.8663	0.9929	417
Kynurenine	0.99	0.85	1.16	0.9049	0.9935	417
PC ae C40:4	1.01	0.86	1.18	0.9052	0.9935	416
PC aa C42:1	1.01	0.85	1.20	0.9055	0.9935	417
PC aa C42:0	0.99	0.83	1.18	0.9110	0.9935	417
Phenylalanine	1.01	0.87	1.17	0.9112	0.9935	417
PC aa C42:6	1.02	0.75	1.38	0.9147	0.9935	120
PC ae C36:5	1.01	0.83	1.24	0.9195	0.9935	417
PC ae C34:1	0.99	0.81	1.21	0.9212	0.9935	417
PC ae C36:0	1.01	0.86	1.18	0.9275	0.9935	415
PC aa C32:2	1.01	0.75	1.36	0.9467	0.9935	120
PC ae C38:0	0.99	0.73	1.34	0.9487	0.9935	120
SM(OH) C24:1	1.01	0.82	1.23	0.9582	0.9935	417
PC ae C38:3	1.00	0.85	1.16	0.9706	0.9935	417
SM(OH) C22:2	1.00	0.81	1.25	0.9715	0.9935	417
C18:2	1.00	0.82	1.22	0.9773	0.9935	417
PC aa C42:5	1.00	0.83	1.20	0.9819	0.9935	417
PC ae C42:1	1.00	0.85	1.17	0.9830	0.9935	417
PC ae C32:1	1.00	0.82	1.22	0.9912	0.9935	417
PC ae C30:0	1.00	0.83	1.21	0.9935	0.9935	416

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 Padj < 0.05 are statistically significant after allowing for multiple testing using a false discovery rate controlling procedure at α = 0.05 (Benjamini-Hochberg). P-values are sorted in ascending order.

Appendix 4.11 Full dataset: Risk of overall prostate cancer per 1 SD increase in metabolite pattern score in EPIC[1]

Metabolite pattern	OR	LL	UL	P-value[2]	n_{case}
1	0.96	0.92	1.01	0.1453	4314
3	0.98	0.93	1.03	0.4269	4314
2	0.97	0.91	1.04	0.4693	4314

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 P-values are sorted in ascending order.

Appendix 4.12 Full dataset: Risk of advanced prostate cancer per 1 SD increase in metabolite pattern score in EPIC[1]

Metabolite pattern	OR	LL	UL	P-value[2]	n_{case}
2	**0.80**	**0.68**	**0.94**	**0.0077**	**930**
1	**0.88**	**0.79**	**0.99**	**0.0297**	**930**
3	0.94	0.84	1.06	0.3278	930

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 P-values are sorted in ascending order. Associations that have a p-value < 0.05 are bolded.

Appendix 4.13 Full dataset: Risk of advanced prostate cancer per 1 SD increase in metabolite pattern score in EPIC[1] by ≤ 10 years of follow-up time

Metabolite pattern	OR	LL	UL	P-value[2]	n_{case}
1	**0.80**	**0.66**	**0.96**	**0.0162**	**390**
2	**0.76**	**0.59**	**0.97**	**0.0273**	**390**
3	0.93	0.77	1.13	0.4683	390

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 P-values are sorted in ascending order. Associations that have a p-value < 0.05 are bolded.

Appendix 4.14 Full dataset: Risk of advanced prostate cancer per 1 SD increase in metabolite pattern score in EPIC[1] by > 10 years of follow-up time

Metabolite pattern	OR	LL	UL	P-value[2]	n_{case}
2	0.84	0.67	1.06	0.1495	540
1	0.95	0.82	1.10	0.4692	540
3	0.98	0.84	1.15	0.8069	540

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 P-values are sorted in ascending order.

Appendix 4.15 Full dataset: Risk of aggressive prostate cancer per 1 SD increase in metabolite pattern score in EPIC[1]

Metabolite pattern	OR	LL	UL	P-value[2]	n_{case}
1	0.95	0.87	1.04	0.2609	1495
3	0.97	0.88	1.06	0.5114	1495
2	0.99	0.88	1.11	0.8371	1495

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 P-values are sorted in ascending order.

Appendix 4.16 Full dataset: Risk of aggressive prostate cancer per 1 SD increase in metabolite pattern score in EPIC[1] by ≤ 10 years of follow-up time

Metabolite pattern	OR	LL	UL	P-value[2]	n_{case}
3	0.93	0.80	1.07	0.2947	669
1	0.97	0.84	1.10	0.6020	669
2	0.99	0.83	1.17	0.8967	669

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 P-values are sorted in ascending order.

Appendix 4.17 Full dataset: Risk of aggressive prostate cancer per 1 SD increase in metabolite pattern score in EPIC[1] by > 10 years of follow-up time

Metabolite pattern	OR	LL	UL	P-value[2]	n_{case}
1	0.96	0.85	1.08	0.4972	790
3	1.04	0.91	1.17	0.5767	790
2	1.01	0.85	1.20	0.8997	790

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 P-values are sorted in ascending order.

Appendix 4.18 Full dataset: Risk of high grade prostate cancer per 1 SD increase in metabolite pattern score in EPIC[1]

Metabolite pattern	OR	LL	UL	P-value[2]	n_{case}
1	0.92	0.78	1.08	0.2886	452
3	1.04	0.88	1.23	0.6204	452
2	1.02	0.82	1.27	0.8558	452

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 P-values are sorted in ascending order.

Appendix 4.19 Full dataset: Risk of high grade prostate cancer per 1 SD increase in metabolite pattern score in EPIC[1] by ≤ 10 years of follow-up time

Metabolite pattern	OR	LL	UL	P-value[2]	n_{case}
1	0.94	0.72	1.23	0.6652	176
2	1.03	0.73	1.45	0.8578	176
3	1.00	0.73	1.36	0.9988	176

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 P-values are sorted in ascending order.

Appendix 4.20 Full dataset: Risk of high grade prostate cancer per 1 SD increase in metabolite pattern score in EPIC[1] by > 10 years of follow-up time

Metabolite pattern	OR	LL	UL	P-value[2]	n_{case}
3	1.10	0.89	1.37	0.3765	270
1	0.94	0.76	1.16	0.5689	270
2	1.08	0.80	1.46	0.6024	270

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 P-values are sorted in ascending order.

Appendix 4.21 Full dataset: Risk of prostate cancer death per 1 SD increase in metabolite pattern score in EPIC[1]

Metabolite pattern	OR	LL	UL	P-value[2]	n_{case}
3	**0.82**	**0.68**	**0.98**	**0.0278**	**448**
2	1.11	0.89	1.40	0.3507	448
1	0.97	0.82	1.15	0.7265	448

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 P-values are sorted in ascending order. Associations with a P-value < 0.05 are bolded.

Appendix 4.22 Full dataset: Risk of prostate cancer death per 1 SD increase in metabolite pattern score in EPIC[1] by ≤ 10 years of follow-up time

Metabolite pattern	OR	LL	UL	P-value[2]	n_{case}
3	**0.77**	**0.61**	**0.99**	**0.0400**	**241**
2	1.13	0.84	1.52	0.4110	241
1	1.03	0.83	1.29	0.7750	241

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 P-values are sorted in ascending order. Associations with a P-value < 0.05 are bolded.

Appendix 4.23 Full dataset: Risk of prostate cancer death per 1 SD increase in metabolite pattern score in EPIC[1] by > 10 years of follow-up time

Metabolite pattern	OR	LL	UL	P-value[2]	n_{case}
3	0.87	0.65	1.18	0.3864	171
1	0.93	0.69	1.25	0.6267	171
2	1.03	0.69	1.54	0.8935	171

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; SD, standard deviation; OR, odds ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit.

1 All analyses were matched on centre, length of follow-up, and age (±6 months), time of day (±1 hour) and fasting status (<3, 3–6, >6 hours) at blood collection and further adjusted for exact age at blood collection (continuously), and baseline values for body mass index (quartiles; unknown), smoking (never; past; current; unknown), alcohol intake (<10; 10-19; 20-39; ≥40 g of alcohol per day; unknown), education (primary; secondary; degree level; unknown) and marital status (married or cohabiting; not married or cohabiting; unknown).

2 P-values are sorted in ascending order.

Appendix 5.1 Hazard ratios for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1]

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Docosahexaenoic Acid to Total Fatty Acids percentage	**6570**	**1.05**	**1.03**	**1.08**	**9.3E-05**	**2.3E-02**
Glucose	**6560**	**0.95**	**0.93**	**0.98**	**1.8E-04**	**2.3E-02**
Degree of Unsaturation	**6570**	**1.05**	**1.02**	**1.07**	**4.1E-04**	**3.4E-02**
Phospholipids to Total Lipids in Small LDL percentage	**6575**	**1.06**	**1.02**	**1.09**	**5.6E-04**	**3.5E-02**
Polyunsaturated Fatty Acids to Monounsaturated Fatty Acids ratio	**6570**	**1.05**	**1.02**	**1.08**	**7.3E-04**	**3.6E-02**
Polyunsaturated Fatty Acids to Total Fatty Acids percentage	**6570**	**1.05**	**1.02**	**1.08**	**1.1E-03**	**4.5E-02**
Monounsaturated Fatty Acids to Total Fatty Acids percentage	**6575**	**0.96**	**0.93**	**0.98**	**1.3E-03**	**4.5E-02**
Omega-3 Fatty Acids to Total Fatty Acids percentage	6575	1.04	1.02	1.07	1.8E-03	5.5E-02
Triglycerides to Total Lipids in Large LDL percentage	6570	0.96	0.93	0.99	2.4E-03	6.6E-02
Triglycerides in HDL	6570	0.96	0.94	0.99	4.3E-03	6.8E-02
Docosahexaenoic Acid	6575	1.04	1.01	1.07	4.3E-03	6.8E-02
Triglycerides in Medium HDL	6575	0.96	0.94	0.99	4.9E-03	6.8E-02
Cholesterol to Total Lipids in Very Small VLDL percentage	6575	1.04	1.01	1.07	5.2E-03	6.8E-02
Cholesteryl Esters to Total Lipids in Very Small VLDL percentage	6575	1.04	1.01	1.07	5.2E-03	6.8E-02
Cholesteryl Esters to Total Lipids in IDL percentage	6575	1.04	1.01	1.07	4.4E-03	6.8E-02
Triglycerides to Total Lipids in IDL percentage	6575	0.96	0.94	0.99	5.0E-03	6.8E-02
Triglycerides to Total Lipids in Medium LDL percentage	6575	0.96	0.94	0.99	4.1E-03	6.8E-02
Triglycerides to Total Lipids in Small LDL percentage	6575	0.96	0.94	0.99	4.2E-03	6.8E-02
Cholesteryl Esters to Total Lipids in Large HDL percentage	6575	1.05	1.02	1.09	5.0E-03	6.8E-02
Free Cholesterol to Total Lipids in Medium LDL percentage	6575	1.07	1.02	1.12	5.7E-03	7.1E-02
Cholesterol to Total Lipids in IDL percentage	6575	1.04	1.01	1.06	6.2E-03	7.4E-02
Triglycerides in LDL	6575	0.97	0.94	0.99	1.1E-02	8.7E-02
Monounsaturated Fatty Acids	6570	0.97	0.94	0.99	7.8E-03	8.7E-02
Triglycerides in Large LDL	6575	0.97	0.94	0.99	1.1E-02	8.7E-02
Triglycerides in Medium LDL	6575	0.97	0.94	0.99	1.1E-02	8.7E-02
Triglycerides in Large HDL	6575	0.97	0.94	0.99	8.4E-03	8.7E-02
Triglycerides in Small HDL	6575	0.97	0.94	0.99	1.1E-02	8.7E-02
Cholesterol to Total Lipids in Medium VLDL percentage	6575	1.04	1.01	1.06	8.8E-03	8.7E-02
Cholesteryl Esters to Total Lipids in Medium VLDL percentage	6575	1.04	1.01	1.07	9.4E-03	8.7E-02
Free Cholesterol to Total Lipids in Small VLDL percentage	6575	1.04	1.01	1.06	8.8E-03	8.7E-02

Appendix 5.1 Hazard ratios for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1]

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Cholesterol to Total Lipids in Large LDL percentage	6575	1.20	1.04	1.39	1.1E-02	8.7E-02
Free Cholesterol to Total Lipids in Small LDL percentage	6570	1.05	1.01	1.10	1.1E-02	8.7E-02
Cholesterol to Total Lipids in Large HDL percentage	6575	1.05	1.01	1.08	1.1E-02	8.7E-02
Triglycerides to Total Lipids in Very Small VLDL percentage	6575	0.97	0.94	0.99	1.2E-02	9.0E-02
Omega-6 Fatty Acids to Omega-3 Fatty Acids ratio	6575	0.97	0.94	0.99	1.4E-02	9.7E-02
Phospholipids to Total Lipids in Small VLDL percentage	6575	1.03	1.01	1.06	1.5E-02	1.0E-01
Free Cholesterol to Total Lipids in Very Small VLDL percentage	6575	1.03	1.01	1.06	1.7E-02	1.2E-01
Alanine	6575	0.97	0.95	1.00	2.1E-02	1.4E-01
Tyrosine	6575	0.97	0.95	1.00	2.2E-02	1.4E-01
Triglycerides in IDL	6575	0.97	0.95	1.00	2.3E-02	1.4E-01
Triglycerides in Small LDL	6575	0.97	0.95	1.00	2.4E-02	1.4E-01
Free Cholesterol to Total Lipids in Medium VLDL percentage	6575	1.03	1.00	1.06	2.3E-02	1.4E-01
Triglycerides to Total Lipids in Very Large HDL percentage	6572	0.97	0.95	1.00	2.4E-02	1.4E-01
Triglycerides to Total Lipids in Large HDL percentage	6575	0.97	0.94	1.00	2.5E-02	1.4E-01
Cholesterol to Total Lipids in Small HDL percentage	6575	1.03	1.00	1.06	2.4E-02	1.4E-01
Cholesterol in Chylomicrons and Extremely Large VLDL	6575	0.97	0.95	1.00	2.9E-02	1.5E-01
Cholesteryl Esters in Medium VLDL	6570	1.03	1.00	1.06	2.9E-02	1.5E-01
Phospholipids in Small HDL	6568	0.97	0.95	1.00	2.9E-02	1.5E-01
Saturated Fatty Acids	6575	0.97	0.95	1.00	3.4E-02	1.7E-01
Cholesteryl Esters in Chylomicrons and Extremely Large VLDL	6575	0.97	0.95	1.00	3.5E-02	1.7E-01
Omega-6 Fatty Acids to Total Fatty Acids percentage	6575	1.03	1.00	1.06	3.7E-02	1.7E-01
Triglycerides in Very Small VLDL	6575	0.97	0.95	1.00	3.6E-02	1.7E-01
Triglycerides in Very Large HDL	6575	0.97	0.95	1.00	3.7E-02	1.7E-01
Average Diameter for LDL Particles	6570	1.03	1.00	1.05	3.8E-02	1.7E-01
Cholesteryl Esters to Total Lipids in Small HDL percentage	6571	1.03	1.00	1.06	3.8E-02	1.7E-01
Cholesterol to Total Lipids in Small VLDL percentage	6575	1.03	1.00	1.05	3.9E-02	1.7E-01
Total Lipids in Small HDL	6575	0.97	0.95	1.00	4.2E-02	1.8E-01
Free Cholesterol to Total Lipids in Large LDL percentage	6569	1.06	1.00	1.12	4.5E-02	1.9E-01
Cholesterol to Total Lipids in Medium LDL percentage	6575	1.14	1.00	1.29	4.6E-02	1.9E-01

Appendix 5.1 Hazard ratios for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1]

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Total Fatty Acids	6575	0.97	0.95	1.00	4.9E-02	2.0E-01
Free Cholesterol in Chylomicrons and Extremely Large VLDL	6575	0.98	0.95	1.00	4.9E-02	2.0E-01
Phospholipids in Medium HDL	6575	0.97	0.95	1.00	5.3E-02	2.1E-01
Cholesteryl Esters to Total Lipids in Very Large VLDL percentage	6575	1.03	1.00	1.05	5.6E-02	2.2E-01
Cholesteryl Esters to Total Lipids in Medium HDL percentage	6570	1.03	1.00	1.06	5.6E-02	2.2E-01
Triglycerides to Total Lipids in Medium HDL percentage	6575	0.98	0.95	1.00	5.9E-02	2.2E-01
Triglycerides to Total Lipids in Small HDL percentage	6575	0.98	0.95	1.00	5.8E-02	2.2E-01
Citrate	6575	0.98	0.95	1.00	6.6E-02	2.4E-01
Cholesteryl Esters to Total Lipids in Large LDL percentage	6575	1.11	0.99	1.25	6.6E-02	2.4E-01
Cholesteryl Esters to Total Lipids in Large VLDL percentage	6570	1.02	1.00	1.05	7.5E-02	2.6E-01
Phospholipids to Total Lipids in Medium VLDL percentage	6575	1.03	1.00	1.06	7.5E-02	2.6E-01
Triglycerides to Total Lipids in Medium VLDL percentage	6548	0.98	0.95	1.00	7.3E-02	2.6E-01
Triglycerides to Total Lipids in Small VLDL percentage	6575	0.98	0.95	1.00	7.3E-02	2.6E-01
Phosphoglycerides	6575	0.98	0.95	1.00	8.0E-02	2.7E-01
Total Lipids in Medium HDL	6570	0.98	0.95	1.00	8.0E-02	2.7E-01
Cholesterol to Total Lipids in Medium HDL percentage	6575	1.03	1.00	1.06	8.0E-02	2.7E-01
Free Cholesterol in Small HDL	6575	0.98	0.95	1.00	8.2E-02	2.7E-01
Total Triglycerides	6575	0.98	0.95	1.00	9.1E-02	2.9E-01
Free Cholesterol in Small LDL	6575	1.02	1.00	1.05	9.2E-02	2.9E-01
Phospholipids to Total Lipids in Medium LDL percentage	6570	1.02	1.00	1.05	9.3E-02	2.9E-01
Omega-3 Fatty Acids	6575	1.02	1.00	1.05	9.6E-02	3.0E-01
Cholesteryl Esters to Total Lipids in Very Large HDL percentage	6575	1.03	0.99	1.06	9.8E-02	3.0E-01
Phosphatidylcholines	6575	0.98	0.95	1.00	1.1E-01	3.3E-01
Valine	6575	0.98	0.95	1.00	1.1E-01	3.3E-01
Saturated Fatty Acids to Total Fatty Acids percentage	6570	0.98	0.95	1.01	1.1E-01	3.4E-01
Cholesterol in Medium VLDL	6572	1.02	0.99	1.04	1.2E-01	3.5E-01
Cholesterol to Total Lipids in Very Large VLDL percentage	6575	1.02	0.99	1.05	1.2E-01	3.5E-01
Apolipoprotein A1	6575	0.98	0.95	1.01	1.2E-01	3.6E-01
Cholesteryl Esters in Very Large HDL	6548	1.02	0.99	1.05	1.3E-01	3.6E-01
Cholesterol in Very Large HDL	6575	1.02	0.99	1.05	1.3E-01	3.6E-01
Concentration of Medium HDL Particles	6575	0.98	0.96	1.01	1.3E-01	3.6E-01

Appendix 5.1 Hazard ratios for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1]

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Cholesteryl Esters to Total Lipids in Small VLDL percentage	6575	1.02	0.99	1.05	1.3E-01	3.7E-01
Phospholipids in Chylomicrons and Extremely Large VLDL	6575	0.98	0.96	1.01	1.4E-01	3.8E-01
Phospholipids in HDL	6570	0.98	0.95	1.01	1.5E-01	3.9E-01
Average Diameter for VLDL Particles	6575	0.98	0.96	1.01	1.5E-01	3.9E-01
Cholesteryl Esters in Large HDL	6570	1.02	0.99	1.05	1.5E-01	3.9E-01
Free Cholesterol to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	6575	0.98	0.96	1.01	1.6E-01	4.2E-01
Leucine	6575	0.98	0.96	1.01	1.6E-01	4.2E-01
Concentration of Chylomicrons and Extremely Large VLDL Particles	6575	0.98	0.96	1.01	1.7E-01	4.2E-01
Triglycerides in Small VLDL	6570	0.98	0.96	1.01	1.7E-01	4.2E-01
Glycoprotein Acetyls	6549	1.02	0.99	1.04	1.7E-01	4.3E-01
Albumin	6575	0.98	0.96	1.01	1.7E-01	4.3E-01
Total Concentration of Branched-Chain Amino Acids (Leucine + Isoleucine + Valine)	6575	0.98	0.96	1.01	1.8E-01	4.3E-01
Total Cholines	6367	0.98	0.96	1.01	1.8E-01	4.3E-01
Phospholipids to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	6575	0.98	0.96	1.01	1.8E-01	4.3E-01
Free Cholesterol to Total Lipids in Large HDL percentage	6575	1.02	0.99	1.05	1.8E-01	4.3E-01
Cholesterol to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	6575	0.98	0.96	1.01	1.8E-01	4.3E-01
Triglycerides in VLDL	6574	0.98	0.96	1.01	2.0E-01	4.5E-01
Concentration of HDL Particles	6575	0.98	0.96	1.01	2.0E-01	4.5E-01
Total Lipids in Chylomicrons and Extremely Large VLDL	6575	0.98	0.96	1.01	2.0E-01	4.5E-01
Phospholipids to Total Lipids in Small HDL percentage	6575	0.98	0.96	1.01	1.9E-01	4.5E-01
Glutamine	6575	0.98	0.96	1.01	2.0E-01	4.5E-01
Free Cholesterol in Medium HDL	6367	0.98	0.96	1.01	2.0E-01	4.5E-01
Concentration of Small HDL Particles	6367	0.98	0.96	1.01	2.0E-01	4.5E-01
Glycine	6570	0.98	0.96	1.01	2.1E-01	4.6E-01
Cholesterol in Large HDL	6575	1.02	0.99	1.04	2.2E-01	4.7E-01
Free Cholesterol in LDL	6575	1.02	0.99	1.05	2.2E-01	4.7E-01
Total Lipids in HDL	6575	0.98	0.96	1.01	2.2E-01	4.7E-01
Apolipoprotein B to Apolipoprotein A1 ratio	6575	1.02	0.99	1.04	2.3E-01	4.7E-01
Linoleic Acid	6575	0.98	0.96	1.01	2.3E-01	4.7E-01
Free Cholesterol in Large LDL	6367	1.02	0.99	1.04	2.3E-01	4.7E-01

Appendix 5.1 Hazard ratios for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1]

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Free Cholesterol in Medium LDL	6575	1.02	0.99	1.05	2.3E-01	4.7E-01
Phospholipids in Very Large HDL	6575	1.02	0.99	1.04	2.3E-01	4.7E-01
Free Cholesterol in Very Large HDL	6575	1.02	0.99	1.04	2.2E-01	4.7E-01
Omega-6 Fatty Acids	6570	0.99	0.96	1.01	2.4E-01	4.7E-01
Linoleic Acid to Total Fatty Acids percentage	6540	1.02	0.99	1.04	2.4E-01	4.7E-01
Cholesteryl Esters in IDL	6575	1.01	0.99	1.04	2.4E-01	4.7E-01
Triglycerides to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	6570	1.02	0.99	1.04	2.4E-01	4.7E-01
Triglycerides to Phosphoglycerides ratio	6575	0.98	0.96	1.01	2.4E-01	4.8E-01
Total Concentration of Lipoprotein Particles	6575	0.98	0.96	1.01	2.5E-01	4.8E-01
Free Cholesterol to Total Lipids in IDL percentage	6575	1.02	0.99	1.04	2.5E-01	4.8E-01
Histidine	6575	1.01	0.99	1.04	2.6E-01	4.8E-01
Cholesterol in IDL	6575	1.01	0.99	1.04	2.6E-01	4.8E-01
Phospholipids in Small LDL	6575	1.01	0.99	1.04	2.6E-01	4.8E-01
Cholesterol in Small HDL	6367	0.99	0.96	1.01	2.6E-01	4.8E-01
Phospholipids to Total Lipids in Very Large VLDL percentage	6575	0.99	0.97	1.01	2.5E-01	4.8E-01
Cholesterol to Total Lipids in Very Large HDL percentage	6575	1.02	0.99	1.04	2.6E-01	4.8E-01
Cholesterol to Total Lipids in Large VLDL percentage	6572	1.01	0.99	1.04	2.7E-01	4.9E-01
Cholesteryl Esters to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	6572	0.99	0.96	1.01	2.8E-01	5.1E-01
Total Phospholipids in Lipoprotein Particles	6575	0.99	0.96	1.01	2.9E-01	5.1E-01
Total Lipids in VLDL	6575	0.99	0.96	1.01	3.0E-01	5.3E-01
Concentration of Very Large VLDL Particles	6575	0.99	0.96	1.01	3.2E-01	5.6E-01
Cholesterol in Large LDL	6575	1.01	0.99	1.04	3.2E-01	5.6E-01
Cholesterol in Medium HDL	6517	0.99	0.96	1.01	3.2E-01	5.6E-01
Phospholipids to Total Lipids in Very Large HDL percentage	6575	1.01	0.99	1.04	3.2E-01	5.6E-01
Phospholipids to Total Lipids in Very Small VLDL percentage	6575	0.99	0.96	1.01	3.4E-01	5.8E-01
Total Lipids in Very Large HDL	6570	1.01	0.99	1.04	3.4E-01	5.8E-01
Total Lipids in Very Large VLDL	6570	0.99	0.96	1.01	3.4E-01	5.8E-01
Free Cholesterol in IDL	6567	1.01	0.99	1.04	3.5E-01	5.8E-01
Clinical LDL Cholesterol	6575	1.01	0.99	1.04	3.6E-01	5.9E-01

Abbreviations: SD, standard deviation; HR, hazard ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit; LDL, low-density lipoprotein cholesterol; HDL,

Appendix 5.1 Hazard ratios for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1]

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]

high-density lipoprotein cholesterol; IDL, intermediate-density lipoprotein cholesterol; VLDL, very low-density lipoprotein cholesterol.

1 Hazard ratios (95% confidence interval) for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features. Analyses were stratified by region and Townsend Deprivation Index with age as the underlying time variable. Adjusted for fasting status, body mass index, height, alcohol intake, family history of prostate cancer, ethnicity, level of education, smoking status, and weekly physical activity.

2 Padj < 0.05 are statistically significant after allowing for multiple testing using a false discovery rate controlling procedure at α = 0.05 (Benjamini-Hochberg). Padj values are sorted in ascending order. Associations with Padj < 0.05 are bolded.

Appendix 5.2 Hazard ratios for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for PSA testing.

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Docosahexaenoic Acid to Total Fatty Acids percentage	**6570**	**1.06**	**1.03**	**1.09**	**9.1E-05**	**1.7E-02**
Glucose	**6560**	**0.95**	**0.93**	**0.98**	**1.3E-04**	**1.7E-02**
Degree of Unsaturation	**6570**	**1.05**	**1.02**	**1.08**	**6.3E-04**	**3.9E-02**
Phospholipids to Total Lipids in Small LDL percentage	**6575**	**1.06**	**1.02**	**1.09**	**5.3E-04**	**3.9E-02**
Polyunsaturated Fatty Acids to Total Fatty Acids percentage	**6570**	**1.05**	**1.02**	**1.08**	**1.1E-03**	**4.6E-02**
Polyunsaturated Fatty Acids to Monounsaturated Fatty Acids ratio	**6570**	**1.05**	**1.02**	**1.08**	**1.1E-03**	**4.6E-02**
Triglycerides to Total Lipids in Large LDL percentage	**6575**	**0.96**	**0.93**	**0.98**	**1.4E-03**	**4.8E-02**
Triglycerides in HDL	6575	0.96	0.94	0.99	3.2E-03	5.7E-02
Docosahexaenoic Acid	6570	1.04	1.01	1.07	3.7E-03	5.7E-02
Monounsaturated Fatty Acids to Total Fatty Acids percentage	6570	0.96	0.93	0.99	2.7E-03	5.7E-02
Triglycerides in Medium HDL	6575	0.96	0.94	0.99	3.5E-03	5.7E-02
Cholesteryl Esters to Total Lipids in Medium VLDL percentage	6575	1.04	1.01	1.07	4.6E-03	5.7E-02
Cholesterol to Total Lipids in Very Small VLDL percentage	6575	1.04	1.01	1.07	4.5E-03	5.7E-02
Cholesteryl Esters to Total Lipids in Very Small VLDL percentage	6575	1.04	1.01	1.07	4.6E-03	5.7E-02
Cholesterol to Total Lipids in IDL percentage	6575	1.04	1.01	1.07	4.8E-03	5.7E-02
Cholesteryl Esters to Total Lipids in IDL percentage	6575	1.04	1.01	1.07	3.9E-03	5.7E-02
Triglycerides to Total Lipids in IDL percentage	6575	0.96	0.93	0.99	3.4E-03	5.7E-02
Cholesterol to Total Lipids in Large LDL percentage	6575	1.25	1.08	1.45	3.5E-03	5.7E-02
Free Cholesterol to Total Lipids in Medium LDL percentage	6575	1.08	1.03	1.13	2.4E-03	5.7E-02
Triglycerides to Total Lipids in Medium LDL percentage	6575	0.96	0.93	0.99	2.7E-03	5.7E-02
Triglycerides to Total Lipids in Small LDL percentage	6575	0.96	0.93	0.99	3.1E-03	5.7E-02
Free Cholesterol to Total Lipids in Small VLDL percentage	6575	1.04	1.01	1.07	5.3E-03	6.1E-02
Omega-3 Fatty Acids to Total Fatty Acids percentage	6570	1.04	1.01	1.06	6.6E-03	6.1E-02
Triglycerides in Large HDL	6575	0.96	0.94	0.99	6.5E-03	6.1E-02
Cholesterol to Total Lipids in Medium VLDL percentage	6575	1.04	1.01	1.07	5.9E-03	6.1E-02
Free Cholesterol to Total Lipids in Small LDL percentage	6575	1.06	1.02	1.11	6.5E-03	6.1E-02
Cholesteryl Esters to Total Lipids in Large HDL percentage	6575	1.05	1.02	1.09	5.7E-03	6.1E-02

Appendix 5.2 Hazard ratios for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for PSA testing.

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Triglycerides to Total Lipids in Very Small VLDL percentage	6575	0.97	0.94	0.99	8.6E-03	7.6E-02
Triglycerides in Small HDL	6575	0.97	0.94	0.99	8.9E-03	7.7E-02
Phospholipids to Total Lipids in Small VLDL percentage	6575	1.04	1.01	1.06	9.3E-03	7.7E-02
Triglycerides in LDL	6575	0.97	0.94	0.99	1.1E-02	8.6E-02
Monounsaturated Fatty Acids	6570	0.97	0.94	0.99	1.2E-02	8.6E-02
Triglycerides in Large LDL	6575	0.97	0.94	0.99	1.2E-02	8.6E-02
Triglycerides in Medium LDL	6575	0.97	0.94	0.99	1.1E-02	8.6E-02
Free Cholesterol to Total Lipids in Very Small VLDL percentage	6575	1.03	1.01	1.06	1.3E-02	8.6E-02
Cholesterol to Total Lipids in Large HDL percentage	6575	1.05	1.01	1.09	1.2E-02	8.6E-02
Cholesterol to Total Lipids in Small HDL percentage	6575	1.03	1.01	1.06	1.3E-02	8.6E-02
Triglycerides to Total Lipids in Large HDL percentage	6575	0.97	0.94	0.99	1.6E-02	1.0E-01
Average Diameter for LDL Particles	6575	1.03	1.01	1.06	1.7E-02	1.0E-01
Cholesteryl Esters in Medium VLDL	6575	1.03	1.01	1.06	1.6E-02	1.0E-01
Free Cholesterol to Total Lipids in Medium VLDL percentage	6575	1.03	1.01	1.06	1.8E-02	1.1E-01
Cholesteryl Esters to Total Lipids in Small HDL percentage	6575	1.04	1.01	1.07	1.9E-02	1.1E-01
Triglycerides to Total Lipids in Very Large HDL percentage	6572	0.97	0.94	1.00	2.1E-02	1.2E-01
Triglycerides in IDL	6575	0.97	0.94	1.00	2.3E-02	1.3E-01
Triglycerides in Small LDL	6575	0.97	0.94	1.00	2.4E-02	1.3E-01
Free Cholesterol to Total Lipids in Large LDL percentage	6575	1.07	1.01	1.13	2.4E-02	1.3E-01
Omega-6 Fatty Acids to Total Fatty Acids percentage	6570	1.03	1.00	1.06	2.5E-02	1.3E-01
Tyrosine	6568	0.97	0.95	1.00	2.7E-02	1.4E-01
Cholesterol to Total Lipids in Small VLDL percentage	6575	1.03	1.00	1.06	2.7E-02	1.4E-01
Cholesterol to Total Lipids in Medium LDL percentage	6575	1.16	1.02	1.32	2.8E-02	1.4E-01
Cholesteryl Esters to Total Lipids in Medium HDL percentage	6575	1.04	1.00	1.07	3.1E-02	1.5E-01
Phospholipids in Small HDL	6575	0.97	0.95	1.00	3.2E-02	1.6E-01
Cholesteryl Esters to Total Lipids in Large LDL percentage	6575	1.14	1.01	1.28	3.4E-02	1.6E-01
Saturated Fatty Acids	6570	0.97	0.95	1.00	3.6E-02	1.6E-01
Alanine	6571	0.97	0.95	1.00	3.6E-02	1.6E-01
Triglycerides in Very Small VLDL	6575	0.97	0.95	1.00	3.6E-02	1.6E-01
Triglycerides in Very Large HDL	6575	0.97	0.95	1.00	3.8E-02	1.7E-01

Appendix 5.2 Hazard ratios for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for PSA testing.

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Omega-6 Fatty Acids to Omega-3 Fatty Acids ratio	6569	0.97	0.95	1.00	4.0E-02	1.7E-01
Cholesterol in Chylomicrons and Extremely Large VLDL	6575	0.97	0.95	1.00	4.0E-02	1.7E-01
Triglycerides to Total Lipids in Medium HDL percentage	6575	0.97	0.95	1.00	4.0E-02	1.7E-01
Triglycerides to Total Lipids in Small HDL percentage	6575	0.97	0.95	1.00	4.3E-02	1.7E-01
Cholesterol to Total Lipids in Medium HDL percentage	6575	1.03	1.00	1.07	4.4E-02	1.7E-01
Cholesteryl Esters in Chylomicrons and Extremely Large VLDL	6575	0.97	0.95	1.00	4.5E-02	1.8E-01
Valine	6570	0.97	0.95	1.00	4.7E-02	1.8E-01
Total Lipids in Small HDL	6575	0.97	0.95	1.00	4.9E-02	1.9E-01
Triglycerides to Total Lipids in Small VLDL percentage	6575	0.97	0.95	1.00	5.3E-02	2.0E-01
Triglycerides to Total Lipids in Medium VLDL percentage	6575	0.98	0.95	1.00	5.6E-02	2.1E-01
Phospholipids in Medium HDL	6575	0.97	0.95	1.00	5.9E-02	2.2E-01
Total Fatty Acids	6570	0.98	0.95	1.00	6.1E-02	2.2E-01
Free Cholesterol in Small LDL	6575	1.03	1.00	1.06	6.2E-02	2.2E-01
Cholesteryl Esters to Total Lipids in Very Large VLDL percentage	6548	1.03	1.00	1.05	6.3E-02	2.2E-01
Phospholipids to Total Lipids in Medium VLDL percentage	6575	1.03	1.00	1.06	6.6E-02	2.3E-01
Cholesteryl Esters to Total Lipids in Large VLDL percentage	6575	1.02	1.00	1.05	6.8E-02	2.3E-01
Saturated Fatty Acids to Total Fatty Acids percentage	6570	0.98	0.95	1.00	7.1E-02	2.4E-01
Free Cholesterol in Chylomicrons and Extremely Large VLDL	6575	0.98	0.95	1.00	7.2E-02	2.4E-01
Phospholipids to Total Lipids in Medium LDL percentage	6575	1.02	1.00	1.05	7.6E-02	2.5E-01
Leucine	6575	0.98	0.95	1.00	8.5E-02	2.8E-01
Total Triglycerides	6575	0.98	0.95	1.00	8.7E-02	2.8E-01
Total Concentration of Branched-Chain Amino Acids (Leucine + Isoleucine + Valine)	6570	0.98	0.95	1.00	8.8E-02	2.8E-01
Cholesterol in Medium VLDL	6575	1.02	1.00	1.05	9.6E-02	2.9E-01
Total Lipids in Medium HDL	6575	0.98	0.95	1.00	9.6E-02	2.9E-01
Free Cholesterol in Small HDL	6575	0.98	0.95	1.00	9.8E-02	3.0E-01
Cholesteryl Esters to Total Lipids in Small VLDL percentage	6575	1.02	1.00	1.05	1.0E-01	3.1E-01
Phosphoglycerides	6570	0.98	0.95	1.00	1.1E-01	3.3E-01
Cholesteryl Esters to Total Lipids in Very Large HDL percentage	6572	1.03	0.99	1.06	1.1E-01	3.3E-01
Cholesterol in Very Large HDL	6575	1.02	0.99	1.05	1.2E-01	3.5E-01

Appendix 5.2 Hazard ratios for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for PSA testing.

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Glycoprotein Acetyls	6575	1.02	0.99	1.05	1.2E-01	3.6E-01
Cholesterol to Total Lipids in Very Large VLDL percentage	6548	1.02	0.99	1.05	1.3E-01	3.6E-01
Average Diameter for VLDL Particles	6575	0.98	0.95	1.01	1.3E-01	3.7E-01
Cholesteryl Esters in Very Large HDL	6575	1.02	0.99	1.05	1.3E-01	3.7E-01
Cholesteryl Esters in Large HDL	6575	1.02	0.99	1.05	1.3E-01	3.7E-01
Free Cholesterol in LDL	6575	1.02	0.99	1.05	1.6E-01	4.0E-01
Phosphatidylcholines	6570	0.98	0.96	1.01	1.5E-01	4.0E-01
Apolipoprotein A1	6575	0.98	0.95	1.01	1.6E-01	4.0E-01
Linoleic Acid to Total Fatty Acids percentage	6570	1.02	0.99	1.05	1.6E-01	4.0E-01
Citrate	6575	0.98	0.96	1.01	1.5E-01	4.0E-01
Triglycerides in Small VLDL	6575	0.98	0.96	1.01	1.6E-01	4.0E-01
Concentration of Medium HDL Particles	6575	0.98	0.96	1.01	1.6E-01	4.0E-01
Omega-3 Fatty Acids	6570	1.02	0.99	1.05	1.7E-01	4.1E-01
Glutamine	6549	0.98	0.96	1.01	1.7E-01	4.1E-01
Free Cholesterol in Large LDL	6575	1.02	0.99	1.05	1.7E-01	4.1E-01
Free Cholesterol in Medium LDL	6575	1.02	0.99	1.05	1.7E-01	4.1E-01
Cholesterol to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	6367	0.98	0.96	1.01	1.6E-01	4.1E-01
Free Cholesterol to Total Lipids in Large HDL percentage	6575	1.02	0.99	1.05	1.7E-01	4.1E-01
Phospholipids to Total Lipids in Small HDL percentage	6575	0.98	0.96	1.01	1.6E-01	4.1E-01
Phospholipids in HDL	6575	0.98	0.96	1.01	1.8E-01	4.2E-01
Albumin	6574	0.98	0.96	1.01	1.8E-01	4.2E-01
Free Cholesterol to Total Lipids in IDL percentage	6575	1.02	0.99	1.04	1.9E-01	4.4E-01
Cholesteryl Esters in IDL	6575	1.02	0.99	1.04	1.9E-01	4.4E-01
Triglycerides in VLDL	6575	0.98	0.96	1.01	1.9E-01	4.4E-01
Cholesterol in Large HDL	6575	1.02	0.99	1.05	2.0E-01	4.4E-01
Free Cholesterol to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	6367	0.98	0.96	1.01	2.0E-01	4.4E-01
Triglycerides to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	6367	1.02	0.99	1.04	2.0E-01	4.4E-01
Triglycerides to Phosphoglycerides ratio	6570	0.98	0.96	1.01	2.0E-01	4.4E-01
Cholesterol in IDL	6575	1.02	0.99	1.04	2.1E-01	4.4E-01
Phospholipids in Small LDL	6575	1.02	0.99	1.04	2.1E-01	4.4E-01
Phospholipids in Chylomicrons and Extremely Large VLDL	6575	0.98	0.96	1.01	2.1E-01	4.5E-01

Appendix 5.2 Hazard ratios for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for PSA testing.

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Apolipoprotein B to Apolipoprotein A1 ratio	6575	1.02	0.99	1.04	2.1E-01	4.5E-01
Phospholipids in Very Large HDL	6575	1.02	0.99	1.05	2.1E-01	4.5E-01
Cholesteryl Esters to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	6367	0.98	0.96	1.01	2.2E-01	4.6E-01
Free Cholesterol in Very Large HDL	6575	1.02	0.99	1.04	2.2E-01	4.6E-01
Phospholipids to Total Lipids in Medium HDL percentage	6575	0.99	0.96	1.01	2.3E-01	4.7E-01
Concentration of Chylomicrons and Extremely Large VLDL Particles	6575	0.98	0.96	1.01	2.4E-01	4.8E-01
Total Cholines	6570	0.98	0.96	1.01	2.5E-01	4.9E-01
Glycine	6540	0.99	0.96	1.01	2.5E-01	4.9E-01
Free Cholesterol in Medium HDL	6575	0.98	0.96	1.01	2.5E-01	4.9E-01
Sphingomyelins	6570	1.02	0.99	1.04	2.5E-01	4.9E-01
Cholesterol to Total Lipids in Large VLDL percentage	6575	1.01	0.99	1.04	2.5E-01	4.9E-01
Cholesterol in Large LDL	6575	1.02	0.99	1.05	2.6E-01	5.0E-01
Concentration of HDL Particles	6575	0.98	0.96	1.01	2.6E-01	5.0E-01
Concentration of Small HDL Particles	6575	0.98	0.96	1.01	2.6E-01	5.0E-01
Total Lipids in Chylomicrons and Extremely Large VLDL	6575	0.99	0.96	1.01	2.6E-01	5.0E-01
Free Cholesterol in IDL	6575	1.01	0.99	1.04	2.7E-01	5.0E-01
Phospholipids to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	6367	0.99	0.96	1.01	2.7E-01	5.0E-01
Total Lipids in HDL	6575	0.99	0.96	1.01	2.7E-01	5.1E-01
Clinical LDL Cholesterol	6575	1.01	0.99	1.04	2.8E-01	5.1E-01
Phospholipids to Total Lipids in Very Large HDL percentage	6572	1.02	0.99	1.04	2.9E-01	5.2E-01
Cholesterol to Total Lipids in Very Large HDL percentage	6572	1.01	0.99	1.04	3.0E-01	5.4E-01
Total Lipids in VLDL	6575	0.99	0.96	1.01	3.1E-01	5.5E-01
Cholesteryl Esters in Very Small VLDL	6575	1.01	0.99	1.04	3.1E-01	5.5E-01
Total Lipids in Very Large HDL	6575	1.01	0.99	1.04	3.1E-01	5.5E-01
Cholesteryl Esters in Large LDL	6575	1.01	0.99	1.04	3.2E-01	5.5E-01
Acetoacetate	6517	1.01	0.99	1.04	3.2E-01	5.6E-01
Concentration of IDL Particles	6575	1.01	0.99	1.04	3.2E-01	5.6E-01
Total Concentration of Lipoprotein Particles	6575	0.99	0.96	1.01	3.4E-01	5.6E-01
Omega-6 Fatty Acids	6570	0.99	0.96	1.01	3.3E-01	5.6E-01
Linoleic Acid	6570	0.99	0.96	1.01	3.3E-01	5.6E-01
Histidine	6567	1.01	0.99	1.04	3.3E-01	5.6E-01

Appendix 5.2 Hazard ratios for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for PSA testing.

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Concentration of Very Large VLDL Particles	6575	0.99	0.96	1.01	3.3E-01	5.6E-01

Abbreviations: PSA, prostate-specific antigen; SD, standard deviation; HR, hazard ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit; LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol; IDL, intermediate-density lipoprotein cholesterol; VLDL, very low-density lipoprotein cholesterol.

1 Hazard ratios (95% confidence interval) for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features, after additional adjustment for PSA testing. Analyses were stratified by region and Townsend Deprivation Index with age as the underlying time variable. Adjusted for PSA testing, fasting status, body mass index, height, alcohol intake, family history of prostate cancer, ethnicity, level of education, smoking status, and weekly physical activity.

2 Padj < 0.05 are statistically significant after allowing for multiple testing using a false discovery rate controlling procedure at $\alpha = 0.05$ (Benjamini-Hochberg). Padj values are sorted in ascending order. Associations with Padj < 0.05 are bolded.

Appendix 5.3 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Metabolites	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj
Concentration of HDL Particles	1309	0.91	0.85	0.96	8.0E-04	3.1E-02	2159	1.00	0.96	1.05	8.4E-01	9.7E-01	3107	1.00	0.97	1.04	8.7E-01	9.9E-01
Leucine	1309	0.91	0.86	0.96	8.1E-04	3.1E-02	2159	0.98	0.94	1.03	4.4E-01	8.1E-01	3107	1.02	0.98	1.05	4.1E-01	7.9E-01
Concentration of Small HDL Particles	1309	0.91	0.86	0.96	6.6E-04	3.1E-02	2159	1.00	0.96	1.05	8.5E-01	9.7E-01	3107	1.00	0.97	1.04	8.1E-01	9.9E-01
Total Lipids in Small HDL	1309	0.90	0.85	0.96	5.6E-04	3.1E-02	2159	0.99	0.95	1.03	6.3E-01	9.2E-01	3107	0.99	0.96	1.03	7.8E-01	9.9E-01
Phospholipids in Small HDL	**1309**	**0.90**	**0.85**	**0.96**	**6.5E-04**	**3.1E-02**	**2159**	**0.99**	**0.94**	**1.03**	**5.2E-01**	**8.6E-01**	**3107**	**0.99**	**0.95**	**1.03**	**6.9E-01**	**9.5E-01**
Cholesterol in Small HDL	1309	0.91	0.86	0.96	7.0E-04	3.1E-02	2159	1.01	0.96	1.06	7.0E-01	9.4E-01	3107	1.01	0.97	1.04	7.9E-01	9.9E-01
Cholesteryl Esters in Small HDL	1309	0.91	0.85	0.96	8.7E-04	3.1E-02	2159	1.01	0.97	1.06	5.4E-01	8.6E-01	3107	1.01	0.97	1.05	6.8E-01	9.5E-01
Total Concentration of Lipoprotein Particles	1309	0.91	0.86	0.96	1.1E-03	3.3E-02	2159	1.01	0.96	1.05	8.2E-01	9.7E-01	3107	1.01	0.97	1.05	7.6E-01	9.8E-01
Albumin	1308	0.91	0.87	0.97	1.2E-03	3.4E-02	2159	1.00	0.96	1.04	9.6E-01	9.9E-01	3107	1.00	0.97	1.04	8.3E-01	9.9E-01
Phospholipids to Total Lipids in Small LDL percentage	1309	1.12	1.04	1.20	1.4E-03	3.4E-02	2159	1.07	1.01	1.13	1.4E-02	2.7E-01	3107	1.02	0.97	1.07	3.9E-01	7.8E-01
Apolipoprotein A1	1309	0.91	0.86	0.97	2.5E-03	4.4E-02	2159	1.00	0.95	1.04	9.0E-01	9.7E-01	3107	1.00	0.96	1.04	9.0E-01	9.9E-01
BCAA	1307	0.92	0.86	0.97	2.4E-03	4.4E-02	2159	0.98	0.94	1.03	4.6E-01	8.2E-01	3104	1.01	0.98	1.05	5.3E-01	8.4E-01
Concentration of Medium HDL Particles	1309	0.93	0.89	0.97	2.4E-03	4.4E-02	2159	1.00	0.96	1.04	9.5E-01	9.9E-01	3107	1.00	0.96	1.04	8.3E-01	9.9E-01
Free Cholesterol in Small HDL	1309	0.91	0.86	0.97	2.0E-03	4.4E-02	2159	0.99	0.95	1.04	7.3E-01	9.4E-01	3107	1.00	0.96	1.03	8.3E-01	9.9E-01
Total Lipids in Medium HDL	1309	0.91	0.86	0.97	2.7E-03	4.5E-02	2159	0.99	0.95	1.04	8.1E-01	9.7E-01	3107	0.99	0.95	1.03	7.1E-01	9.6E-01
Phospholipids in Medium HDL	1309	0.92	0.86	0.97	3.5E-03	5.0E-02	2159	0.99	0.95	1.04	6.8E-01	9.4E-01	3107	0.99	0.95	1.03	5.8E-01	8.7E-01
Cholesterol in Medium HDL	1309	0.93	0.88	0.98	3.7E-03	5.0E-02	2159	1.01	0.96	1.06	7.3E-01	9.4E-01	3107	1.00	0.96	1.04	9.2E-01	9.9E-01
Cholesteryl Esters in Medium HDL	1309	0.93	0.88	0.98	3.6E-03	5.0E-02	2159	1.01	0.96	1.06	6.9E-01	9.4E-01	3107	1.00	0.96	1.04	8.6E-01	9.9E-01
Free Cholesterol in Medium HDL	1309	0.93	0.88	0.98	5.2E-03	6.8E-02	2159	1.00	0.96	1.05	9.1E-01	9.7E-01	3107	1.00	0.96	1.04	8.6E-01	9.9E-01
Isoleucine	1308	0.93	0.88	0.98	6.7E-03	8.3E-02	2159	1.00	0.96	1.05	9.0E-01	9.7E-01	3107	1.02	0.99	1.06	1.9E-01	6.8E-01
Valine	1307	0.92	0.87	0.98	7.6E-03	9.1E-02	2159	0.98	0.94	1.02	3.5E-01	7.5E-01	3104	1.00	0.97	1.04	8.5E-01	9.9E-01

Appendix 5.3 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Metabolites	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj
Phosphatidylcholines	1307	0.93	0.88	0.98	9.5E-03	1.1E-01	2159	0.99	0.95	1.03	6.3E-01	9.2E-01	3104	1.00	0.96	1.03	8.3E-01	9.9E-01
Phosphoglycerides	1307	0.93	0.88	0.98	1.0E-02	1.1E-01	2159	0.99	0.94	1.03	5.7E-01	8.7E-01	3104	0.99	0.96	1.03	7.1E-01	9.6E-01
Total Lipids in HDL	1309	0.93	0.87	0.98	1.3E-02	1.3E-01	2159	1.00	0.95	1.05	9.8E-01	1.0E+00	3107	1.00	0.96	1.04	9.2E-01	9.9E-01
Glutamine	1304	0.93	0.88	0.98	1.2E-02	1.3E-01	2151	1.01	0.96	1.05	7.8E-01	9.5E-01	3094	0.99	0.96	1.03	6.4E-01	9.4E-01
Phospholipids in HDL	1309	0.93	0.88	0.98	1.4E-02	1.3E-01	2159	1.00	0.95	1.04	8.4E-01	9.7E-01	3107	0.99	0.96	1.03	7.6E-01	9.8E-01
Total Cholines	1307	0.93	0.88	0.99	1.4E-02	1.3E-01	2159	0.99	0.95	1.04	7.3E-01	9.4E-01	3104	1.00	0.96	1.04	9.7E-01	1.0E+00
Citrate	1309	0.93	0.88	0.99	1.4E-02	1.3E-01	2159	1.00	0.95	1.04	9.0E-01	9.7E-01	3107	0.98	0.95	1.02	3.4E-01	7.2E-01
Pyruvate	1302	0.94	0.89	0.99	1.6E-02	1.4E-01	2152	1.01	0.97	1.05	7.1E-01	9.4E-01	3095	1.02	0.99	1.06	2.2E-01	6.8E-01
Cholesteryl Esters in HDL	1309	0.93	0.88	0.99	2.1E-02	1.8E-01	2159	1.02	0.97	1.07	4.2E-01	8.0E-01	3107	1.01	0.97	1.05	5.8E-01	8.7E-01
HDL Cholesterol	1309	0.93	0.88	0.99	2.3E-02	1.9E-01	2159	1.02	0.97	1.06	5.0E-01	8.6E-01	3107	1.01	0.97	1.05	6.4E-01	9.4E-01
Total Phospholipids in Lipoprotein Particles	1309	0.94	0.89	0.99	3.0E-02	2.3E-01	2159	0.99	0.95	1.04	7.4E-01	9.4E-01	3107	1.00	0.97	1.04	8.6E-01	9.9E-01
Lactate	1305	0.94	0.89	1.00	3.5E-02	2.6E-01	2156	0.99	0.95	1.04	7.6E-01	9.4E-01	3102	1.01	0.97	1.04	7.0E-01	9.5E-01
Free Cholesterol in HDL	1309	0.94	0.89	1.00	4.2E-02	3.1E-01	2159	1.00	0.96	1.05	8.8E-01	9.7E-01	3107	1.00	0.96	1.04	9.2E-01	9.9E-01
Triglycerides to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	1269	1.07	1.00	1.15	4.4E-02	3.1E-01	2084	1.03	0.99	1.08	1.7E-01	6.0E-01	3014	0.99	0.96	1.02	4.6E-01	8.0E-01
Phospholipids to Total Lipids in IDL percentage	1309	1.06	1.00	1.12	5.3E-02	3.6E-01	2159	0.99	0.95	1.04	7.5E-01	9.4E-01	3107	0.98	0.94	1.01	1.9E-01	6.8E-01
Phospholipids to Total Lipids in Very Small VLDL percentage	1309	1.05	1.00	1.12	6.9E-02	4.7E-01	2159	0.98	0.94	1.02	3.4E-01	7.5E-01	3107	0.97	0.93	1.00	6.8E-02	5.6E-01
Total Esterified Cholesterol	1309	0.95	0.90	1.01	8.1E-02	5.1E-01	2159	1.01	0.97	1.06	5.5E-01	8.6E-01	3107	1.02	0.98	1.06	2.7E-01	6.9E-01
Total Lipids in Lipoprotein Particles	1309	0.95	0.90	1.01	8.1E-02	5.1E-01	2159	0.99	0.95	1.03	6.6E-01	9.2E-01	3107	1.00	0.97	1.04	8.2E-01	9.9E-01
Total Fatty Acids	1307	0.95	0.90	1.01	8.8E-02	5.1E-01	2159	0.97	0.93	1.02	2.6E-01	6.8E-01	3104	0.98	0.95	1.02	4.2E-01	7.9E-01
Omega-6 Fatty Acids	1307	0.95	0.90	1.01	9.0E-02	5.1E-01	2159	0.98	0.94	1.03	4.7E-01	8.3E-01	3104	1.00	0.96	1.04	1.0E+00	1.0E+00
Saturated Fatty Acids	1307	0.95	0.90	1.01	9.0E-02	5.1E-01	2159	0.97	0.93	1.02	2.1E-01	6.6E-01	3104	0.98	0.95	1.02	3.7E-01	7.6E-01
Glucose	1303	0.95	0.90	1.01	9.0E-02	5.1E-01	2157	0.95	0.91	0.99	1.9E-02	2.7E-01	3100	0.96	0.92	0.99	1.8E-02	4.4E-01
Triglycerides in Medium HDL	1309	0.95	0.90	1.01	8.3E-02	5.1E-01	2159	0.95	0.91	1.00	3.6E-02	2.7E-01	3107	0.98	0.94	1.01	2.3E-01	6.8E-01

Appendix 5.3 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Metabolites	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj
Cholesterol to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	1269	0.95	0.90	1.01	9.5E-02	5.3E-01	2084	0.99	0.94	1.03	5.0E-01	8.6E-01	3014	0.99	0.96	1.03	7.6E-01	9.8E-01
Tyrosine	1307	0.95	0.90	1.01	1.0E-01	5.4E-01	2158	0.95	0.91	0.99	2.7E-02	2.7E-01	3103	0.99	0.96	1.03	6.6E-01	9.5E-01
Cholesteryl Esters to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	1269	0.95	0.90	1.01	1.0E-01	5.4E-01	2084	0.98	0.94	1.03	4.5E-01	8.2E-01	3014	1.00	0.97	1.04	9.0E-01	9.9E-01
Total Cholesterol	1309	0.96	0.90	1.01	1.0E-01	5.4E-01	2159	1.01	0.97	1.06	6.2E-01	9.2E-01	3107	1.02	0.98	1.06	2.8E-01	6.9E-01
Monounsaturated Fatty Acids	1307	0.95	0.90	1.01	1.1E-01	5.4E-01	2159	0.96	0.92	1.01	9.2E-02	3.8E-01	3104	0.97	0.94	1.01	1.6E-01	6.8E-01
Triglycerides in HDL	1309	0.96	0.90	1.01	1.2E-01	6.1E-01	2159	0.95	0.91	0.99	2.8E-02	2.7E-01	3107	0.98	0.94	1.01	1.9E-01	6.8E-01
Free Cholesterol to Total Lipids in Medium HDL percentage	1309	0.97	0.93	1.01	1.2E-01	6.1E-01	2159	1.02	0.97	1.07	3.7E-01	7.6E-01	3107	1.01	0.97	1.05	7.6E-01	9.8E-01
Phospholipids to Total Lipids in Medium HDL percentage	1309	1.06	0.98	1.14	1.3E-01	6.1E-01	2159	0.98	0.95	1.01	1.9E-01	6.3E-01	3107	0.98	0.96	1.01	1.5E-01	6.8E-01
Polyunsaturated Fatty Acids	1307	0.96	0.91	1.01	1.4E-01	6.5E-01	2159	1.00	0.96	1.04	9.2E-01	9.8E-01	3104	1.00	0.97	1.04	8.1E-01	9.9E-01
Total Lipids in Large LDL	1309	0.96	0.91	1.02	1.6E-01	7.3E-01	2159	1.02	0.97	1.06	4.8E-01	8.4E-01	3107	1.02	0.98	1.06	2.8E-01	7.0E-01
Total Free Cholesterol	1309	0.96	0.91	1.02	1.9E-01	7.4E-01	2159	1.00	0.96	1.05	8.2E-01	9.7E-01	3107	1.02	0.98	1.06	3.0E-01	7.1E-01
Total Lipids in LDL	1309	0.96	0.91	1.02	1.8E-01	7.4E-01	2159	1.01	0.97	1.05	6.6E-01	9.2E-01	3107	1.02	0.98	1.06	3.3E-01	7.1E-01
Sphingomyelins	1307	0.96	0.91	1.02	1.7E-01	7.4E-01	2159	1.03	0.98	1.07	2.4E-01	6.7E-01	3104	1.02	0.99	1.06	2.1E-01	6.8E-01
Linoleic Acid	1307	0.96	0.91	1.02	1.8E-01	7.4E-01	2159	0.98	0.93	1.02	2.6E-01	6.8E-01	3104	1.00	0.97	1.04	9.5E-01	1.0E+00
3-Hydroxybutyrate	1280	0.96	0.91	1.02	1.9E-01	7.4E-01	2109	1.00	0.96	1.04	1.0E+00	1.0E+00	3038	1.02	0.98	1.05	4.3E-01	7.9E-01
Triglycerides in Chylomicrons and Extremely Large VLDL	1309	1.04	0.98	1.10	1.8E-01	7.4E-01	2159	0.98	0.94	1.02	3.9E-01	7.7E-01	3107	1.00	0.96	1.03	8.5E-01	9.9E-01
Phospholipids in Large LDL	1309	0.96	0.91	1.02	1.7E-01	7.4E-01	2159	1.01	0.97	1.06	5.2E-01	8.6E-01	3107	1.02	0.99	1.06	2.3E-01	6.8E-01
Total Lipids in Medium LDL	1309	0.96	0.91	1.02	1.9E-01	7.4E-01	2159	1.00	0.96	1.04	1.0E+00	1.0E+00	3107	1.01	0.98	1.05	4.4E-01	8.0E-01

Appendix 5.3 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Metabolites	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj
Phospholipids in Medium LDL	1309	0.96	0.91	1.02	1.9E-01	7.4E-01	2159	1.01	0.96	1.05	8.1E-01	9.7E-01	3107	1.02	0.98	1.05	4.0E-01	7.8E-01
Triglycerides in Large HDL	1309	0.96	0.91	1.02	1.9E-01	7.4E-01	2159	0.95	0.91	1.00	3.6E-02	2.7E-01	3107	0.98	0.94	1.01	2.2E-01	6.8E-01
Free Cholesterol to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	1269	0.96	0.91	1.02	1.8E-01	7.4E-01	2084	1.00	0.96	1.04	9.6E-01	9.9E-01	3014	0.98	0.94	1.01	2.2E-01	6.8E-01
Clinical LDL Cholesterol	1309	0.97	0.92	1.02	2.7E-01	8.0E-01	2159	1.02	0.97	1.06	4.8E-01	8.4E-01	3107	1.03	0.99	1.07	1.4E-01	6.8E-01
LDL Cholesterol	1309	0.98	0.94	1.02	2.5E-01	8.0E-01	2159	1.02	0.97	1.07	5.2E-01	8.6E-01	3107	1.02	0.98	1.07	2.5E-01	6.9E-01
Triglycerides in LDL	1309	0.97	0.91	1.02	2.5E-01	8.0E-01	2159	0.96	0.91	1.00	5.7E-02	3.2E-01	3107	0.97	0.94	1.01	1.7E-01	6.8E-01
Phospholipids in LDL	1309	0.97	0.92	1.02	2.3E-01	8.0E-01	2159	1.01	0.97	1.06	6.0E-01	9.1E-01	3107	1.02	0.98	1.06	2.8E-01	6.9E-01
Cholesteryl Esters in LDL	1309	0.98	0.94	1.02	2.5E-01	8.0E-01	2159	1.01	0.96	1.06	6.3E-01	9.2E-01	3107	1.02	0.98	1.06	3.0E-01	7.1E-01
Apolipoprotein B to Apolipoprotein A1 ratio	1309	1.03	0.98	1.09	2.7E-01	8.0E-01	2159	1.00	0.96	1.05	8.8E-01	9.7E-01	3107	1.02	0.98	1.05	3.7E-01	7.6E-01
Polyunsaturated Fatty Acids to Total Fatty Acids percentage	1307	1.04	0.97	1.10	2.7E-01	8.0E-01	2159	1.06	1.01	1.11	2.5E-02	2.7E-01	3104	1.04	1.00	1.09	3.1E-02	4.6E-01
Glycine	1305	0.97	0.92	1.02	2.6E-01	8.0E-01	2147	1.00	0.95	1.04	8.8E-01	9.7E-01	3088	0.98	0.95	1.02	3.3E-01	7.1E-01
Acetate	1303	1.03	0.98	1.09	2.4E-01	8.0E-01	2150	0.99	0.95	1.04	7.7E-01	9.4E-01	3090	1.00	0.96	1.03	8.2E-01	9.9E-01
Acetone	1309	0.97	0.91	1.02	2.5E-01	8.0E-01	2159	1.01	0.97	1.05	6.4E-01	9.2E-01	3107	1.02	0.99	1.06	1.9E-01	6.8E-01
Creatinine	1281	1.03	0.98	1.09	2.3E-01	8.0E-01	2114	1.00	0.96	1.05	8.8E-01	9.7E-01	3051	1.00	0.96	1.04	9.7E-01	1.0E+00
Cholesteryl Esters in IDL	1309	0.97	0.92	1.02	2.6E-01	8.0E-01	2159	1.02	0.98	1.07	3.5E-01	7.5E-01	3107	1.03	0.99	1.07	9.4E-02	6.1E-01
Cholesteryl Esters in Large LDL	1309	0.97	0.93	1.02	2.6E-01	8.0E-01	2159	1.02	0.98	1.07	3.6E-01	7.5E-01	3107	1.02	0.99	1.07	2.2E-01	6.8E-01
Triglycerides in Large LDL	1309	0.96	0.91	1.02	2.2E-01	8.0E-01	2159	0.96	0.92	1.00	7.1E-02	3.4E-01	3107	0.97	0.94	1.01	1.7E-01	6.8E-01
Cholesterol in Medium LDL	1309	0.98	0.93	1.02	2.6E-01	8.0E-01	2159	1.00	0.96	1.05	8.9E-01	9.7E-01	3107	1.02	0.98	1.06	3.3E-01	7.1E-01
Cholesteryl Esters in Medium LDL	1309	0.98	0.94	1.02	2.6E-01	8.0E-01	2159	1.00	0.95	1.04	8.6E-01	9.7E-01	3107	1.02	0.98	1.05	4.3E-01	7.9E-01
Triglycerides in Medium LDL	1309	0.97	0.91	1.03	2.6E-01	8.0E-01	2159	0.96	0.91	1.00	5.1E-02	3.0E-01	3107	0.97	0.94	1.01	1.8E-01	6.8E-01
Triglycerides in Small HDL	1309	0.96	0.91	1.02	2.3E-01	8.0E-01	2159	0.95	0.91	1.00	4.0E-02	2.7E-01	3107	0.98	0.94	1.02	2.4E-01	6.9E-01

Appendix 5.3 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Metabolites	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj
Cholesteryl Esters to Total Lipids in Medium VLDL percentage	1309	1.04	0.98	1.11	2.2E-01	8.0E-01	2159	1.03	0.98	1.08	2.3E-01	6.6E-01	3107	1.04	1.00	1.09	5.0E-02	5.2E-01
Polyunsaturated Fatty Acids to Monounsaturated Fatty Acids ratio	1307	1.03	0.97	1.10	2.8E-01	8.1E-01	2159	1.06	1.01	1.11	1.8E-02	2.7E-01	3104	1.05	1.01	1.09	2.5E-02	4.4E-01
Total Lipids in IDL	1309	0.97	0.92	1.02	2.8E-01	8.1E-01	2159	1.01	0.97	1.06	5.3E-01	8.6E-01	3107	1.02	0.99	1.06	2.1E-01	6.8E-01
Total Cholesterol Minus HDL-C	1309	0.97	0.92	1.03	2.9E-01	8.2E-01	2159	1.01	0.97	1.05	7.5E-01	9.4E-01	3107	1.02	0.98	1.06	2.7E-01	6.9E-01
Cholesterol in IDL	1309	0.97	0.92	1.03	2.9E-01	8.2E-01	2159	1.02	0.98	1.06	3.6E-01	7.5E-01	3107	1.03	0.99	1.07	1.1E-01	6.6E-01
Free Cholesterol in Very Large HDL	1309	1.03	0.97	1.09	2.9E-01	8.2E-01	2159	1.01	0.97	1.06	6.4E-01	9.2E-01	3107	1.01	0.98	1.05	4.8E-01	8.0E-01
Free Cholesterol in Large LDL	1309	0.98	0.94	1.02	3.0E-01	8.3E-01	2159	1.03	0.98	1.08	2.4E-01	6.7E-01	3107	1.03	0.99	1.07	1.4E-01	6.8E-01
Free Cholesterol in LDL	1309	0.98	0.94	1.02	3.2E-01	8.5E-01	2159	1.03	0.98	1.08	2.5E-01	6.8E-01	3107	1.03	0.99	1.08	1.3E-01	6.8E-01
Docosahexaenoic Acid to Total Fatty Acids percentage	1307	1.03	0.97	1.09	3.2E-01	8.5E-01	2159	1.08	1.03	1.13	2.2E-03	2.7E-01	3104	1.05	1.01	1.09	1.3E-02	4.4E-01
Cholesterol in Large LDL	1309	0.98	0.94	1.02	3.2E-01	8.5E-01	2159	1.02	0.98	1.08	3.2E-01	7.5E-01	3107	1.03	0.99	1.07	1.9E-01	6.8E-01
Cholesteryl Esters in Small LDL	1309	0.98	0.94	1.02	3.2E-01	8.5E-01	2159	0.99	0.95	1.04	8.0E-01	9.6E-01	3107	1.02	0.98	1.06	4.6E-01	8.0E-01
Free Cholesterol to Total Lipids in Very Small VLDL percentage	1309	1.03	0.97	1.09	3.2E-01	8.5E-01	2159	1.04	0.99	1.08	1.3E-01	5.0E-01	3107	1.03	0.99	1.07	1.2E-01	6.7E-01
Monounsaturated Fatty Acids to Total Fatty Acids percentage	1307	0.97	0.91	1.03	3.3E-01	8.6E-01	2159	0.95	0.90	0.99	2.2E-02	2.7E-01	3104	0.96	0.92	1.00	3.1E-02	4.6E-01
Glycoprotein Acetyls	1309	1.03	0.97	1.09	3.3E-01	8.6E-01	2159	1.03	0.98	1.08	2.2E-01	6.6E-01	3107	1.01	0.97	1.05	7.5E-01	9.8E-01
Free Cholesterol in Medium LDL	1309	0.98	0.94	1.02	3.5E-01	8.7E-01	2159	1.03	0.98	1.08	2.9E-01	7.1E-01	3107	1.03	0.99	1.08	1.4E-01	6.8E-01
Triglycerides to Total Lipids in Very Large VLDL percentage	1306	1.03	0.97	1.10	3.5E-01	8.7E-01	2150	0.97	0.94	0.99	2.1E-02	2.7E-01	3092	1.01	0.97	1.05	6.1E-01	9.1E-01
Cholesterol to Total Lipids in Medium VLDL percentage	1309	1.03	0.97	1.09	3.4E-01	8.7E-01	2159	1.03	0.99	1.08	1.8E-01	6.2E-01	3107	1.04	1.00	1.08	3.7E-02	4.8E-01

Appendix 5.3 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Metabolites	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj
Free Cholesterol to Total Lipids in Medium VLDL percentage	1309	1.03	0.97	1.09	3.6E-01	8.8E-01	2159	1.02	0.98	1.07	3.2E-01	7.5E-01	3107	1.04	1.00	1.08	6.0E-02	5.2E-01
Free Cholesterol in IDL	1309	0.98	0.92	1.03	3.7E-01	9.1E-01	2159	1.02	0.98	1.06	4.1E-01	7.9E-01	3107	1.02	0.99	1.06	2.0E-01	6.8E-01
Cholesteryl Esters to Total Lipids in IDL percentage	1309	0.98	0.92	1.03	3.9E-01	9.3E-01	2159	1.05	1.00	1.10	3.9E-02	2.7E-01	3107	1.06	1.02	1.10	2.6E-03	4.4E-01
Saturated Fatty Acids to Total Fatty Acids percentage	1307	0.98	0.92	1.03	4.0E-01	9.3E-01	2159	0.97	0.93	1.02	2.5E-01	6.8E-01	3104	0.98	0.95	1.02	4.3E-01	7.9E-01
Cholesteryl Esters in Medium VLDL	1309	1.02	0.97	1.08	4.1E-01	9.3E-01	2159	1.02	0.97	1.06	4.3E-01	8.1E-01	3107	1.04	1.00	1.08	4.8E-02	5.2E-01
Phospholipids in IDL	1309	0.98	0.93	1.03	4.1E-01	9.3E-01	2159	1.01	0.97	1.06	5.4E-01	8.6E-01	3107	1.02	0.98	1.06	2.7E-01	6.9E-01
Cholesterol in Small LDL	1309	0.98	0.94	1.03	4.1E-01	9.3E-01	2159	1.00	0.96	1.05	8.9E-01	9.7E-01	3107	1.02	0.98	1.06	3.2E-01	7.1E-01
Cholesteryl Esters to Total Lipids in Large VLDL percentage	1309	1.02	0.97	1.08	4.0E-01	9.3E-01	2159	1.03	0.98	1.07	2.2E-01	6.6E-01	3107	1.02	0.98	1.06	3.1E-01	7.1E-01
Triglycerides to Total Lipids in Very Large HDL percentage	1309	0.98	0.92	1.03	4.1E-01	9.3E-01	2159	0.97	0.92	1.01	1.3E-01	5.0E-01	3104	0.97	0.93	1.01	1.4E-01	6.8E-01
Total Lipids in Large HDL	1309	0.98	0.92	1.04	4.2E-01	9.5E-01	2159	1.01	0.97	1.06	6.4E-01	9.2E-01	3107	1.01	0.97	1.05	7.5E-01	9.8E-01
Omega-6 Fatty Acids to Total Fatty Acids percentage	1307	1.02	0.96	1.09	4.3E-01	9.5E-01	2159	1.03	0.98	1.08	2.6E-01	6.8E-01	3104	1.03	0.99	1.07	1.1E-01	6.6E-01
Triglycerides in IDL	1309	0.98	0.92	1.04	4.4E-01	9.5E-01	2159	0.96	0.91	1.00	6.0E-02	3.2E-01	3107	0.98	0.94	1.01	2.1E-01	6.8E-01
Cholesterol in Very Large HDL	1309	1.02	0.97	1.08	4.3E-01	9.5E-01	2159	1.02	0.97	1.06	4.6E-01	8.2E-01	3107	1.02	0.98	1.06	2.6E-01	6.9E-01
Phospholipids in Large HDL	1309	0.98	0.92	1.04	4.3E-01	9.5E-01	2159	1.01	0.97	1.06	6.3E-01	9.2E-01	3107	1.00	0.96	1.04	8.9E-01	9.9E-01
Cholesteryl Esters to Total Lipids in Small LDL percentage	1309	0.99	0.96	1.02	4.5E-01	9.5E-01	2159	0.99	0.97	1.02	4.5E-01	8.2E-01	3107	1.00	0.95	1.06	9.0E-01	9.9E-01
Cholesteryl Esters to Total Lipids in Very Large HDL percentage	1309	0.98	0.94	1.03	4.5E-01	9.5E-01	2159	1.03	0.97	1.09	3.6E-01	7.5E-01	3104	1.09	1.01	1.18	2.5E-02	4.4E-01
Free Cholesterol to Total Lipids in Large HDL percentage	1309	1.03	0.96	1.10	4.5E-01	9.5E-01	2159	1.03	0.97	1.08	3.4E-01	7.5E-01	3107	1.01	0.97	1.06	5.0E-01	8.1E-01

Appendix 5.3 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Metabolites	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj
Remnant Cholesterol (Non-HDL, Non-LDL - Cholesterol)	1309	0.98	0.93	1.03	4.7E-01	9.5E-01	2159	1.00	0.96	1.04	9.8E-01	1.0E+00	3107	1.02	0.98	1.06	3.2E-01	7.1E-01
VLDL Cholesterol	1309	0.99	0.94	1.05	8.0E-01	9.5E-01	2159	0.99	0.95	1.03	5.3E-01	8.6E-01	3107	1.01	0.97	1.04	6.7E-01	9.5E-01
Total Triglycerides	1309	0.98	0.93	1.04	5.7E-01	9.5E-01	2159	0.96	0.92	1.01	1.0E-01	4.1E-01	3107	0.99	0.95	1.02	4.8E-01	8.0E-01
Triglycerides in VLDL	1309	0.99	0.93	1.05	7.2E-01	9.5E-01	2159	0.97	0.93	1.01	1.4E-01	5.2E-01	3107	0.99	0.96	1.03	6.7E-01	9.5E-01
Phospholipids in VLDL	1309	0.99	0.94	1.05	7.3E-01	9.5E-01	2159	0.98	0.94	1.02	2.9E-01	7.1E-01	3107	1.00	0.96	1.03	9.0E-01	9.9E-01
Cholesteryl Esters in VLDL	1309	1.00	0.94	1.05	8.7E-01	9.5E-01	2159	0.99	0.95	1.03	6.9E-01	9.4E-01	3107	1.01	0.98	1.05	4.9E-01	8.1E-01
Free Cholesterol in VLDL	1309	0.99	0.94	1.05	7.4E-01	9.5E-01	2159	0.98	0.94	1.02	3.8E-01	7.6E-01	3107	1.00	0.97	1.04	9.1E-01	9.9E-01
Total Lipids in VLDL	1309	0.99	0.93	1.05	6.9E-01	9.5E-01	2159	0.97	0.93	1.02	2.2E-01	6.6E-01	3107	1.00	0.96	1.03	8.1E-01	9.9E-01
Concentration of VLDL Particles	1309	0.99	0.94	1.05	7.7E-01	9.5E-01	2159	0.98	0.94	1.03	4.1E-01	7.9E-01	3107	1.00	0.97	1.04	9.6E-01	1.0E+00
Concentration of LDL Particles	1309	0.99	0.93	1.04	5.9E-01	9.5E-01	2159	1.00	0.96	1.05	8.6E-01	9.7E-01	3107	1.02	0.98	1.05	3.8E-01	7.6E-01
Average Diameter for VLDL Particles	1309	0.98	0.93	1.04	5.7E-01	9.5E-01	2159	0.97	0.93	1.01	1.8E-01	6.2E-01	3107	0.99	0.95	1.03	5.3E-01	8.4E-01
Average Diameter for LDL Particles	1309	1.00	0.95	1.06	8.6E-01	9.5E-01	2159	1.06	1.01	1.11	1.3E-02	2.7E-01	3107	1.02	0.98	1.05	4.1E-01	7.9E-01
Average Diameter for HDL Particles	1309	0.99	0.94	1.05	8.5E-01	9.5E-01	2159	1.00	0.96	1.05	9.5E-01	9.9E-01	3107	1.00	0.96	1.04	9.9E-01	1.0E+00
Triglycerides to Phosphoglycerides ratio	1307	1.02	0.96	1.08	5.7E-01	9.5E-01	2159	0.96	0.92	1.01	1.0E-01	4.2E-01	3104	0.99	0.95	1.02	4.8E-01	8.0E-01
Apolipoprotein B	1309	0.99	0.93	1.04	5.9E-01	9.5E-01	2159	1.00	0.96	1.05	9.3E-01	9.9E-01	3107	1.02	0.98	1.05	3.7E-01	7.6E-01
Degree of Unsaturation	1307	1.02	0.96	1.08	4.9E-01	9.5E-01	2159	1.07	1.02	1.12	3.3E-03	2.7E-01	3104	1.04	1.01	1.08	2.5E-02	4.4E-01
Omega-3 Fatty Acids	1307	0.99	0.93	1.05	7.2E-01	9.5E-01	2159	1.04	1.00	1.09	6.1E-02	3.2E-01	3104	1.02	0.98	1.06	2.7E-01	6.9E-01
Omega-3 Fatty Acids to Total Fatty Acids percentage	1307	1.02	0.96	1.08	5.7E-01	9.5E-01	2159	1.07	1.02	1.12	4.1E-03	2.7E-01	3104	1.03	1.00	1.07	7.3E-02	5.6E-01
Linoleic Acid to Total Fatty Acids percentage	1307	1.02	0.96	1.08	6.2E-01	9.5E-01	2159	1.00	0.95	1.05	9.3E-01	9.9E-01	3104	1.03	0.99	1.07	1.5E-01	6.8E-01
Omega-6 Fatty Acids to Omega-3 Fatty Acids ratio	1306	0.98	0.93	1.04	5.2E-01	9.5E-01	2159	0.95	0.91	0.99	1.8E-02	2.7E-01	3104	0.98	0.94	1.01	2.3E-01	6.8E-01
Alanine	1307	0.99	0.93	1.04	6.5E-01	9.5E-01	2159	0.99	0.94	1.03	5.3E-01	8.6E-01	3105	0.95	0.92	0.99	1.1E-02	4.4E-01
Histidine	1307	0.98	0.93	1.04	4.8E-01	9.5E-01	2158	1.00	0.96	1.04	9.9E-01	1.0E+00	3102	1.04	1.00	1.08	3.6E-02	4.8E-01

Appendix 5.3 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Metabolites	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj
Phenylalanine	1308	1.00	0.94	1.05	9.0E-01	9.5E-01	2158	0.98	0.93	1.02	2.6E-01	6.8E-01	3107	1.01	0.97	1.04	6.9E-01	9.5E-01
Acetoacetate	1294	0.99	0.94	1.05	7.7E-01	9.5E-01	2146	0.99	0.95	1.04	7.7E-01	9.4E-01	3077	1.03	0.99	1.07	1.1E-01	6.5E-01
Concentration of Chylomicrons and Extremely Large VLDL Particles	1309	1.01	0.95	1.06	8.3E-01	9.5E-01	2159	0.97	0.92	1.01	1.1E-01	4.2E-01	3107	0.99	0.95	1.02	4.3E-01	7.9E-01
Total Lipids in Chylomicrons and Extremely Large VLDL	1309	1.00	0.95	1.06	8.6E-01	9.5E-01	2159	0.96	0.92	1.00	6.2E-02	3.2E-01	3107	0.99	0.96	1.03	6.7E-01	9.5E-01
Phospholipids in Chylomicrons and Extremely Large VLDL	1309	1.01	0.96	1.07	6.5E-01	9.5E-01	2159	0.96	0.93	1.01	8.8E-02	3.8E-01	3107	0.98	0.95	1.02	3.0E-01	7.1E-01
Cholesterol in Chylomicrons and Extremely Large VLDL	1309	0.98	0.93	1.04	5.8E-01	9.5E-01	2159	0.95	0.91	0.99	1.5E-02	2.7E-01	3107	0.99	0.95	1.02	4.3E-01	7.9E-01
Cholesteryl Esters in Chylomicrons and Extremely Large VLDL	1309	0.98	0.93	1.04	5.0E-01	9.5E-01	2159	0.95	0.91	0.99	1.3E-02	2.7E-01	3107	0.99	0.95	1.03	5.8E-01	8.7E-01
Free Cholesterol in Chylomicrons and Extremely Large VLDL	1309	0.99	0.94	1.05	7.6E-01	9.5E-01	2159	0.96	0.92	1.00	4.6E-02	2.8E-01	3107	0.98	0.95	1.02	3.2E-01	7.1E-01
Concentration of Very Large VLDL Particles	1309	0.99	0.94	1.05	7.3E-01	9.5E-01	2159	0.97	0.93	1.02	2.2E-01	6.6E-01	3107	1.00	0.96	1.03	8.3E-01	9.9E-01
Total Lipids in Very Large VLDL	1309	0.99	0.94	1.05	8.5E-01	9.5E-01	2159	0.97	0.93	1.01	1.9E-01	6.3E-01	3107	1.00	0.96	1.03	8.8E-01	9.9E-01
Phospholipids in Very Large VLDL	1309	1.01	0.95	1.06	8.4E-01	9.5E-01	2159	0.98	0.94	1.02	2.2E-01	6.6E-01	3107	1.00	0.96	1.03	8.0E-01	9.9E-01
Cholesterol in Very Large VLDL	1309	0.99	0.94	1.05	8.5E-01	9.5E-01	2159	0.98	0.94	1.02	2.8E-01	7.1E-01	3107	1.00	0.97	1.04	9.4E-01	1.0E+00
Cholesteryl Esters in Very Large VLDL	1309	1.00	0.94	1.05	8.8E-01	9.5E-01	2159	0.98	0.94	1.02	4.0E-01	7.9E-01	3107	1.01	0.97	1.04	6.9E-01	9.5E-01
Triglycerides in Very Large VLDL	1309	1.01	0.95	1.07	7.7E-01	9.5E-01	2159	0.96	0.92	1.01	9.1E-02	3.8E-01	3107	1.00	0.97	1.04	9.3E-01	9.9E-01
Concentration of Large VLDL Particles	1309	0.99	0.94	1.05	7.4E-01	9.5E-01	2159	0.97	0.93	1.02	2.5E-01	6.8E-01	3107	1.00	0.96	1.03	8.7E-01	9.9E-01
Total Lipids in Large VLDL	1309	0.99	0.94	1.05	7.6E-01	9.5E-01	2159	0.98	0.94	1.02	3.2E-01	7.5E-01	3107	1.00	0.96	1.04	1.0E+00	1.0E+00

Appendix 5.3 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Metabolites	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj
Phospholipids in Large VLDL	1309	1.00	0.95	1.06	8.9E-01	9.5E-01	2159	0.99	0.95	1.03	6.6E-01	9.2E-01	3107	1.00	0.97	1.04	1.0E+00	1.0E+00
Cholesterol in Large VLDL	1309	0.99	0.94	1.05	8.3E-01	9.5E-01	2159	0.98	0.94	1.03	4.4E-01	8.1E-01	3107	1.00	0.97	1.04	8.8E-01	9.9E-01
Free Cholesterol in Large VLDL	1309	0.99	0.94	1.05	6.9E-01	9.5E-01	2159	0.98	0.94	1.02	3.4E-01	7.5E-01	3107	1.00	0.96	1.04	9.6E-01	1.0E+00
Triglycerides in Large VLDL	1309	0.99	0.94	1.05	7.5E-01	9.5E-01	2159	0.98	0.94	1.02	3.6E-01	7.5E-01	3107	1.00	0.96	1.04	9.8E-01	1.0E+00
Concentration of Medium VLDL Particles	1309	0.99	0.94	1.05	7.6E-01	9.5E-01	2159	0.99	0.95	1.04	7.4E-01	9.4E-01	3107	1.01	0.98	1.05	4.8E-01	8.0E-01
Total Lipids in Medium VLDL	1309	0.99	0.94	1.05	7.1E-01	9.5E-01	2159	0.99	0.95	1.03	6.7E-01	9.2E-01	3107	1.01	0.97	1.05	5.6E-01	8.7E-01
Phospholipids in Medium VLDL	1309	1.00	0.94	1.05	8.9E-01	9.5E-01	2159	1.00	0.95	1.04	8.2E-01	9.7E-01	3107	1.02	0.98	1.05	3.6E-01	7.5E-01
Cholesterol in Medium VLDL	1309	1.01	0.95	1.06	7.9E-01	9.5E-01	2159	1.01	0.97	1.05	6.5E-01	9.2E-01	3107	1.03	1.00	1.07	8.5E-02	6.1E-01
Triglycerides in Medium VLDL	1309	0.99	0.93	1.04	6.1E-01	9.5E-01	2159	0.98	0.94	1.02	3.6E-01	7.5E-01	3107	1.00	0.96	1.04	1.0E+00	1.0E+00
Concentration of Small VLDL Particles	1309	1.00	0.94	1.05	9.0E-01	9.5E-01	2159	0.98	0.94	1.02	3.6E-01	7.5E-01	3107	1.00	0.96	1.04	9.6E-01	1.0E+00
Total Lipids in Small VLDL	1309	0.99	0.94	1.05	7.8E-01	9.5E-01	2159	0.98	0.94	1.02	3.9E-01	7.7E-01	3107	1.00	0.96	1.04	9.9E-01	1.0E+00
Phospholipids in Small VLDL	1309	0.99	0.94	1.05	8.4E-01	9.5E-01	2159	0.99	0.95	1.04	7.3E-01	9.4E-01	3107	1.01	0.97	1.05	6.0E-01	9.0E-01
Free Cholesterol in Small VLDL	1309	1.00	0.94	1.05	8.9E-01	9.5E-01	2159	1.00	0.96	1.04	9.7E-01	9.9E-01	3107	1.02	0.98	1.06	2.8E-01	7.0E-01
Triglycerides in Small VLDL	1309	0.99	0.94	1.05	8.2E-01	9.5E-01	2159	0.97	0.92	1.01	1.4E-01	5.0E-01	3107	0.99	0.95	1.03	5.4E-01	8.5E-01
Concentration of Very Small VLDL Particles	1309	0.99	0.94	1.05	8.1E-01	9.5E-01	2159	0.99	0.95	1.03	5.5E-01	8.6E-01	3107	1.00	0.97	1.04	8.5E-01	9.9E-01
Total Lipids in Very Small VLDL	1309	0.99	0.94	1.05	8.3E-01	9.5E-01	2159	0.99	0.95	1.03	5.3E-01	8.6E-01	3107	1.00	0.97	1.04	9.3E-01	9.9E-01
Cholesterol in Very Small VLDL	1309	1.00	0.94	1.05	8.9E-01	9.5E-01	2159	1.00	0.96	1.05	8.7E-01	9.7E-01	3107	1.02	0.98	1.05	3.7E-01	7.6E-01
Cholesteryl Esters in Very Small VLDL	1309	1.00	0.94	1.05	8.7E-01	9.5E-01	2159	1.01	0.97	1.05	6.9E-01	9.4E-01	3107	1.02	0.98	1.06	2.6E-01	6.9E-01
Triglycerides in Very Small VLDL	1309	0.99	0.93	1.05	7.0E-01	9.5E-01	2159	0.95	0.91	1.00	4.7E-02	2.8E-01	3107	0.98	0.94	1.02	2.5E-01	6.9E-01

Appendix 5.3 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Metabolites	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj
Concentration of IDL Particles	1309	0.98	0.93	1.04	5.7E-01	9.5E-01	2159	1.01	0.97	1.05	7.0E-01	9.4E-01	3107	1.03	0.99	1.06	1.7E-01	6.8E-01
Concentration of Large LDL Particles	1309	0.99	0.94	1.04	6.5E-01	9.5E-01	2159	1.01	0.97	1.06	5.8E-01	8.8E-01	3107	1.02	0.98	1.06	3.3E-01	7.1E-01
Concentration of Medium LDL Particles	1309	0.98	0.93	1.03	4.6E-01	9.5E-01	2159	0.99	0.95	1.03	6.5E-01	9.2E-01	3107	1.01	0.98	1.05	4.8E-01	8.0E-01
Concentration of Small LDL Particles	1309	0.99	0.93	1.04	6.3E-01	9.5E-01	2159	0.99	0.95	1.04	7.3E-01	9.4E-01	3107	1.01	0.98	1.05	4.6E-01	8.0E-01
Total Lipids in Small LDL	1309	0.98	0.93	1.03	4.7E-01	9.5E-01	2159	1.00	0.96	1.05	9.5E-01	9.9E-01	3107	1.02	0.98	1.05	4.1E-01	7.9E-01
Free Cholesterol in Small LDL	1309	1.00	0.94	1.05	8.6E-01	9.5E-01	2159	1.03	0.98	1.08	2.8E-01	7.1E-01	3107	1.04	0.99	1.08	9.2E-02	6.1E-01
Triglycerides in Small LDL	1309	0.98	0.92	1.04	5.0E-01	9.5E-01	2159	0.95	0.91	1.00	4.2E-02	2.7E-01	3107	0.98	0.94	1.02	2.5E-01	6.9E-01
Concentration of Very Large HDL Particles	1309	1.01	0.96	1.07	6.2E-01	9.5E-01	2159	1.00	0.96	1.05	9.3E-01	9.9E-01	3107	1.01	0.97	1.05	6.6E-01	9.5E-01
Total Lipids in Very Large HDL	1309	1.02	0.97	1.08	4.6E-01	9.5E-01	2159	1.00	0.96	1.05	8.4E-01	9.7E-01	3107	1.01	0.98	1.05	4.5E-01	8.0E-01
Phospholipids in Very Large HDL	1309	1.01	0.95	1.07	7.1E-01	9.5E-01	2159	1.02	0.98	1.07	3.4E-01	7.5E-01	3107	1.01	0.98	1.05	4.7E-01	8.0E-01
Cholesteryl Esters in Very Large HDL	1309	1.01	0.95	1.07	7.4E-01	9.5E-01	2159	1.02	0.98	1.07	3.6E-01	7.5E-01	3107	1.03	0.99	1.07	2.0E-01	6.8E-01
Triglycerides in Very Large HDL	1309	0.99	0.93	1.05	6.7E-01	9.5E-01	2159	0.95	0.91	1.00	4.1E-02	2.7E-01	3107	0.98	0.94	1.02	2.9E-01	7.1E-01
Concentration of Large HDL Particles	1309	0.98	0.93	1.04	5.4E-01	9.5E-01	2159	1.01	0.97	1.06	5.4E-01	8.6E-01	3107	1.01	0.97	1.05	7.5E-01	9.8E-01
Cholesterol in Large HDL	1309	0.99	0.94	1.05	8.0E-01	9.5E-01	2159	1.03	0.98	1.08	2.3E-01	6.6E-01	3107	1.02	0.98	1.06	3.2E-01	7.1E-01
Cholesteryl Esters in Large HDL	1309	0.99	0.94	1.05	8.2E-01	9.5E-01	2159	1.03	0.99	1.08	1.7E-01	5.9E-01	3107	1.02	0.98	1.06	2.5E-01	6.9E-01
Free Cholesterol in Large HDL	1309	1.00	0.94	1.05	8.9E-01	9.5E-01	2159	1.02	0.97	1.07	4.2E-01	8.0E-01	3107	1.01	0.97	1.05	5.8E-01	8.7E-01
Phospholipids to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	1269	1.02	0.96	1.07	5.8E-01	9.5E-01	2084	0.99	0.95	1.03	7.2E-01	9.4E-01	3014	0.97	0.94	1.00	4.1E-02	4.8E-01
Phospholipids to Total Lipids in Very Large VLDL percentage	1306	0.99	0.94	1.03	5.9E-01	9.5E-01	2150	0.98	0.94	1.01	1.5E-01	5.4E-01	3092	1.00	0.96	1.03	9.1E-01	9.9E-01

Appendix 5.3 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Metabolites	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj
Cholesterol to Total Lipids in Very Large VLDL percentage	1306	1.01	0.95	1.07	7.9E-01	9.5E-01	2150	1.04	0.99	1.09	9.0E-02	3.8E-01	3092	1.01	0.98	1.05	5.1E-01	8.1E-01
Cholesteryl Esters to Total Lipids in Very Large VLDL percentage	1306	1.01	0.95	1.07	7.4E-01	9.5E-01	2150	1.04	1.00	1.09	6.8E-02	3.4E-01	3092	1.02	0.98	1.06	3.0E-01	7.1E-01
Free Cholesterol to Total Lipids in Very Large VLDL percentage	1306	1.01	0.95	1.06	8.0E-01	9.5E-01	2150	1.02	0.98	1.06	3.9E-01	7.7E-01	3092	0.99	0.96	1.03	6.7E-01	9.5E-01
Phospholipids to Total Lipids in Large VLDL percentage	1309	1.02	0.96	1.08	5.7E-01	9.5E-01	2159	1.01	0.97	1.05	7.6E-01	9.4E-01	3107	1.00	0.97	1.03	9.6E-01	1.0E+00
Cholesterol to Total Lipids in Large VLDL percentage	1309	1.01	0.96	1.07	7.1E-01	9.5E-01	2159	1.02	0.98	1.07	3.5E-01	7.5E-01	3107	1.01	0.97	1.05	5.6E-01	8.7E-01
Free Cholesterol to Total Lipids in Large VLDL percentage	1309	0.98	0.94	1.04	5.5E-01	9.5E-01	2159	1.00	0.96	1.04	9.1E-01	9.7E-01	3107	0.99	0.96	1.03	7.6E-01	9.8E-01
Phospholipids to Total Lipids in Medium VLDL percentage	1309	1.03	0.96	1.09	4.5E-01	9.5E-01	2159	1.02	0.97	1.07	5.4E-01	8.6E-01	3107	1.04	0.99	1.08	1.1E-01	6.5E-01
Triglycerides to Total Lipids in Medium VLDL percentage	1309	0.99	0.94	1.04	6.9E-01	9.5E-01	2159	0.97	0.94	1.01	2.1E-01	6.6E-01	3107	0.98	0.94	1.01	1.9E-01	6.8E-01
Phospholipids to Total Lipids in Small VLDL percentage	1309	1.01	0.95	1.07	8.2E-01	9.5E-01	2159	1.04	1.00	1.09	7.5E-02	3.4E-01	3107	1.04	1.00	1.08	5.7E-02	5.2E-01
Cholesterol to Total Lipids in Small VLDL percentage	1309	1.01	0.95	1.07	7.3E-01	9.5E-01	2159	1.03	0.98	1.07	2.3E-01	6.6E-01	3107	1.03	1.00	1.07	7.7E-02	5.6E-01
Cholesteryl Esters to Total Lipids in Small VLDL percentage	1309	1.01	0.96	1.07	6.9E-01	9.5E-01	2159	1.02	0.97	1.06	4.5E-01	8.2E-01	3107	1.02	0.99	1.06	2.0E-01	6.8E-01
Free Cholesterol to Total Lipids in Small VLDL percentage	1309	1.01	0.95	1.07	8.1E-01	9.5E-01	2159	1.04	0.99	1.09	9.1E-02	3.8E-01	3107	1.04	1.01	1.09	2.4E-02	4.4E-01
Cholesterol to Total Lipids in Very Small VLDL percentage	1309	1.00	0.95	1.06	8.8E-01	9.5E-01	2159	1.05	1.00	1.10	4.4E-02	2.8E-01	3107	1.05	1.01	1.09	2.1E-02	4.4E-01

Appendix 5.3 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Metabolites	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj
Triglycerides to Total Lipids in Very Small VLDL percentage	1309	0.99	0.94	1.05	7.8E-01	9.5E-01	2159	0.95	0.91	1.00	4.1E-02	2.7E-01	3107	0.97	0.93	1.00	7.7E-02	5.6E-01
Cholesterol to Total Lipids in IDL percentage	1309	0.98	0.93	1.04	5.7E-01	9.5E-01	2159	1.05	1.00	1.10	3.9E-02	2.7E-01	3107	1.05	1.01	1.10	8.2E-03	4.4E-01
Free Cholesterol to Total Lipids in IDL percentage	1309	1.01	0.95	1.07	7.9E-01	9.5E-01	2159	1.02	0.98	1.07	2.9E-01	7.2E-01	3107	1.01	0.98	1.05	5.3E-01	8.4E-01
Triglycerides to Total Lipids in IDL percentage	1309	1.01	0.95	1.07	7.4E-01	9.5E-01	2159	0.95	0.90	0.99	1.9E-02	2.7E-01	3107	0.95	0.92	0.99	1.8E-02	4.4E-01
Phospholipids to Total Lipids in Large LDL percentage	1309	1.02	0.96	1.07	5.6E-01	9.5E-01	2159	0.99	0.95	1.03	5.5E-01	8.6E-01	3107	1.01	0.98	1.05	5.1E-01	8.1E-01
Cholesterol to Total Lipids in Large LDL percentage	1309	1.02	0.90	1.15	8.0E-01	9.5E-01	2159	1.35	1.04	1.76	2.2E-02	2.7E-01	3107	1.24	1.00	1.53	5.2E-02	5.2E-01
Cholesteryl Esters to Total Lipids in Large LDL percentage	1309	1.03	0.86	1.24	7.7E-01	9.5E-01	2159	1.19	0.97	1.46	9.2E-02	3.8E-01	3107	1.10	0.93	1.30	2.5E-01	6.9E-01
Triglycerides to Total Lipids in Large LDL percentage	1309	1.01	0.95	1.07	7.8E-01	9.5E-01	2159	0.94	0.90	0.99	1.1E-02	2.7E-01	3107	0.95	0.91	0.99	1.2E-02	4.4E-01
Phospholipids to Total Lipids in Medium LDL percentage	1309	1.01	0.95	1.06	8.1E-01	9.5E-01	2159	1.05	1.01	1.10	2.7E-02	2.7E-01	3107	1.01	0.97	1.04	6.7E-01	9.5E-01
Cholesterol to Total Lipids in Medium LDL percentage	1309	1.01	0.88	1.16	8.8E-01	9.5E-01	2159	1.11	0.89	1.37	3.6E-01	7.5E-01	3107	1.25	1.03	1.51	2.6E-02	4.4E-01
Cholesteryl Esters to Total Lipids in Medium LDL percentage	1309	1.01	0.92	1.10	8.8E-01	9.5E-01	2159	0.99	0.97	1.02	5.6E-01	8.7E-01	3107	1.02	0.93	1.11	6.9E-01	9.5E-01
Free Cholesterol to Total Lipids in Medium LDL percentage	1309	1.01	0.94	1.09	7.9E-01	9.5E-01	2159	1.12	1.03	1.22	9.0E-03	2.7E-01	3107	1.06	0.99	1.14	7.0E-02	5.6E-01
Triglycerides to Total Lipids in Medium LDL percentage	1309	1.01	0.95	1.07	8.5E-01	9.5E-01	2159	0.95	0.91	0.99	2.5E-02	2.7E-01	3107	0.95	0.92	0.99	1.5E-02	4.4E-01
Cholesterol to Total Lipids in Small LDL percentage	1309	0.99	0.96	1.03	6.4E-01	9.5E-01	2159	1.01	0.90	1.14	8.3E-01	9.7E-01	3107	1.12	0.97	1.30	1.3E-01	6.8E-01

Appendix 5.3 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Metabolites	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj
Free Cholesterol to Total Lipids in Small LDL percentage	1309	1.02	0.95	1.10	5.9E-01	9.5E-01	2159	1.07	0.99	1.15	7.2E-02	3.4E-01	3107	1.06	1.00	1.12	6.1E-02	5.2E-01
Triglycerides to Total Lipids in Small LDL percentage	1309	1.00	0.94	1.06	8.7E-01	9.5E-01	2159	0.94	0.90	0.99	1.5E-02	2.7E-01	3107	0.96	0.93	1.00	4.2E-02	4.8E-01
Phospholipids to Total Lipids in Very Large HDL percentage	1309	0.99	0.94	1.05	7.8E-01	9.5E-01	2159	1.03	0.98	1.08	3.3E-01	7.5E-01	3104	1.02	0.98	1.06	4.1E-01	7.9E-01
Cholesterol to Total Lipids in Very Large HDL percentage	1309	0.99	0.95	1.04	7.9E-01	9.5E-01	2159	1.04	0.99	1.09	1.4E-01	5.2E-01	3104	1.01	0.97	1.05	5.0E-01	8.1E-01
Phospholipids to Total Lipids in Large HDL percentage	1309	1.01	0.93	1.10	7.8E-01	9.5E-01	2159	1.00	0.95	1.05	9.6E-01	9.9E-01	3107	0.99	0.96	1.02	4.5E-01	8.0E-01
Cholesterol to Total Lipids in Large HDL percentage	1309	1.02	0.96	1.10	5.0E-01	9.5E-01	2159	1.07	1.00	1.14	5.2E-02	3.0E-01	3107	1.05	0.99	1.10	8.8E-02	6.1E-01
Cholesteryl Esters to Total Lipids in Large HDL percentage	1309	1.02	0.96	1.09	5.2E-01	9.5E-01	2159	1.08	1.01	1.15	2.8E-02	2.7E-01	3107	1.05	1.00	1.11	5.9E-02	5.2E-01
Triglycerides to Total Lipids in Large HDL percentage	1309	0.99	0.93	1.05	7.2E-01	9.5E-01	2159	0.95	0.91	1.00	3.4E-02	2.7E-01	3107	0.98	0.94	1.01	2.0E-01	6.8E-01
Cholesterol to Total Lipids in Medium HDL percentage	1309	0.98	0.93	1.03	4.7E-01	9.5E-01	2159	1.05	1.00	1.11	6.0E-02	3.2E-01	3107	1.03	0.99	1.08	1.4E-01	6.8E-01
Cholesteryl Esters to Total Lipids in Medium HDL percentage	1309	0.99	0.93	1.04	6.5E-01	9.5E-01	2159	1.05	1.00	1.11	6.3E-02	3.2E-01	3107	1.04	0.99	1.08	1.2E-01	6.8E-01
Phospholipids to Total Lipids in Small HDL percentage	1309	0.99	0.94	1.05	7.9E-01	9.5E-01	2159	0.97	0.93	1.02	2.6E-01	6.8E-01	3107	0.99	0.95	1.02	4.5E-01	8.0E-01
Cholesterol to Total Lipids in Small HDL percentage	1309	1.00	0.94	1.05	8.7E-01	9.5E-01	2159	1.06	1.01	1.11	2.1E-02	2.7E-01	3107	1.03	0.99	1.07	1.5E-01	6.8E-01
Cholesteryl Esters to Total Lipids in Small HDL percentage	1309	0.99	0.93	1.06	7.6E-01	9.5E-01	2159	1.06	1.01	1.12	2.7E-02	2.7E-01	3107	1.03	0.99	1.08	1.7E-01	6.8E-01

Appendix 5.3 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Metabolites	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj
Free Cholesterol to Total Lipids in Small HDL percentage	1309	1.01	0.96	1.07	6.6E-01	9.5E-01	2159	1.01	0.96	1.05	7.6E-01	9.4E-01	3107	1.00	0.97	1.04	8.8E-01	9.9E-01
Triglycerides to Total Lipids in Small HDL percentage	1309	1.01	0.95	1.07	7.1E-01	9.5E-01	2159	0.95	0.91	1.00	3.3E-02	2.7E-01	3107	0.98	0.94	1.01	2.2E-01	6.8E-01
Cholesterol in Small VLDL	1309	1.00	0.94	1.05	9.0E-01	9.6E-01	2159	0.99	0.95	1.04	7.6E-01	9.4E-01	3107	1.01	0.98	1.05	4.8E-01	8.0E-01
Free Cholesterol in Very Large VLDL	1309	1.00	0.94	1.05	9.3E-01	9.7E-01	2159	0.97	0.93	1.02	2.4E-01	6.7E-01	3107	1.00	0.96	1.03	8.7E-01	9.9E-01
Cholesteryl Esters in Small VLDL	1309	1.00	0.94	1.05	9.2E-01	9.7E-01	2159	0.99	0.95	1.03	6.2E-01	9.2E-01	3107	1.01	0.97	1.05	6.3E-01	9.4E-01
Triglycerides to Total Lipids in Large VLDL percentage	1309	1.00	0.95	1.05	9.3E-01	9.7E-01	2159	1.00	0.96	1.05	8.5E-01	9.7E-01	3107	1.00	0.96	1.04	9.2E-01	9.9E-01
Free Cholesterol to Total Lipids in Large LDL percentage	1309	1.00	0.94	1.07	9.2E-01	9.7E-01	2159	1.09	0.99	1.21	7.2E-02	3.4E-01	3107	1.07	0.99	1.16	1.0E-01	6.5E-01
Docosahexaenoic Acid	1307	1.00	0.94	1.06	9.8E-01	9.9E-01	2159	1.06	1.01	1.11	1.3E-02	2.7E-01	3104	1.04	1.00	1.08	3.8E-02	4.8E-01
Cholesteryl Esters in Large VLDL	1309	1.00	0.95	1.06	9.8E-01	9.9E-01	2159	0.99	0.95	1.03	5.8E-01	8.8E-01	3107	1.01	0.97	1.05	6.8E-01	9.5E-01
Free Cholesterol in Medium VLDL	1309	1.00	0.95	1.05	9.9E-01	9.9E-01	2159	1.00	0.96	1.04	9.7E-01	9.9E-01	3107	1.02	0.99	1.06	2.4E-01	6.9E-01
Phospholipids in Very Small VLDL	1309	1.00	0.95	1.06	9.9E-01	9.9E-01	2159	0.98	0.94	1.03	4.6E-01	8.2E-01	3107	1.00	0.96	1.03	8.9E-01	9.9E-01
Free Cholesterol in Very Small VLDL	1309	1.00	0.95	1.05	9.7E-01	9.9E-01	2159	0.99	0.95	1.04	7.3E-01	9.4E-01	3107	1.01	0.97	1.04	7.3E-01	9.7E-01
Triglycerides to Total Lipids in Small VLDL percentage	1309	1.00	0.94	1.06	9.8E-01	9.9E-01	2159	0.96	0.92	1.01	1.1E-01	4.3E-01	3107	0.98	0.94	1.01	2.0E-01	6.8E-01
Cholesteryl Esters to Total Lipids in Very Small VLDL percentage	1309	1.00	0.94	1.06	9.7E-01	9.9E-01	2159	1.05	1.00	1.10	4.1E-02	2.7E-01	3107	1.05	1.01	1.09	1.9E-02	4.4E-01
Free Cholesterol to Total Lipids in Very Large HDL percentage	1309	1.00	0.94	1.06	9.7E-01	9.9E-01	2159	1.00	0.96	1.05	8.8E-01	9.7E-01	3104	0.98	0.95	1.02	3.3E-01	7.1E-01
Triglycerides to Total Lipids in Medium HDL percentage	1309	1.00	0.94	1.06	9.9E-01	9.9E-01	2159	0.95	0.91	1.00	4.0E-02	2.7E-01	3107	0.98	0.94	1.02	2.9E-01	7.1E-01

Appendix 5.3 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Metabolites	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj
Phospholipids in Small LDL	1309	1.00	0.95	1.06	9.9E-01	9.9E-01	2159	1.01	0.97	1.06	5.1E-01	8.6E-01	3107	1.02	0.98	1.06	2.8E-01	6.9E-01

Abbreviations: HR, hazard ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit; LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol; IDL, intermediate-density lipoprotein cholesterol; VLDL, very low-density lipoprotein cholesterol; BCAA, Total concentration of branched-chain amino acids (Leucine + Isoleucine + Valine).

1 Hazard ratios (95% confidence interval) for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features by time since blood collection. Analyses were stratified by region and Townsend Deprivation Index with age as the underlying time variable. Adjusted for fasting status, body mass index, height, alcohol intake, family history of prostate cancer, ethnicity, level of education, smoking status, and weekly physical activity.

2 Padj < 0.05 are statistically significant after allowing for multiple testing using a false discovery rate controlling procedure at α = 0.05 (Benjamini-Hochberg). Padj values are sorted in ascending order in the < 3 years group. Associations with a Padj < 0.05 in both the overall dataset and in this dataset are bolded.

Appendix 5.4 Hazard ratios for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes.

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Docosahexaenoic Acid to Total Fatty Acids percentage	**6570**	**1.06**	**1.03**	**1.08**	**4.9E-05**	**1.2E-02**
Phospholipids to Total Lipids in Small LDL percentage	**6575**	**1.06**	**1.03**	**1.10**	**1.4E-04**	**1.7E-02**
Degree of Unsaturation	6570	1.04	1.02	1.07	8.5E-04	5.9E-02
Omega-3 Fatty Acids to Total Fatty Acids percentage	6570	1.04	1.02	1.07	9.5E-04	5.9E-02
Polyunsaturated Fatty Acids to Total Fatty Acids percentage	6570	1.04	1.02	1.07	2.2E-03	9.1E-02
Polyunsaturated Fatty Acids to Monounsaturated Fatty Acids ratio	6570	1.04	1.02	1.07	2.0E-03	9.1E-02
Triglycerides in LDL	6575	0.96	0.94	0.99	5.9E-03	1.0E-01
Triglycerides in HDL	6575	0.96	0.94	0.99	4.3E-03	1.0E-01
Monounsaturated Fatty Acids	6570	0.96	0.94	0.99	4.3E-03	1.0E-01
Docosahexaenoic Acid	6570	1.04	1.01	1.06	5.8E-03	1.0E-01
Monounsaturated Fatty Acids to Total Fatty Acids percentage	6570	0.96	0.94	0.99	4.0E-03	1.0E-01
Omega-6 Fatty Acids to Omega-3 Fatty Acids ratio	6569	0.97	0.94	0.99	6.7E-03	1.0E-01
Triglycerides in Large LDL	6575	0.96	0.94	0.99	5.5E-03	1.0E-01
Triglycerides in Medium LDL	6575	0.96	0.94	0.99	6.7E-03	1.0E-01
Triglycerides in Large HDL	6575	0.96	0.94	0.99	5.6E-03	1.0E-01
Triglycerides in Medium HDL	6575	0.96	0.94	0.99	5.2E-03	1.0E-01
Saturated Fatty Acids	6570	0.97	0.94	0.99	1.1E-02	1.6E-01
Triglycerides in IDL	6575	0.97	0.94	0.99	1.2E-02	1.7E-01
Phosphoglycerides	6570	0.97	0.94	0.99	1.8E-02	1.8E-01
Total Fatty Acids	6570	0.97	0.94	0.99	1.5E-02	1.8E-01
Tyrosine	6568	0.97	0.95	1.00	2.0E-02	1.8E-01
Glucose	6560	0.97	0.94	0.99	1.8E-02	1.8E-01
Cholesterol in Chylomicrons and Extremely Large VLDL	6575	0.97	0.95	1.00	2.1E-02	1.8E-01
Cholesteryl Esters in Chylomicrons and Extremely Large VLDL	6575	0.97	0.95	1.00	2.1E-02	1.8E-01
Triglycerides in Small LDL	6575	0.97	0.94	0.99	1.7E-02	1.8E-01
Triglycerides in Very Large HDL	6575	0.97	0.95	0.99	1.9E-02	1.8E-01
Phospholipids in Small HDL	6575	0.97	0.94	1.00	2.0E-02	1.8E-01
Triglycerides in Small HDL	6575	0.97	0.94	0.99	1.5E-02	1.8E-01
Free Cholesterol to Total Lipids in Medium LDL percentage	6575	1.06	1.01	1.11	1.7E-02	1.8E-01
Cholesteryl Esters to Total Lipids in Large HDL percentage	6575	1.04	1.01	1.08	1.8E-02	1.8E-01
Phosphatidylcholines	6570	0.97	0.95	1.00	2.4E-02	1.8E-01
Triglycerides to Total Lipids in Large LDL percentage	6575	0.97	0.94	1.00	2.5E-02	1.8E-01

Appendix 5.4 Hazard ratios for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes.

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Free Cholesterol to Total Lipids in Small LDL percentage	6575	1.05	1.01	1.09	2.5E-02	1.8E-01
Triglycerides to Total Lipids in Small LDL percentage	6575	0.97	0.94	1.00	2.4E-02	1.8E-01
Total Lipids in Small HDL	6575	0.97	0.95	1.00	2.6E-02	1.9E-01
Triglycerides in Very Small VLDL	6575	0.97	0.95	1.00	3.0E-02	2.1E-01
Free Cholesterol in Small HDL	6575	0.97	0.95	1.00	3.3E-02	2.2E-01
Triglycerides to Total Lipids in Medium LDL percentage	6575	0.97	0.95	1.00	3.6E-02	2.4E-01
Total Cholines	6570	0.97	0.95	1.00	3.8E-02	2.4E-01
Cholesterol to Total Lipids in Large HDL percentage	6575	1.04	1.00	1.07	3.9E-02	2.4E-01
Linoleic Acid	6570	0.97	0.95	1.00	4.0E-02	2.4E-01
Phospholipids in Medium HDL	6575	0.97	0.95	1.00	4.2E-02	2.5E-01
Omega-6 Fatty Acids	6570	0.97	0.95	1.00	4.7E-02	2.6E-01
Free Cholesterol in Chylomicrons and Extremely Large VLDL	6575	0.97	0.95	1.00	4.5E-02	2.6E-01
Cholesteryl Esters to Total Lipids in IDL percentage	6575	1.03	1.00	1.05	4.7E-02	2.6E-01
Phospholipids to Total Lipids in Medium LDL percentage	6575	1.03	1.00	1.05	4.6E-02	2.6E-01
Triglycerides to Total Lipids in Large HDL percentage	6575	0.97	0.95	1.00	4.9E-02	2.6E-01
Cholesterol to Total Lipids in Small HDL percentage	6575	1.03	1.00	1.05	5.2E-02	2.7E-01
Total Phospholipids in Lipoprotein Particles	6575	0.98	0.95	1.00	5.4E-02	2.7E-01
Triglycerides to Total Lipids in IDL percentage	6575	0.97	0.95	1.00	5.5E-02	2.7E-01
Cholesteryl Esters to Total Lipids in Small HDL percentage	6575	1.03	1.00	1.06	5.6E-02	2.7E-01
Total Lipids in Medium HDL	6575	0.97	0.95	1.00	5.9E-02	2.8E-01
Triglycerides to Total Lipids in Very Large HDL percentage	6572	0.97	0.95	1.00	5.8E-02	2.8E-01
Apolipoprotein A1	6575	0.98	0.95	1.00	6.4E-02	2.9E-01
Total Triglycerides	6575	0.98	0.95	1.00	7.4E-02	3.1E-01
Omega-6 Fatty Acids to Total Fatty Acids percentage	6570	1.03	1.00	1.05	7.0E-02	3.1E-01
Cholesterol to Total Lipids in Very Small VLDL percentage	6575	1.02	1.00	1.05	7.4E-02	3.1E-01
Cholesterol to Total Lipids in IDL percentage	6575	1.03	1.00	1.05	6.9E-02	3.1E-01
Cholesterol to Total Lipids in Large LDL percentage	6575	1.14	0.99	1.31	7.1E-02	3.1E-01
Cholesteryl Esters to Total Lipids in Medium HDL percentage	6575	1.03	1.00	1.06	7.2E-02	3.1E-01
Triglycerides to Total Lipids in Medium HDL percentage	6575	0.98	0.95	1.00	7.5E-02	3.1E-01

Appendix 5.4 Hazard ratios for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes.

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Total Lipids in Lipoprotein Particles	6575	0.98	0.95	1.00	7.8E-02	3.1E-01
Cholesteryl Esters to Total Lipids in Very Small VLDL percentage	6575	1.02	1.00	1.05	7.8E-02	3.1E-01
Phospholipids in HDL	6575	0.98	0.95	1.00	8.7E-02	3.2E-01
Total Concentration of Lipoprotein Particles	6575	0.98	0.95	1.00	9.0E-02	3.2E-01
Concentration of HDL Particles	6575	0.98	0.95	1.00	9.2E-02	3.2E-01
Average Diameter for LDL Particles	6575	1.02	1.00	1.05	9.4E-02	3.2E-01
Saturated Fatty Acids to Total Fatty Acids percentage	6570	0.98	0.95	1.00	8.8E-02	3.2E-01
Alanine	6571	0.98	0.95	1.00	8.9E-02	3.2E-01
Concentration of Medium HDL Particles	6575	0.98	0.95	1.00	8.4E-02	3.2E-01
Cholesteryl Esters to Total Lipids in Medium VLDL percentage	6575	1.02	1.00	1.05	9.3E-02	3.2E-01
Free Cholesterol to Total Lipids in Small VLDL percentage	6575	1.02	1.00	1.05	9.3E-02	3.2E-01
Triglycerides to Total Lipids in Small HDL percentage	6575	0.98	0.95	1.00	8.9E-02	3.2E-01
Free Cholesterol to Total Lipids in Very Small VLDL percentage	6575	1.02	1.00	1.05	9.9E-02	3.3E-01
Cholesterol to Total Lipids in Medium VLDL percentage	6575	1.02	1.00	1.05	1.0E-01	3.3E-01
Omega-3 Fatty Acids	6570	1.02	1.00	1.05	1.1E-01	3.6E-01
Citrate	6575	0.98	0.96	1.01	1.2E-01	3.6E-01
Free Cholesterol in Medium HDL	6575	0.98	0.95	1.00	1.1E-01	3.6E-01
Concentration of Small HDL Particles	6575	0.98	0.95	1.00	1.1E-01	3.6E-01
Cholesterol to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	6367	0.98	0.96	1.00	1.2E-01	3.6E-01
Triglycerides to Total Lipids in Very Small VLDL percentage	6575	0.98	0.95	1.01	1.1E-01	3.6E-01
Total Lipids in HDL	6575	0.98	0.95	1.01	1.2E-01	3.6E-01
Glutamine	6549	0.98	0.96	1.01	1.2E-01	3.7E-01
Phospholipids to Total Lipids in Small VLDL percentage	6575	1.02	0.99	1.05	1.3E-01	3.7E-01
Cholesterol to Total Lipids in Medium HDL percentage	6575	1.02	0.99	1.06	1.3E-01	3.8E-01
Cholesteryl Esters to Total Lipids in Large LDL percentage	6575	1.09	0.97	1.22	1.3E-01	3.8E-01
Glycoprotein Acetyls	6575	1.02	0.99	1.05	1.4E-01	3.9E-01
Phospholipids in Chylomicrons and Extremely Large VLDL	6575	0.98	0.96	1.01	1.4E-01	4.0E-01
Total Lipids in VLDL	6575	0.98	0.96	1.01	1.5E-01	4.1E-01
Cholesterol in Small HDL	6575	0.98	0.96	1.01	1.5E-01	4.1E-01

Appendix 5.4 Hazard ratios for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes.

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Cholesteryl Esters to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	6367	0.98	0.96	1.01	1.5E-01	4.1E-01
Free Cholesterol to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	6367	0.98	0.96	1.01	1.5E-01	4.1E-01
Cholesteryl Esters to Total Lipids in Very Large HDL percentage	6572	1.02	0.99	1.06	1.5E-01	4.1E-01
Triglycerides in Small VLDL	6575	0.98	0.96	1.01	1.6E-01	4.2E-01
Phospholipids in VLDL	6575	0.98	0.96	1.01	1.6E-01	4.2E-01
Concentration of Chylomicrons and Extremely Large VLDL Particles	6575	0.98	0.96	1.01	1.6E-01	4.2E-01
Triglycerides in VLDL	6575	0.98	0.96	1.01	1.7E-01	4.3E-01
Free Cholesterol to Total Lipids in Large LDL percentage	6575	1.04	0.98	1.09	1.7E-01	4.3E-01
Glycine	6540	0.98	0.96	1.01	1.8E-01	4.5E-01
Phospholipids to Total Lipids in Very Large VLDL percentage	6548	0.99	0.96	1.01	1.8E-01	4.5E-01
Total Lipids in Chylomicrons and Extremely Large VLDL	6575	0.98	0.96	1.01	1.9E-01	4.5E-01
Cholesteryl Esters to Total Lipids in Very Large VLDL percentage	6548	1.02	0.99	1.04	1.9E-01	4.5E-01
Free Cholesterol in HDL	6575	0.98	0.96	1.01	1.9E-01	4.6E-01
Phospholipids in Very Small VLDL	6575	0.98	0.96	1.01	1.9E-01	4.6E-01
Average Diameter for VLDL Particles	6575	0.98	0.96	1.01	2.0E-01	4.7E-01
Albumin	6574	0.98	0.96	1.01	2.0E-01	4.7E-01
Total Lipids in Very Small VLDL	6575	0.98	0.96	1.01	2.0E-01	4.7E-01
Free Cholesterol to Total Lipids in Medium VLDL percentage	6575	1.02	0.99	1.04	2.0E-01	4.7E-01
Phospholipids to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	6367	0.99	0.96	1.01	2.1E-01	4.7E-01
Polyunsaturated Fatty Acids	6570	0.98	0.96	1.01	2.1E-01	4.7E-01
Free Cholesterol in VLDL	6575	0.98	0.96	1.01	2.2E-01	4.8E-01
Concentration of VLDL Particles	6575	0.98	0.96	1.01	2.2E-01	4.8E-01
Concentration of Very Small VLDL Particles	6575	0.98	0.96	1.01	2.1E-01	4.8E-01
Triglycerides to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	6367	1.02	0.99	1.04	2.2E-01	4.9E-01
Cholesterol in Medium HDL	6575	0.98	0.96	1.01	2.4E-01	5.1E-01
Cholesterol to Total Lipids in Medium LDL percentage	6575	1.08	0.95	1.23	2.4E-01	5.1E-01
Valine	6570	0.98	0.96	1.01	2.4E-01	5.1E-01
Concentration of Very Large VLDL Particles	6575	0.99	0.96	1.01	2.5E-01	5.2E-01

Appendix 5.4 Hazard ratios for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes.						
Metabolite	**Cases**	**HR**	**LL**	**UL**	**P-value[2]**	**Padj[2]**
Concentration of Large VLDL Particles	6575	0.99	0.96	1.01	2.5E-01	5.2E-01
Cholesterol to Total Lipids in Very Large HDL percentage	6572	1.02	0.99	1.04	2.5E-01	5.2E-01
Total Lipids in Small VLDL	6575	0.99	0.96	1.01	2.6E-01	5.2E-01
Histidine	6567	1.01	0.99	1.04	2.6E-01	5.3E-01
Cholesteryl Esters in Small HDL	6575	0.98	0.96	1.01	2.6E-01	5.3E-01
Cholesterol to Total Lipids in Small VLDL percentage	6575	1.02	0.99	1.04	2.6E-01	5.3E-01
Total Lipids in Very Large VLDL	6575	0.99	0.96	1.01	2.7E-01	5.3E-01
Free Cholesterol in Very Large VLDL	6575	0.99	0.96	1.01	2.7E-01	5.3E-01
Cholesterol in Very Large VLDL	6575	0.99	0.96	1.01	2.8E-01	5.4E-01
Cholesteryl Esters in Medium HDL	6575	0.99	0.96	1.01	2.8E-01	5.4E-01
Cholesteryl Esters to Total Lipids in Large VLDL percentage	6575	1.01	0.99	1.04	2.8E-01	5.4E-01
Concentration of Small VLDL Particles	6575	0.99	0.96	1.01	2.9E-01	5.5E-01
Concentration of Medium LDL Particles	6575	0.99	0.96	1.01	2.9E-01	5.5E-01
Leucine	6575	0.99	0.96	1.01	2.9E-01	5.5E-01
Cholesteryl Esters in Medium VLDL	6575	1.01	0.99	1.04	2.9E-01	5.5E-01
VLDL Cholesterol	6575	0.99	0.96	1.01	3.0E-01	5.6E-01
Free Cholesterol in Large VLDL	6575	0.99	0.96	1.01	3.1E-01	5.6E-01
Phospholipids to Total Lipids in Small HDL percentage	6575	0.99	0.96	1.01	3.1E-01	5.7E-01
Triglycerides to Phosphoglycerides ratio	6570	0.99	0.96	1.01	3.2E-01	5.8E-01
Phospholipids in Very Large VLDL	6575	0.99	0.96	1.01	3.3E-01	5.9E-01
Triglycerides in Medium VLDL	6575	0.99	0.96	1.01	3.3E-01	5.9E-01
Total Concentration of Branched-Chain Amino Acids (Leucine + Isoleucine + Valine)	6570	0.99	0.96	1.01	3.4E-01	6.0E-01
Total Lipids in Medium LDL	6575	0.99	0.96	1.01	3.4E-01	6.0E-01
Cholesterol to Total Lipids in Very Large VLDL percentage	6548	1.01	0.99	1.04	3.4E-01	6.0E-01
Free Cholesterol in Very Small VLDL	6575	0.99	0.96	1.01	3.4E-01	6.0E-01
Cholesteryl Esters in Very Large HDL	6575	1.01	0.99	1.04	3.5E-01	6.0E-01
Total Lipids in Large VLDL	6575	0.99	0.96	1.01	3.5E-01	6.0E-01
Cholesteryl Esters in Small LDL	6575	0.99	0.97	1.01	3.5E-01	6.0E-01
Triglycerides to Total Lipids in Small VLDL percentage	6575	0.99	0.96	1.01	3.5E-01	6.0E-01
Cholesteryl Esters in Large HDL	6575	1.01	0.99	1.04	3.6E-01	6.0E-01

Appendix 5.4 Hazard ratios for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes.

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Cholesteryl Esters in Medium LDL	6575	0.99	0.97	1.01	3.6E-01	6.1E-01

Abbreviations: SD, standard deviation; HR, hazard ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit; LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol; IDL, intermediate-density lipoprotein cholesterol; VLDL, very low-density lipoprotein cholesterol.

1 Hazard ratios (95% confidence interval) for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features, after additional adjustment for diabetes. Analyses were stratified by region and Townsend Deprivation Index with age as the underlying time variable. Adjusted for diabetes, fasting status, body mass index, height, alcohol intake, family history of prostate cancer, ethnicity, level of education, smoking status, and weekly physical activity.

2 Padj < 0.05 are statistically significant after allowing for multiple testing using a false discovery rate controlling procedure at $\alpha = 0.05$ (Benjamini-Hochberg). Padj values are sorted in ascending order. Associations with Padj < 0.05 are bolded.

Appendix 5.5 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1] after additional adjustment for diabetes, stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Metabolites	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj
Total Concentration of Lipoprotein Particles	1309	0.901	0.850	0.955	4.4E-04	2.0E-02	2159	1.00	0.95	1.04	9.3E-01	1.0E+00	3107	1.00	0.96	1.04	0.94	1.00
Concentration of HDL Particles	1309	0.900	0.849	0.954	4.2E-04	2.0E-02	2159	1.00	0.95	1.05	9.8E-01	1.0E+00	3107	1.00	0.96	1.04	0.92	1.00
Concentration of Small HDL Particles	1309	0.902	0.852	0.955	3.9E-04	2.0E-02	2159	1.00	0.96	1.05	9.9E-01	1.0E+00	3107	1.00	0.96	1.04	0.98	1.00
Total Lipids in Small HDL	1309	0.901	0.850	0.955	4.1E-04	2.0E-02	2159	0.99	0.94	1.03	5.5E-01	9.1E-01	3107	0.99	0.95	1.03	0.68	1.00
Phospholipids in Small HDL	1309	0.902	0.851	0.956	5.2E-04	2.0E-02	2159	0.98	0.94	1.03	4.7E-01	8.8E-01	3107	0.99	0.95	1.03	0.62	1.00
Cholesterol in Small HDL	1309	0.902	0.852	0.955	4.2E-04	2.0E-02	2159	1.01	0.96	1.05	8.3E-01	9.8E-01	3107	1.00	0.96	1.04	0.95	1.00
Cholesteryl Esters in Small HDL	1309	0.902	0.850	0.956	5.6E-04	2.0E-02	2159	1.01	0.96	1.06	6.4E-01	9.4E-01	3107	1.00	0.97	1.05	0.82	1.00
Phospholipids to Total Lipids in Small LDL percentage	1309	1.126	1.050	1.206	7.9E-04	2.5E-02	2159	1.08	1.02	1.14	7.7E-03	3.5E-01	3107	1.03	0.98	1.08	0.27	1.00
Leucine	1309	0.912	0.862	0.964	1.2E-03	3.0E-02	2159	0.99	0.95	1.03	5.6E-01	9.1E-01	3107	1.02	0.98	1.06	0.28	1.00
Free Cholesterol in Small HDL	1309	0.909	0.858	0.963	1.1E-03	3.0E-02	2159	0.99	0.94	1.03	5.7E-01	9.1E-01	3107	0.99	0.95	1.03	0.63	1.00
Albumin	1308	0.914	0.866	0.966	1.3E-03	3.0E-02	2159	1.00	0.96	1.04	1.0E+00	1.0E+00	3107	1.01	0.97	1.04	0.78	1.00
Apolipoprotein A1	1309	0.909	0.857	0.965	1.6E-03	3.3E-02	2159	0.99	0.95	1.04	7.6E-01	9.7E-01	3107	0.99	0.96	1.03	0.73	1.00
Concentration of Medium HDL Particles	1309	0.928	0.885	0.973	1.9E-03	3.6E-02	2159	1.00	0.95	1.04	8.6E-01	1.0E+00	3107	0.99	0.96	1.03	0.73	1.00
Total Lipids in Medium HDL	1309	0.913	0.861	0.968	2.2E-03	3.9E-02	2159	0.99	0.95	1.04	7.4E-01	9.7E-01	3107	0.99	0.95	1.03	0.65	1.00
Phospholipids in Medium HDL	1309	0.914	0.862	0.970	3.1E-03	4.5E-02	2159	0.99	0.94	1.04	6.4E-01	9.4E-01	3107	0.99	0.95	1.03	0.54	1.00
Cholesterol in Medium HDL	1309	0.927	0.882	0.974	2.8E-03	4.5E-02	2159	1.01	0.96	1.05	8.1E-01	9.8E-01	3107	1.00	0.96	1.04	0.98	1.00
Cholesteryl Esters in Medium HDL	1309	0.927	0.882	0.975	2.9E-03	4.5E-02	2159	1.01	0.96	1.05	7.6E-01	9.7E-01	3107	1.00	0.96	1.04	0.95	1.00
Free Cholesterol in Medium HDL	1309	0.926	0.879	0.975	3.4E-03	4.8E-02	2159	1.00	0.95	1.05	9.6E-01	1.0E+00	3107	0.99	0.95	1.03	0.70	1.00
BCAA	1307	0.919	0.868	0.973	3.7E-03	4.9E-02	2159	0.99	0.95	1.03	6.1E-01	9.3E-01	3104	1.02	0.98	1.06	0.36	1.00
Phosphatidylcholines	1307	0.920	0.869	0.974	4.1E-03	5.1E-02	2159	0.98	0.94	1.03	4.0E-01	8.4E-01	3104	0.99	0.95	1.03	0.51	1.00
Phosphoglycerides	1307	0.921	0.870	0.975	4.7E-03	5.5E-02	2159	0.98	0.94	1.02	3.6E-01	8.0E-01	3104	0.98	0.95	1.02	0.43	1.00

Appendix 5.5 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1] after additional adjustment for diabetes, stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Total Cholines	1307	0.922	0.871	0.977	5.5E-03	6.2E-02	2159	0.98	0.94	1.03	4.5E-01	8.6E-01	3104	0.99	0.95	1.03	0.59	1.00
Total Lipids in HDL	1309	0.923	0.870	0.980	8.4E-03	9.0E-02	2159	0.99	0.95	1.04	8.3E-01	9.8E-01	3107	0.99	0.96	1.03	0.74	1.00
Glutamine	1304	0.929	0.879	0.982	9.2E-03	9.2E-02	2151	1.00	0.96	1.05	8.9E-01	1.0E+00	3094	0.99	0.95	1.02	0.52	1.00
Isoleucine	1308	0.929	0.879	0.982	9.2E-03	9.2E-02	2159	1.01	0.96	1.05	7.7E-01	9.7E-01	3107	1.03	0.99	1.07	0.12	1.00
Phospholipids in HDL	1309	0.925	0.872	0.982	1.0E-02	9.7E-02	2159	0.99	0.95	1.04	7.2E-01	9.7E-01	3107	0.99	0.95	1.03	0.63	1.00
Total Phospholipids in Lipoprotein Particles	1309	0.929	0.878	0.983	1.1E-02	1.0E-01	2159	0.98	0.94	1.03	4.1E-01	8.4E-01	3107	0.99	0.96	1.03	0.67	1.00
Valine	1307	0.929	0.876	0.984	1.2E-02	1.1E-01	2159	0.98	0.94	1.03	4.9E-01	8.9E-01	3104	1.01	0.97	1.05	0.61	1.00
Cholesteryl Esters in HDL	1309	0.928	0.874	0.985	1.4E-02	1.2E-01	2159	1.01	0.97	1.06	5.6E-01	9.1E-01	3107	1.01	0.97	1.05	0.78	1.00
HDL Cholesterol	1309	0.928	0.875	0.985	1.4E-02	1.2E-01	2159	1.01	0.96	1.06	6.7E-01	9.5E-01	3107	1.00	0.96	1.04	0.87	1.00
Pyruvate	1302	0.936	0.887	0.989	1.9E-02	1.5E-01	2152	1.01	0.97	1.06	6.5E-01	9.4E-01	3095	1.03	0.99	1.06	0.19	1.00
Citrate	1309	0.935	0.883	0.989	1.9E-02	1.5E-01	2159	1.00	0.96	1.05	9.8E-01	1.0E+00	3107	0.99	0.95	1.02	0.45	1.00
Free Cholesterol in HDL	1309	0.935	0.882	0.991	2.4E-02	1.8E-01	2159	1.00	0.95	1.04	8.8E-01	1.0E+00	3107	0.99	0.96	1.03	0.79	1.00
Total Esterified Cholesterol	1309	0.938	0.886	0.992	2.5E-02	1.9E-01	2159	1.00	0.96	1.05	9.8E-01	1.0E+00	3107	1.01	0.97	1.05	0.74	1.00
Total Cholesterol	1309	0.941	0.889	0.996	3.4E-02	2.3E-01	2159	1.00	0.95	1.04	9.0E-01	1.0E+00	3107	1.01	0.97	1.04	0.75	1.00
Total Lipids in Lipoprotein Particles	1309	0.941	0.890	0.996	3.4E-02	2.3E-01	2159	0.98	0.94	1.02	3.5E-01	7.9E-01	3107	0.99	0.96	1.03	0.70	1.00
Phospholipids to Total Lipids in IDL percentage	1309	1.064	1.005	1.126	3.3E-02	2.3E-01	2159	1.00	0.96	1.04	9.4E-01	1.0E+00	3107	0.98	0.95	1.02	0.30	1.00
Phospholipids to Total Lipids in Very Small VLDL percentage	1309	1.063	1.003	1.126	4.0E-02	2.6E-01	2159	0.99	0.94	1.03	5.2E-01	9.1E-01	3107	0.97	0.94	1.01	0.15	1.00
Triglycerides to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	1269	1.072	1.003	1.146	4.2E-02	2.7E-01	2084	1.03	0.99	1.08	1.7E-01	5.9E-01	3014	0.99	0.96	1.02	0.47	1.00
Omega-6 Fatty Acids	1307	0.944	0.892	0.998	4.4E-02	2.7E-01	2159	0.97	0.93	1.02	2.4E-01	6.8E-01	3104	0.99	0.95	1.03	0.57	1.00
Lactate	1305	0.945	0.894	0.999	4.5E-02	2.8E-01	2156	1.00	0.95	1.04	8.9E-01	1.0E+00	3102	1.01	0.97	1.05	0.55	1.00
Total Fatty Acids	1307	0.946	0.893	1.002	6.0E-02	3.5E-01	2159	0.97	0.93	1.01	1.7E-01	5.9E-01	3104	0.98	0.94	1.02	0.25	1.00
Saturated Fatty Acids	1307	0.946	0.892	1.003	6.3E-02	3.5E-01	2159	0.97	0.92	1.01	1.3E-01	5.4E-01	3104	0.98	0.94	1.01	0.23	1.00
Total Lipids in Large LDL	1309	0.948	0.897	1.003	6.2E-02	3.5E-01	2159	1.00	0.96	1.05	9.0E-01	1.0E+00	3107	1.01	0.97	1.04	0.74	1.00
Phospholipids in Large LDL	1309	0.948	0.897	1.003	6.4E-02	3.5E-01	2159	1.00	0.96	1.05	9.7E-01	1.0E+00	3107	1.01	0.97	1.05	0.66	1.00

Appendix 5.5 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1] after additional adjustment for diabetes, stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Cholesteryl Esters to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	1269	0.949	0.896	1.004	6.7E-02	3.6E-01	2084	0.98	0.94	1.02	3.4E-01	7.8E-01	3014	1.00	0.96	1.03	0.92	1.00
Total Free Cholesterol	1309	0.950	0.898	1.005	7.4E-02	3.8E-01	2159	0.99	0.95	1.04	6.9E-01	9.5E-01	3107	1.00	0.97	1.04	0.80	1.00
Cholesterol to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	1269	0.949	0.896	1.005	7.3E-02	3.8E-01	2084	0.98	0.94	1.03	4.2E-01	8.4E-01	3014	0.99	0.96	1.03	0.64	1.00
Free Cholesterol to Total Lipids in Medium HDL percentage	1309	0.964	0.927	1.003	7.3E-02	3.8E-01	2159	1.01	0.97	1.06	5.8E-01	9.2E-01	3107	1.00	0.96	1.04	0.90	1.00
Total Lipids in LDL	1309	0.951	0.900	1.006	7.8E-02	3.8E-01	2159	1.00	0.95	1.04	9.0E-01	1.0E+00	3107	1.00	0.97	1.04	0.80	1.00
Sphingomyelins	1307	0.950	0.898	1.006	7.9E-02	3.8E-01	2159	1.02	0.97	1.06	5.0E-01	8.9E-01	3104	1.01	0.97	1.05	0.55	1.00
Polyunsaturated Fatty Acids	1307	0.951	0.898	1.006	7.8E-02	3.8E-01	2159	0.99	0.95	1.03	6.2E-01	9.3E-01	3104	1.00	0.96	1.03	0.80	1.00
Monounsaturated Fatty Acids	1307	0.951	0.897	1.008	9.1E-02	4.0E-01	2159	0.96	0.92	1.00	7.1E-02	4.2E-01	3104	0.97	0.93	1.01	0.12	1.00
Linoleic Acid	1307	0.953	0.902	1.008	9.3E-02	4.0E-01	2159	0.96	0.92	1.01	1.1E-01	5.2E-01	3104	0.99	0.95	1.03	0.59	1.00
Total Lipids in Medium LDL	1309	0.954	0.902	1.008	9.3E-02	4.0E-01	2159	0.99	0.95	1.03	6.1E-01	9.3E-01	3107	1.00	0.97	1.04	0.89	1.00
Phospholipids in Medium LDL	1309	0.953	0.902	1.007	9.0E-02	4.0E-01	2159	0.99	0.95	1.04	7.9E-01	9.7E-01	3107	1.00	0.97	1.04	0.83	1.00
Triglycerides in Medium HDL	1309	0.951	0.897	1.007	8.6E-02	4.0E-01	2159	0.95	0.91	1.00	3.7E-02	3.5E-01	3107	0.98	0.94	1.01	0.23	1.00
Phospholipids to Total Lipids in Medium HDL percentage	1309	1.064	0.990	1.144	9.1E-02	4.0E-01	2159	0.98	0.95	1.02	3.0E-01	7.5E-01	3107	0.98	0.95	1.01	0.27	1.00
Cholesteryl Esters in IDL	1309	0.953	0.901	1.009	9.6E-02	4.1E-01	2159	1.01	0.96	1.05	8.0E-01	9.7E-01	3107	1.02	0.98	1.06	0.41	1.00
Phospholipids in LDL	1309	0.954	0.903	1.009	1.0E-01	4.1E-01	2159	1.00	0.96	1.04	9.6E-01	1.0E+00	3107	1.01	0.97	1.05	0.72	1.00
Tyrosine	1307	0.954	0.901	1.009	1.0E-01	4.1E-01	2158	0.95	0.91	0.99	2.6E-02	3.5E-01	3103	0.99	0.96	1.03	0.65	1.00
Total Lipids in IDL	1309	0.955	0.903	1.010	1.1E-01	4.3E-01	2159	1.00	0.96	1.04	9.6E-01	1.0E+00	3107	1.01	0.97	1.05	0.68	1.00
Cholesterol in IDL	1309	0.955	0.902	1.010	1.1E-01	4.3E-01	2159	1.01	0.96	1.05	8.2E-01	9.8E-01	3107	1.01	0.98	1.05	0.47	1.00
Clinical LDL Cholesterol	1309	0.956	0.904	1.011	1.1E-01	4.4E-01	2159	1.00	0.96	1.05	9.3E-01	1.0E+00	3107	1.01	0.98	1.05	0.49	1.00
LDL Cholesterol	1309	0.972	0.937	1.007	1.1E-01	4.4E-01	2159	1.00	0.96	1.05	9.1E-01	1.0E+00	3107	1.01	0.97	1.05	0.63	1.00

Appendix 5.5 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1] after additional adjustment for diabetes, stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Total Cholesterol Minus HDL-C	1309	0.958	0.906	1.013	1.3E-01	4.6E-01	2159	0.99	0.95	1.04	7.7E-01	9.7E-01	3107	1.01	0.97	1.04	0.75	1.00
Triglycerides in HDL	1309	0.956	0.902	1.013	1.3E-01	4.6E-01	2159	0.95	0.91	0.99	2.7E-02	3.5E-01	3107	0.98	0.94	1.01	0.19	1.00
Cholesteryl Esters in LDL	1309	0.973	0.939	1.007	1.2E-01	4.6E-01	2159	1.00	0.95	1.05	9.7E-01	1.0E+00	3107	1.01	0.97	1.05	0.70	1.00
Cholesteryl Esters in Large LDL	1309	0.968	0.929	1.009	1.2E-01	4.6E-01	2159	1.01	0.96	1.06	6.9E-01	9.5E-01	3107	1.01	0.97	1.05	0.58	1.00
Cholesterol in Medium LDL	1309	0.970	0.933	1.009	1.3E-01	4.6E-01	2159	0.99	0.95	1.04	7.2E-01	9.7E-01	3107	1.01	0.97	1.05	0.72	1.00
Cholesteryl Esters in Medium LDL	1309	0.971	0.934	1.009	1.3E-01	4.7E-01	2159	0.99	0.95	1.03	5.0E-01	8.9E-01	3107	1.00	0.97	1.04	0.85	1.00
Free Cholesterol in Large LDL	1309	0.971	0.934	1.010	1.4E-01	4.9E-01	2159	1.02	0.97	1.07	5.3E-01	9.1E-01	3107	1.02	0.98	1.06	0.45	1.00
Free Cholesterol in LDL	1309	0.971	0.933	1.011	1.5E-01	5.2E-01	2159	1.01	0.97	1.07	5.5E-01	9.1E-01	3107	1.02	0.98	1.06	0.43	1.00
Free Cholesterol in IDL	1309	0.960	0.907	1.016	1.6E-01	5.3E-01	2159	1.00	0.96	1.05	8.9E-01	1.0E+00	3107	1.01	0.97	1.05	0.68	1.00
Cholesterol in Large LDL	1309	0.975	0.942	1.011	1.7E-01	5.6E-01	2159	1.01	0.96	1.06	6.4E-01	9.4E-01	3107	1.01	0.97	1.06	0.52	1.00
Triglycerides in Large HDL	1309	0.960	0.906	1.018	1.7E-01	5.6E-01	2159	0.95	0.91	0.99	2.9E-02	3.5E-01	3107	0.98	0.94	1.01	0.19	1.00
Cholesteryl Esters in Small LDL	1309	0.974	0.939	1.011	1.7E-01	5.6E-01	2159	0.98	0.95	1.02	4.4E-01	8.6E-01	3107	1.00	0.96	1.04	0.89	1.00
Free Cholesterol to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	1269	0.963	0.911	1.017	1.8E-01	5.7E-01	2084	1.00	0.96	1.04	9.8E-01	1.0E+00	3014	0.98	0.94	1.01	0.22	1.00
Triglycerides in Chylomicrons and Extremely Large VLDL	1309	1.040	0.982	1.101	1.8E-01	5.7E-01	2159	0.98	0.94	1.02	3.9E-01	8.4E-01	3107	1.00	0.96	1.03	0.84	1.00
Phospholipids in IDL	1309	0.962	0.910	1.018	1.8E-01	5.7E-01	2159	1.00	0.96	1.04	9.5E-01	1.0E+00	3107	1.00	0.97	1.04	0.79	1.00
Triglycerides in Large LDL	1309	0.962	0.907	1.020	1.9E-01	5.8E-01	2159	0.96	0.91	1.00	5.1E-02	3.5E-01	3107	0.97	0.93	1.01	0.12	1.00
Free Cholesterol in Medium LDL	1309	0.973	0.934	1.014	1.9E-01	5.8E-01	2159	1.01	0.97	1.06	5.6E-01	9.1E-01	3107	1.02	0.98	1.06	0.40	1.00
Cholesteryl Esters to Total Lipids in IDL percentage	1309	0.964	0.910	1.020	2.0E-01	6.1E-01	2159	1.04	0.99	1.09	1.1E-01	5.4E-01	3107	1.05	1.01	1.09	0.02	0.98
3-Hydroxybutyrate	1280	0.965	0.911	1.021	2.1E-01	6.3E-01	2109	1.00	0.96	1.05	9.3E-01	1.0E+00	3038	1.02	0.98	1.06	0.36	1.00
Triglycerides in LDL	1309	0.964	0.909	1.023	2.2E-01	6.5E-01	2159	0.95	0.91	1.00	4.3E-02	3.5E-01	3107	0.97	0.93	1.01	0.13	1.00
Cholesterol in Small LDL	1309	0.976	0.939	1.015	2.2E-01	6.5E-01	2159	0.99	0.95	1.04	7.1E-01	9.7E-01	3107	1.01	0.97	1.05	0.73	1.00

Appendix 5.5 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1] after additional adjustment for diabetes, stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Creatinine	1281	1.034	0.979	1.092	2.3E-01	6.5E-01	2114	1.00	0.96	1.05	8.9E-01	1.0E+00	3051	1.00	0.96	1.04	0.98	1.00
Remnant Cholesterol (Non-HDL, Non-LDL - Cholesterol)	1309	0.967	0.915	1.022	2.4E-01	6.5E-01	2159	0.99	0.94	1.03	5.3E-01	9.1E-01	3107	1.00	0.97	1.04	0.84	1.00
Glucose	1303	0.963	0.906	1.024	2.3E-01	6.5E-01	2157	0.96	0.92	1.01	1.2E-01	5.4E-01	3100	0.97	0.93	1.01	0.18	1.00
Triglycerides in Medium LDL	1309	0.965	0.910	1.023	2.4E-01	6.5E-01	2159	0.95	0.91	1.00	4.0E-02	3.5E-01	3107	0.97	0.94	1.01	0.14	1.00
Glycine	1305	0.968	0.918	1.022	2.4E-01	6.5E-01	2147	1.00	0.95	1.04	8.5E-01	9.9E-01	3088	0.98	0.95	1.02	0.30	1.00
Triglycerides in Small HDL	1309	0.966	0.912	1.024	2.4E-01	6.6E-01	2159	0.95	0.91	1.00	4.4E-02	3.5E-01	3107	0.98	0.94	1.02	0.26	1.00
Acetate	1303	1.033	0.977	1.093	2.5E-01	6.8E-01	2150	0.99	0.95	1.04	7.4E-01	9.7E-01	3090	0.99	0.96	1.03	0.78	1.00
Concentration of Medium LDL Particles	1309	0.969	0.916	1.024	2.6E-01	6.9E-01	2159	0.98	0.94	1.02	3.1E-01	7.6E-01	3107	1.00	0.96	1.04	0.98	1.00
Acetone	1309	0.969	0.916	1.025	2.7E-01	6.9E-01	2159	1.01	0.97	1.06	5.9E-01	9.3E-01	3107	1.03	0.99	1.06	0.16	1.00
Total Lipids in Small LDL	1309	0.969	0.917	1.025	2.7E-01	6.9E-01	2159	0.99	0.95	1.03	6.4E-01	9.4E-01	3107	1.00	0.97	1.04	0.87	1.00
Triglycerides to Total Lipids in Very Large VLDL percentage	1306	1.038	0.972	1.109	2.7E-01	6.9E-01	2150	0.97	0.94	1.00	3.9E-02	3.5E-01	3092	1.02	0.98	1.06	0.42	1.00
Concentration of IDL Particles	1309	0.970	0.917	1.026	2.9E-01	7.3E-01	2159	0.99	0.95	1.04	7.6E-01	9.7E-01	3107	1.01	0.97	1.05	0.59	1.00
Docosahexaenoic Acid to Total Fatty Acids percentage	1307	1.032	0.973	1.095	3.0E-01	7.4E-01	2159	1.08	1.03	1.13	1.6E-03	3.5E-01	3104	1.05	1.01	1.09	0.01	0.98
Glycoprotein Acetyls	1309	1.031	0.973	1.092	3.0E-01	7.5E-01	2159	1.03	0.98	1.08	2.0E-01	6.2E-01	3107	1.01	0.97	1.05	0.69	1.00
Polyunsaturated Fatty Acids to Total Fatty Acids percentage	1307	1.033	0.971	1.098	3.1E-01	7.5E-01	2159	1.05	1.00	1.11	3.3E-02	3.5E-01	3104	1.04	1.00	1.08	0.04	0.98
Total Lipids in Large HDL	1309	0.970	0.915	1.030	3.2E-01	7.9E-01	2159	1.01	0.96	1.05	8.3E-01	9.8E-01	3107	1.00	0.96	1.04	1.00	1.00
Cholesterol to Total Lipids in IDL percentage	1309	0.971	0.917	1.029	3.3E-01	7.9E-01	2159	1.04	0.99	1.09	1.2E-01	5.4E-01	3107	1.04	1.00	1.08	0.05	0.98
Apolipoprotein B	1309	0.973	0.920	1.029	3.4E-01	8.0E-01	2159	0.99	0.95	1.03	6.2E-01	9.3E-01	3107	1.00	0.97	1.04	0.87	1.00
Polyunsaturated Fatty Acids to Monounsaturated Fatty Acids ratio	1307	1.030	0.969	1.095	3.4E-01	8.0E-01	2159	1.05	1.01	1.11	2.8E-02	3.5E-01	3104	1.04	1.00	1.08	0.04	0.98
Concentration of LDL Particles	1309	0.973	0.920	1.029	3.4E-01	8.0E-01	2159	0.99	0.95	1.04	6.9E-01	9.5E-01	3107	1.00	0.97	1.04	0.87	1.00

Appendix 5.5 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1] after additional adjustment for diabetes, stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Phospholipids in Large HDL	1309	0.973	0.917	1.032	3.6E-01	8.4E-01	2159	1.01	0.96	1.06	7.8E-01	9.7E-01	3107	1.00	0.96	1.04	0.93	1.00
Cholesteryl Esters to Total Lipids in Very Large HDL percentage	1309	0.981	0.940	1.023	3.7E-01	8.5E-01	2159	1.02	0.97	1.08	4.4E-01	8.5E-01	3104	1.08	1.00	1.17	0.04	0.98
Saturated Fatty Acids to Total Fatty Acids percentage	1307	0.974	0.919	1.033	3.8E-01	8.6E-01	2159	0.97	0.93	1.02	2.2E-01	6.6E-01	3104	0.98	0.95	1.02	0.39	1.00
Cholesteryl Esters to Total Lipids in Medium VLDL percentage	1309	1.029	0.966	1.095	3.8E-01	8.6E-01	2159	1.02	0.97	1.07	5.0E-01	8.9E-01	3107	1.03	0.99	1.07	0.19	1.00
Concentration of Large LDL Particles	1309	0.976	0.923	1.032	3.9E-01	8.6E-01	2159	1.00	0.96	1.04	9.9E-01	1.0E+00	3107	1.00	0.97	1.04	0.80	1.00
Triglycerides in IDL	1309	0.975	0.920	1.034	3.9E-01	8.6E-01	2159	0.95	0.91	1.00	4.4E-02	3.5E-01	3107	0.97	0.94	1.01	0.16	1.00
Concentration of Small LDL Particles	1309	0.976	0.923	1.032	3.9E-01	8.6E-01	2159	0.98	0.94	1.02	3.8E-01	8.2E-01	3107	1.00	0.97	1.04	0.94	1.00
Concentration of Large HDL Particles	1309	0.975	0.920	1.033	4.0E-01	8.6E-01	2159	1.01	0.96	1.06	7.7E-01	9.7E-01	3107	1.00	0.96	1.04	0.93	1.00
Apolipoprotein B to Apolipoprotein A1 ratio	1309	1.024	0.968	1.084	4.1E-01	8.7E-01	2159	0.99	0.95	1.04	7.7E-01	9.7E-01	3107	1.01	0.97	1.04	0.74	1.00
Monounsaturated Fatty Acids to Total Fatty Acids percentage	1307	0.975	0.918	1.035	4.1E-01	8.7E-01	2159	0.95	0.91	1.00	3.6E-02	3.5E-01	3104	0.96	0.93	1.00	0.06	0.98
Free Cholesterol in Very Large HDL	1309	1.025	0.967	1.086	4.1E-01	8.7E-01	2159	1.00	0.96	1.05	9.2E-01	1.0E+00	3107	1.00	0.97	1.04	0.81	1.00
Cholesterol to Total Lipids in Medium HDL percentage	1309	0.980	0.934	1.028	4.1E-01	8.7E-01	2159	1.05	0.99	1.11	7.9E-02	4.5E-01	3107	1.03	0.98	1.08	0.19	1.00
Cholesteryl Esters to Total Lipids in Small LDL percentage	1309	0.989	0.964	1.015	4.2E-01	8.8E-01	2159	0.99	0.97	1.01	4.0E-01	8.4E-01	3107	1.00	0.96	1.04	1.00	1.00
VLDL Cholesterol	1309	0.984	0.931	1.040	5.6E-01	9.0E-01	2159	0.98	0.93	1.02	2.6E-01	6.9E-01	3107	1.00	0.96	1.03	0.86	1.00
Total Triglycerides	1309	0.982	0.927	1.041	5.5E-01	9.0E-01	2159	0.96	0.92	1.01	9.0E-02	4.7E-01	3107	0.98	0.95	1.02	0.43	1.00
Triglycerides in VLDL	1309	0.989	0.933	1.047	7.0E-01	9.0E-01	2159	0.97	0.92	1.01	1.3E-01	5.4E-01	3107	0.99	0.95	1.03	0.62	1.00
Phospholipids in VLDL	1309	0.984	0.931	1.041	5.8E-01	9.0E-01	2159	0.97	0.93	1.01	1.7E-01	5.9E-01	3107	0.99	0.95	1.03	0.61	1.00
Cholesteryl Esters in VLDL	1309	0.984	0.932	1.040	5.8E-01	9.0E-01	2159	0.98	0.94	1.02	3.3E-01	7.7E-01	3107	1.00	0.96	1.04	0.99	1.00
Free Cholesterol in VLDL	1309	0.984	0.930	1.040	5.6E-01	9.0E-01	2159	0.97	0.93	1.02	2.1E-01	6.3E-01	3107	0.99	0.96	1.03	0.72	1.00

Appendix 5.5 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1] after additional adjustment for diabetes, stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Total Lipids in VLDL	1309	0.984	0.930	1.042	5.8E-01	9.0E-01	2159	0.97	0.93	1.01	1.5E-01	5.6E-01	3107	0.99	0.95	1.03	0.60	1.00
Concentration of VLDL Particles	1309	0.984	0.931	1.041	5.8E-01	9.0E-01	2159	0.97	0.93	1.02	2.3E-01	6.7E-01	3107	0.99	0.96	1.03	0.67	1.00
Average Diameter for VLDL Particles	1309	0.985	0.930	1.044	6.1E-01	9.0E-01	2159	0.97	0.93	1.02	2.1E-01	6.5E-01	3107	0.99	0.95	1.03	0.61	1.00
Triglycerides to Phosphoglycerides ratio	1307	1.019	0.961	1.081	5.2E-01	9.0E-01	2159	0.96	0.92	1.01	1.3E-01	5.4E-01	3104	0.99	0.95	1.03	0.55	1.00
Degree of Unsaturation	1307	1.018	0.961	1.079	5.4E-01	9.0E-01	2159	1.07	1.02	1.12	4.7E-03	3.5E-01	3104	1.04	1.00	1.08	0.04	0.98
Omega-3 Fatty Acids	1307	0.989	0.934	1.047	7.0E-01	9.0E-01	2159	1.04	1.00	1.09	6.7E-02	4.2E-01	3104	1.02	0.98	1.06	0.29	1.00
Omega-3 Fatty Acids to Total Fatty Acids percentage	1307	1.019	0.962	1.078	5.3E-01	9.0E-01	2159	1.07	1.02	1.12	2.9E-03	3.5E-01	3104	1.04	1.00	1.08	0.06	0.98
Omega-6 Fatty Acids to Total Fatty Acids percentage	1307	1.021	0.961	1.086	5.0E-01	9.0E-01	2159	1.02	0.98	1.07	3.4E-01	7.7E-01	3104	1.03	0.99	1.07	0.16	1.00
Omega-6 Fatty Acids to Omega-3 Fatty Acids ratio	1306	0.979	0.925	1.036	4.6E-01	9.0E-01	2159	0.95	0.90	0.99	1.3E-02	3.5E-01	3104	0.97	0.94	1.01	0.17	1.00
Histidine	1307	0.980	0.928	1.036	4.8E-01	9.0E-01	2158	1.00	0.96	1.04	9.9E-01	1.0E+00	3102	1.04	1.00	1.08	0.04	0.98
Phospholipids in Chylomicrons and Extremely Large VLDL	1309	1.013	0.958	1.072	6.4E-01	9.0E-01	2159	0.96	0.93	1.01	8.7E-02	4.7E-01	3107	0.98	0.95	1.02	0.30	1.00
Cholesterol in Chylomicrons and Extremely Large VLDL	1309	0.983	0.930	1.040	5.5E-01	9.0E-01	2159	0.95	0.91	0.99	1.2E-02	3.5E-01	3107	0.98	0.95	1.02	0.38	1.00
Cholesteryl Esters in Chylomicrons and Extremely Large VLDL	1309	0.979	0.927	1.035	4.6E-01	9.0E-01	2159	0.95	0.91	0.99	9.6E-03	3.5E-01	3107	0.99	0.95	1.02	0.49	1.00
Concentration of Very Large VLDL Particles	1309	0.989	0.935	1.045	6.9E-01	9.0E-01	2159	0.97	0.93	1.01	1.9E-01	6.2E-01	3107	0.99	0.96	1.03	0.75	1.00
Cholesterol in Very Large VLDL	1309	0.990	0.937	1.046	7.2E-01	9.0E-01	2159	0.97	0.93	1.01	1.9E-01	6.2E-01	3107	1.00	0.96	1.03	0.82	1.00
Cholesteryl Esters in Very Large VLDL	1309	0.990	0.938	1.045	7.1E-01	9.0E-01	2159	0.98	0.94	1.02	2.4E-01	6.8E-01	3107	1.00	0.96	1.04	1.00	1.00
Concentration of Large VLDL Particles	1309	0.988	0.934	1.045	6.8E-01	9.0E-01	2159	0.97	0.93	1.02	2.0E-01	6.3E-01	3107	0.99	0.96	1.03	0.74	1.00
Total Lipids in Large VLDL	1309	0.989	0.935	1.046	7.0E-01	9.0E-01	2159	0.98	0.93	1.02	2.6E-01	6.9E-01	3107	1.00	0.96	1.03	0.87	1.00
Cholesterol in Large VLDL	1309	0.990	0.937	1.046	7.1E-01	9.0E-01	2159	0.98	0.94	1.02	3.1E-01	7.6E-01	3107	1.00	0.96	1.03	0.89	1.00

Appendix 5.5 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1] after additional adjustment for diabetes, stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Free Cholesterol in Large VLDL	1309	0.986	0.933	1.042	6.2E-01	9.0E-01	2159	0.98	0.93	1.02	2.7E-01	7.0E-01	3107	1.00	0.96	1.03	0.80	1.00
Triglycerides in Large VLDL	1309	0.990	0.935	1.047	7.2E-01	9.0E-01	2159	0.98	0.94	1.02	3.2E-01	7.7E-01	3107	1.00	0.96	1.04	0.94	1.00
Concentration of Medium VLDL Particles	1309	0.983	0.930	1.038	5.3E-01	9.0E-01	2159	0.98	0.94	1.03	4.3E-01	8.5E-01	3107	1.00	0.97	1.04	0.90	1.00
Total Lipids in Medium VLDL	1309	0.982	0.930	1.038	5.3E-01	9.0E-01	2159	0.98	0.94	1.03	4.2E-01	8.4E-01	3107	1.00	0.97	1.04	0.91	1.00
Phospholipids in Medium VLDL	1309	0.987	0.936	1.042	6.4E-01	9.0E-01	2159	0.98	0.94	1.03	4.7E-01	8.9E-01	3107	1.01	0.97	1.04	0.75	1.00
Cholesteryl Esters in Medium VLDL	1309	1.012	0.954	1.073	6.9E-01	9.0E-01	2159	1.00	0.96	1.05	9.0E-01	1.0E+00	3107	1.02	0.98	1.07	0.24	1.00
Free Cholesterol in Medium VLDL	1309	0.989	0.936	1.045	7.0E-01	9.0E-01	2159	0.99	0.95	1.03	5.6E-01	9.1E-01	3107	1.01	0.97	1.05	0.61	1.00
Triglycerides in Medium VLDL	1309	0.983	0.929	1.040	5.6E-01	9.0E-01	2159	0.98	0.94	1.02	3.0E-01	7.4E-01	3107	1.00	0.96	1.03	0.87	1.00
Total Lipids in Small VLDL	1309	0.986	0.933	1.043	6.3E-01	9.0E-01	2159	0.97	0.93	1.02	2.5E-01	6.8E-01	3107	0.99	0.96	1.03	0.70	1.00
Phospholipids in Small VLDL	1309	0.986	0.933	1.042	6.1E-01	9.0E-01	2159	0.98	0.94	1.03	4.3E-01	8.5E-01	3107	1.00	0.96	1.04	0.97	1.00
Cholesterol in Small VLDL	1309	0.987	0.934	1.043	6.5E-01	9.0E-01	2159	0.98	0.94	1.03	4.2E-01	8.4E-01	3107	1.00	0.97	1.04	0.94	1.00
Cholesteryl Esters in Small VLDL	1309	0.988	0.936	1.044	6.8E-01	9.0E-01	2159	0.98	0.94	1.02	3.3E-01	7.7E-01	3107	1.00	0.96	1.04	0.92	1.00
Free Cholesterol in Small VLDL	1309	0.986	0.933	1.042	6.1E-01	9.0E-01	2159	0.99	0.95	1.03	6.1E-01	9.3E-01	3107	1.01	0.97	1.05	0.68	1.00
Concentration of Very Small VLDL Particles	1309	0.983	0.930	1.039	5.5E-01	9.0E-01	2159	0.98	0.93	1.02	2.6E-01	6.9E-01	3107	0.99	0.96	1.03	0.64	1.00
Total Lipids in Very Small VLDL	1309	0.984	0.931	1.041	5.8E-01	9.0E-01	2159	0.98	0.93	1.02	2.5E-01	6.9E-01	3107	0.99	0.95	1.03	0.59	1.00
Cholesterol in Very Small VLDL	1309	0.983	0.930	1.040	5.6E-01	9.0E-01	2159	0.99	0.95	1.03	6.2E-01	9.3E-01	3107	1.00	0.97	1.04	0.94	1.00
Cholesteryl Esters in Very Small VLDL	1309	0.982	0.928	1.038	5.2E-01	9.0E-01	2159	0.99	0.95	1.04	7.6E-01	9.7E-01	3107	1.01	0.97	1.04	0.79	1.00
Free Cholesterol in Very Small VLDL	1309	0.989	0.935	1.045	6.8E-01	9.0E-01	2159	0.98	0.94	1.02	3.6E-01	8.1E-01	3107	0.99	0.96	1.03	0.73	1.00
Triglycerides in Very Small VLDL	1309	0.988	0.932	1.047	6.9E-01	9.0E-01	2159	0.95	0.91	1.00	4.2E-02	3.5E-01	3107	0.98	0.94	1.01	0.23	1.00
Phospholipids in Small LDL	1309	0.990	0.935	1.048	7.2E-01	9.0E-01	2159	1.00	0.96	1.05	8.7E-01	1.0E+00	3107	1.01	0.97	1.05	0.65	1.00

Appendix 5.5 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1] after additional adjustment for diabetes, stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Free Cholesterol in Small LDL	1309	0.987	0.937	1.040	6.3E-01	9.0E-01	2159	1.02	0.97	1.07	5.4E-01	9.1E-01	3107	1.02	0.98	1.07	0.27	1.00
Triglycerides in Small LDL	1309	0.979	0.923	1.038	4.8E-01	9.0E-01	2159	0.95	0.91	1.00	3.5E-02	3.5E-01	3107	0.98	0.94	1.01	0.21	1.00
Total Lipids in Very Large HDL	1309	1.015	0.958	1.074	6.2E-01	9.0E-01	2159	1.00	0.95	1.04	8.6E-01	1.0E+00	3107	1.01	0.97	1.04	0.77	1.00
Cholesterol in Very Large HDL	1309	1.015	0.959	1.075	6.0E-01	9.0E-01	2159	1.01	0.96	1.05	7.3E-01	9.7E-01	3107	1.01	0.97	1.05	0.54	1.00
Triglycerides in Very Large HDL	1309	0.985	0.930	1.044	6.1E-01	9.0E-01	2159	0.95	0.91	0.99	2.8E-02	3.5E-01	3107	0.98	0.94	1.01	0.21	1.00
Cholesterol in Large HDL	1309	0.986	0.931	1.044	6.3E-01	9.0E-01	2159	1.02	0.97	1.07	3.7E-01	8.1E-01	3107	1.01	0.97	1.05	0.54	1.00
Cholesteryl Esters in Large HDL	1309	0.987	0.932	1.045	6.5E-01	9.0E-01	2159	1.03	0.98	1.08	2.8E-01	7.1E-01	3107	1.02	0.98	1.06	0.44	1.00
Free Cholesterol in Large HDL	1309	0.989	0.934	1.048	7.0E-01	9.0E-01	2159	1.01	0.96	1.06	6.4E-01	9.4E-01	3107	1.00	0.96	1.04	0.89	1.00
Phospholipids to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	1269	1.017	0.961	1.076	5.5E-01	9.0E-01	2084	0.99	0.95	1.04	7.5E-01	9.7E-01	3014	0.97	0.94	1.00	0.05	0.98
Phospholipids to Total Lipids in Very Large VLDL percentage	1306	0.985	0.941	1.032	5.3E-01	9.0E-01	2150	0.97	0.94	1.01	1.2E-01	5.4E-01	3092	1.00	0.96	1.03	0.80	1.00
Phospholipids to Total Lipids in Large VLDL percentage	1309	1.016	0.960	1.074	5.9E-01	9.0E-01	2159	1.01	0.96	1.05	7.9E-01	9.7E-01	3107	1.00	0.97	1.03	0.90	1.00
Cholesteryl Esters to Total Lipids in Large VLDL percentage	1309	1.017	0.960	1.076	5.7E-01	9.0E-01	2159	1.02	0.97	1.07	4.1E-01	8.4E-01	3107	1.01	0.97	1.05	0.61	1.00
Free Cholesterol to Total Lipids in Large VLDL percentage	1309	0.982	0.933	1.033	4.8E-01	9.0E-01	2159	1.00	0.96	1.04	9.6E-01	1.0E+00	3107	0.99	0.96	1.03	0.61	1.00
Phospholipids to Total Lipids in Medium VLDL percentage	1309	1.012	0.951	1.078	7.0E-01	9.0E-01	2159	1.00	0.96	1.05	9.3E-01	1.0E+00	3107	1.02	0.98	1.07	0.37	1.00
Cholesterol to Total Lipids in Medium VLDL percentage	1309	1.017	0.959	1.079	5.7E-01	9.0E-01	2159	1.02	0.97	1.07	4.4E-01	8.6E-01	3107	1.03	0.99	1.07	0.16	1.00
Free Cholesterol to Total Lipids in Medium VLDL percentage	1309	1.016	0.958	1.077	5.9E-01	9.0E-01	2159	1.01	0.96	1.06	6.8E-01	9.5E-01	3107	1.02	0.98	1.06	0.24	1.00

Appendix 5.5 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1] after additional adjustment for diabetes, stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Triglycerides to Total Lipids in Small VLDL percentage	1309	1.011	0.954	1.072	7.1E-01	9.0E-01	2159	0.97	0.93	1.02	2.6E-01	6.9E-01	3107	0.99	0.95	1.03	0.50	1.00
Cholesteryl Esters to Total Lipids in Very Small VLDL percentage	1309	0.988	0.931	1.049	6.9E-01	9.0E-01	2159	1.04	0.99	1.09	1.4E-01	5.4E-01	3107	1.03	0.99	1.08	0.11	1.00
Free Cholesterol to Total Lipids in Very Small VLDL percentage	1309	1.022	0.963	1.084	4.7E-01	9.0E-01	2159	1.03	0.98	1.07	2.8E-01	7.1E-01	3107	1.02	0.98	1.06	0.30	1.00
Triglycerides to Total Lipids in IDL percentage	1309	1.022	0.963	1.085	4.7E-01	9.0E-01	2159	0.96	0.91	1.00	6.4E-02	4.1E-01	3107	0.97	0.93	1.00	0.08	1.00
Phospholipids to Total Lipids in Large LDL percentage	1309	1.016	0.962	1.072	5.8E-01	9.0E-01	2159	0.99	0.94	1.03	5.3E-01	9.1E-01	3107	1.01	0.98	1.05	0.53	1.00
Triglycerides to Total Lipids in Large LDL percentage	1309	1.020	0.960	1.083	5.2E-01	9.0E-01	2159	0.95	0.91	1.00	3.7E-02	3.5E-01	3107	0.96	0.92	1.00	0.05	0.98
Phospholipids to Total Lipids in Medium LDL percentage	1309	1.010	0.956	1.068	7.2E-01	9.0E-01	2159	1.05	1.01	1.10	1.6E-02	3.5E-01	3107	1.01	0.98	1.05	0.51	1.00
Triglycerides to Total Lipids in Medium LDL percentage	1309	1.016	0.958	1.078	6.0E-01	9.0E-01	2159	0.96	0.91	1.00	7.0E-02	4.2E-01	3107	0.96	0.93	1.00	0.06	0.98
Cholesterol to Total Lipids in Small LDL percentage	1309	0.992	0.962	1.023	5.9E-01	9.0E-01	2159	1.00	0.94	1.07	9.5E-01	1.0E+00	3107	1.07	0.92	1.24	0.35	1.00
Free Cholesterol to Total Lipids in Small LDL percentage	1309	1.017	0.943	1.096	6.6E-01	9.0E-01	2159	1.06	0.99	1.14	1.0E-01	5.2E-01	3107	1.05	0.99	1.12	0.10	1.00
Phospholipids to Total Lipids in Very Large HDL percentage	1309	0.989	0.936	1.044	6.8E-01	9.0E-01	2159	1.02	0.97	1.07	4.2E-01	8.4E-01	3104	1.01	0.97	1.05	0.56	1.00
Triglycerides to Total Lipids in Very Large HDL percentage	1309	0.980	0.924	1.039	4.9E-01	9.0E-01	2159	0.97	0.93	1.02	1.9E-01	6.2E-01	3104	0.98	0.94	1.01	0.22	1.00
Phospholipids to Total Lipids in Large HDL percentage	1309	1.026	0.934	1.127	5.9E-01	9.0E-01	2159	1.01	0.95	1.07	7.9E-01	9.7E-01	3107	0.99	0.96	1.03	0.71	1.00
Cholesterol to Total Lipids in Large HDL percentage	1309	1.016	0.952	1.085	6.3E-01	9.0E-01	2159	1.06	0.99	1.13	9.6E-02	5.0E-01	3107	1.03	0.98	1.09	0.18	1.00

Appendix 5.5 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1] after additional adjustment for diabetes, stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Cholesteryl Esters to Total Lipids in Large HDL percentage	1309	1.016	0.952	1.084	6.3E-01	9.0E-01	2159	1.07	1.00	1.14	5.1E-02	3.5E-01	3107	1.04	0.99	1.09	0.12	1.00
Free Cholesterol to Total Lipids in Large HDL percentage	1309	1.019	0.953	1.090	5.8E-01	9.0E-01	2159	1.02	0.97	1.07	5.2E-01	9.1E-01	3107	1.01	0.97	1.05	0.78	1.00
Cholesteryl Esters to Total Lipids in Medium HDL percentage	1309	0.986	0.932	1.042	6.1E-01	9.0E-01	2159	1.05	1.00	1.11	7.2E-02	4.2E-01	3107	1.03	0.99	1.08	0.14	1.00
Cholesteryl Esters to Total Lipids in Small HDL percentage	1309	0.987	0.925	1.054	7.0E-01	9.0E-01	2159	1.06	1.00	1.12	3.4E-02	3.5E-01	3107	1.03	0.98	1.07	0.21	1.00
Triglycerides to Total Lipids in Small HDL percentage	1309	1.014	0.956	1.075	6.5E-01	9.0E-01	2159	0.95	0.91	1.00	4.2E-02	3.5E-01	3107	0.98	0.94	1.02	0.27	1.00
Free Cholesterol in Chylomicrons and Extremely Large VLDL	1309	0.991	0.936	1.049	7.5E-01	9.3E-01	2159	0.96	0.92	1.00	4.4E-02	3.5E-01	3107	0.98	0.95	1.02	0.30	1.00
Average Diameter for HDL Particles	1309	0.991	0.936	1.050	7.7E-01	9.4E-01	2159	1.00	0.95	1.04	9.2E-01	1.0E+00	3107	1.00	0.96	1.03	0.83	1.00
Triglycerides in Very Large VLDL	1309	1.008	0.953	1.066	7.8E-01	9.4E-01	2159	0.96	0.92	1.01	8.6E-02	4.7E-01	3107	1.00	0.96	1.04	0.96	1.00
Concentration of Small VLDL Particles	1309	0.991	0.937	1.049	7.6E-01	9.4E-01	2159	0.97	0.93	1.02	2.4E-01	6.8E-01	3107	0.99	0.96	1.03	0.70	1.00
Phospholipids in Very Small VLDL	1309	0.992	0.938	1.049	7.7E-01	9.4E-01	2159	0.97	0.93	1.02	2.3E-01	6.8E-01	3107	0.99	0.95	1.02	0.48	1.00
Cholesterol to Total Lipids in Very Small VLDL percentage	1309	0.991	0.934	1.052	7.8E-01	9.4E-01	2159	1.04	0.99	1.09	1.4E-01	5.5E-01	3107	1.03	0.99	1.07	0.11	1.00
Cholesterol to Total Lipids in Very Large HDL percentage	1309	0.993	0.947	1.041	7.7E-01	9.4E-01	2159	1.04	0.99	1.09	1.3E-01	5.4E-01	3104	1.01	0.98	1.05	0.48	1.00
Cholesterol to Total Lipids in Small HDL percentage	1309	0.991	0.935	1.051	7.7E-01	9.4E-01	2159	1.05	1.00	1.10	3.2E-02	3.5E-01	3107	1.02	0.99	1.06	0.22	1.00
Linoleic Acid to Total Fatty Acids percentage	1307	1.008	0.948	1.071	8.0E-01	9.4E-01	2159	0.99	0.94	1.04	6.6E-01	9.4E-01	3104	1.02	0.98	1.06	0.32	1.00
Acetoacetate	1294	0.993	0.938	1.050	7.9E-01	9.4E-01	2146	0.99	0.95	1.04	8.0E-01	9.7E-01	3077	1.03	0.99	1.07	0.10	1.00
Total Lipids in Very Large VLDL	1309	0.993	0.939	1.050	8.0E-01	9.4E-01	2159	0.97	0.93	1.01	1.6E-01	5.9E-01	3107	0.99	0.96	1.03	0.79	1.00

Appendix 5.5 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1] after additional adjustment for diabetes, stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Cholesteryl Esters to Total Lipids in Large LDL percentage	1309	1.020	0.881	1.180	7.9E-01	9.4E-01	2159	1.17	0.95	1.43	1.3E-01	5.4E-01	3107	1.08	0.92	1.28	0.36	1.00
Triglycerides in Small VLDL	1309	0.993	0.937	1.052	8.1E-01	9.5E-01	2159	0.97	0.92	1.01	1.3E-01	5.4E-01	3107	0.99	0.95	1.03	0.52	1.00
Triglycerides to Total Lipids in Large HDL percentage	1309	0.993	0.935	1.054	8.1E-01	9.5E-01	2159	0.95	0.91	1.00	4.9E-02	3.5E-01	3107	0.98	0.94	1.02	0.28	1.00
Alanine	1307	0.994	0.940	1.052	8.4E-01	9.6E-01	2159	0.99	0.95	1.04	8.0E-01	9.7E-01	3105	0.96	0.93	1.00	0.03	0.98
Concentration of Chylomicrons and Extremely Large VLDL Particles	1309	1.006	0.951	1.065	8.3E-01	9.6E-01	2159	0.96	0.92	1.01	1.0E-01	5.2E-01	3107	0.99	0.95	1.02	0.42	1.00
Concentration of Very Large HDL Particles	1309	1.006	0.951	1.065	8.2E-01	9.6E-01	2159	0.99	0.95	1.04	7.4E-01	9.7E-01	3107	1.00	0.96	1.04	0.93	1.00
Cholesterol to Total Lipids in Large LDL percentage	1309	1.010	0.922	1.107	8.3E-01	9.6E-01	2159	1.28	0.99	1.67	6.0E-02	3.9E-01	3107	1.16	0.94	1.44	0.16	1.00
Free Cholesterol to Total Lipids in Small HDL percentage	1309	1.006	0.952	1.063	8.3E-01	9.6E-01	2159	1.00	0.96	1.04	9.7E-01	1.0E+00	3107	0.99	0.96	1.03	0.77	1.00
Free Cholesterol in Very Large VLDL	1309	0.995	0.941	1.051	8.5E-01	9.6E-01	2159	0.97	0.93	1.01	1.8E-01	6.1E-01	3107	0.99	0.96	1.03	0.71	1.00
Cholesteryl Esters in Large VLDL	1309	0.995	0.942	1.051	8.6E-01	9.7E-01	2159	0.98	0.94	1.02	3.8E-01	8.3E-01	3107	1.00	0.96	1.04	0.98	1.00
Cholesterol in Medium VLDL	1309	0.995	0.941	1.052	8.6E-01	9.7E-01	2159	1.00	0.95	1.04	8.4E-01	9.9E-01	3107	1.02	0.98	1.06	0.35	1.00
Total Lipids in Chylomicrons and Extremely Large VLDL	1309	1.005	0.950	1.063	8.7E-01	9.7E-01	2159	0.96	0.92	1.00	6.0E-02	3.9E-01	3107	0.99	0.96	1.03	0.65	1.00
Phospholipids in Very Large HDL	1309	1.005	0.948	1.066	8.6E-01	9.7E-01	2159	1.02	0.97	1.06	5.0E-01	8.9E-01	3107	1.01	0.97	1.05	0.72	1.00
Free Cholesterol to Total Lipids in Medium LDL percentage	1309	1.006	0.939	1.078	8.7E-01	9.7E-01	2159	1.11	1.02	1.21	1.7E-02	3.5E-01	3107	1.05	0.99	1.13	0.12	1.00
Free Cholesterol to Total Lipids in Small VLDL percentage	1309	0.995	0.938	1.056	8.8E-01	9.7E-01	2159	1.03	0.98	1.08	2.5E-01	6.8E-01	3107	1.03	0.99	1.07	0.11	1.00
Phospholipids to Total Lipids in Small VLDL percentage	1309	0.996	0.939	1.056	8.9E-01	9.7E-01	2159	1.03	0.98	1.08	2.0E-01	6.3E-01	3107	1.03	0.99	1.07	0.21	1.00

Appendix 5.5 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1] after additional adjustment for diabetes, stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Phospholipids to Total Lipids in Small HDL percentage	1309	0.996	0.940	1.055	8.9E-01	9.7E-01	2159	0.98	0.94	1.02	3.4E-01	7.7E-01	3107	0.99	0.95	1.03	0.58	1.00
Phospholipids in Very Large VLDL	1309	1.004	0.950	1.060	9.0E-01	9.8E-01	2159	0.97	0.93	1.01	1.8E-01	6.1E-01	3107	0.99	0.96	1.03	0.68	1.00
Cholesterol to Total Lipids in Large VLDL percentage	1309	1.004	0.950	1.061	9.0E-01	9.8E-01	2159	1.01	0.97	1.06	5.6E-01	9.1E-01	3107	1.00	0.97	1.04	0.89	1.00
Cholesteryl Esters in Very Large HDL	1309	1.003	0.947	1.062	9.2E-01	9.8E-01	2159	1.01	0.97	1.06	5.6E-01	9.1E-01	3107	1.02	0.98	1.06	0.40	1.00
Cholesteryl Esters to Total Lipids in Very Large VLDL percentage	1306	1.003	0.947	1.062	9.2E-01	9.8E-01	2150	1.03	0.99	1.08	1.4E-01	5.4E-01	3092	1.01	0.97	1.05	0.54	1.00
Triglycerides to Total Lipids in Very Small VLDL percentage	1309	1.003	0.945	1.064	9.2E-01	9.8E-01	2159	0.96	0.92	1.01	1.3E-01	5.4E-01	3107	0.98	0.94	1.02	0.27	1.00
Cholesteryl Esters to Total Lipids in Medium LDL percentage	1309	1.003	0.939	1.073	9.2E-01	9.8E-01	2159	0.99	0.97	1.01	5.1E-01	9.0E-01	3107	1.01	0.94	1.08	0.82	1.00
Free Cholesterol to Total Lipids in Very Large HDL percentage	1309	1.003	0.948	1.062	9.1E-01	9.8E-01	2159	1.01	0.96	1.06	6.9E-01	9.5E-01	3104	0.99	0.95	1.03	0.51	1.00
Triglycerides to Total Lipids in Small LDL percentage	1309	1.003	0.945	1.064	9.3E-01	9.8E-01	2159	0.95	0.91	1.00	3.7E-02	3.5E-01	3107	0.97	0.93	1.01	0.11	1.00
Phenylalanine	1308	0.998	0.943	1.055	9.4E-01	9.8E-01	2158	0.98	0.94	1.02	3.0E-01	7.4E-01	3107	1.01	0.97	1.05	0.63	1.00
Phospholipids in Large VLDL	1309	1.002	0.949	1.058	9.4E-01	9.8E-01	2159	0.99	0.95	1.03	5.8E-01	9.2E-01	3107	1.00	0.96	1.03	0.89	1.00
Free Cholesterol to Total Lipids in Very Large VLDL percentage	1306	1.001	0.948	1.057	9.7E-01	9.8E-01	2150	1.01	0.97	1.06	5.9E-01	9.2E-01	3092	0.98	0.95	1.02	0.41	1.00
Triglycerides to Total Lipids in Large VLDL percentage	1309	1.002	0.947	1.060	9.5E-01	9.8E-01	2159	1.01	0.96	1.06	6.8E-01	9.5E-01	3107	1.01	0.97	1.05	0.68	1.00
Triglycerides to Total Lipids in Medium VLDL percentage	1309	0.999	0.943	1.057	9.6E-01	9.8E-01	2159	0.98	0.94	1.03	4.5E-01	8.6E-01	3107	0.99	0.95	1.03	0.51	1.00
Cholesterol to Total Lipids in Small VLDL percentage	1309	0.999	0.942	1.058	9.6E-01	9.8E-01	2159	1.02	0.97	1.06	5.0E-01	8.9E-01	3107	1.02	0.98	1.06	0.27	1.00

Appendix 5.5 Hazard ratios and 95% CIs of prostate cancer risk associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1] after additional adjustment for diabetes, stratified by time since blood collection

	< 3 years						3-7 years						> 7 years					
Cholesteryl Esters to Total Lipids in Small VLDL percentage	1309	1.002	0.947	1.060	9.6E-01	9.8E-01	2159	1.01	0.96	1.05	8.0E-01	9.7E-01	3107	1.01	0.98	1.05	0.50	1.00
Free Cholesterol to Total Lipids in IDL percentage	1309	1.001	0.945	1.060	9.7E-01	9.8E-01	2159	1.02	0.97	1.06	4.8E-01	8.9E-01	3107	1.00	0.97	1.04	0.86	1.00
Free Cholesterol to Total Lipids in Large LDL percentage	1309	0.999	0.947	1.053	9.7E-01	9.8E-01	2159	1.07	0.97	1.18	1.7E-01	5.9E-01	3107	1.05	0.97	1.13	0.26	1.00
Cholesterol to Total Lipids in Medium LDL percentage	1309	1.003	0.922	1.092	9.4E-01	9.8E-01	2159	1.05	0.85	1.30	6.5E-01	9.4E-01	3107	1.18	0.97	1.43	0.10	1.00
Triglycerides to Total Lipids in Medium HDL percentage	1309	1.002	0.945	1.063	9.4E-01	9.8E-01	2159	0.96	0.91	1.00	4.7E-02	3.5E-01	3107	0.98	0.94	1.02	0.33	1.00
Cholesterol to Total Lipids in Very Large VLDL percentage	1306	1.001	0.945	1.060	9.8E-01	9.9E-01	2150	1.03	0.99	1.08	1.8E-01	6.1E-01	3092	1.00	0.97	1.04	0.81	1.00
Average Diameter for LDL Particles	1309	1.000	0.945	1.059	9.9E-01	1.0E+00	2159	1.05	1.01	1.10	2.4E-02	3.5E-01	3107	1.01	0.97	1.05	0.59	1.00
Docosahexaenoic Acid	1307	1.000	0.944	1.059	1.0E+00	1.0E+00	2159	1.06	1.01	1.11	1.5E-02	3.5E-01	3104	1.04	1.00	1.08	0.05	0.98

Abbreviations: HR, hazard ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit; LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol; IDL, intermediate-density lipoprotein cholesterol; VLDL, very low-density lipoprotein cholesterol; BCAA, Total concentration of branched-chain amino acids (Leucine + Isoleucine + Valine).

1 Hazard ratios (95% confidence interval) for prostate cancer associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features by time since blood collection, after additional adjustment for diabetes. Analyses were stratified by region and Townsend Deprivation Index with age as the underlying time variable. Adjusted for diabetes, fasting status, body mass index, height, alcohol intake, family history of prostate cancer, ethnicity, level of education, smoking status, and weekly physical activity.

2 Padj < 0.05 are statistically significant after allowing for multiple testing using a false discovery rate controlling procedure at α = 0.05 (Benjamini-Hochberg). Padj values in the < 3 years group are sorted in ascending order.

Appendix 5.6 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1]

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Glycoprotein Acetyls	**596**	**1.25**	**1.14**	**1.36**	**4.9E-07**	**1.2E-04**
Albumin	**596**	**0.86**	**0.79**	**0.93**	**1.1E-04**	**1.3E-02**
Citrate	**596**	**1.18**	**1.08**	**1.28**	**1.6E-04**	**1.3E-02**
Glucose	595	1.14	1.05	1.23	1.0E-03	6.3E-02
Cholesteryl Esters to Total Lipids in Small HDL percentage	596	0.87	0.79	0.95	1.5E-03	7.6E-02
Phospholipids to Total Lipids in Small HDL percentage	596	1.14	1.04	1.24	3.1E-03	1.3E-01
Histidine	596	0.90	0.83	0.97	7.3E-03	2.1E-01
Triglycerides in IDL	596	1.12	1.03	1.22	8.0E-03	2.1E-01
Cholesterol to Total Lipids in Small HDL percentage	596	0.89	0.82	0.97	8.0E-03	2.1E-01
Phospholipids in Large HDL	596	1.13	1.03	1.24	9.3E-03	2.1E-01
Triglycerides in Large HDL	596	1.12	1.03	1.23	9.9E-03	2.1E-01
Phospholipids in HDL	596	1.12	1.03	1.23	1.0E-02	2.1E-01
Phospholipids in Very Small VLDL	596	1.11	1.02	1.21	1.2E-02	2.2E-01
Phospholipids in Medium HDL	596	1.12	1.02	1.22	1.6E-02	2.4E-01
Total Lipids in Very Small VLDL	596	1.10	1.02	1.20	1.7E-02	2.4E-01
Average Diameter for HDL Particles	596	1.11	1.02	1.20	1.8E-02	2.4E-01
Pyruvate	595	1.11	1.02	1.21	1.9E-02	2.4E-01
Saturated Fatty Acids to Total Fatty Acids percentage	596	1.11	1.02	1.21	1.9E-02	2.4E-01
Linoleic Acid to Total Fatty Acids percentage	596	0.90	0.83	0.98	2.0E-02	2.4E-01
Total Lipids in Large HDL	596	1.11	1.02	1.22	2.0E-02	2.4E-01
Free Cholesterol in HDL	596	1.11	1.02	1.21	2.1E-02	2.4E-01
Total Lipids in HDL	596	1.11	1.02	1.21	2.2E-02	2.4E-01
Triglycerides to Total Lipids in Medium LDL percentage	596	1.10	1.01	1.20	2.3E-02	2.4E-01
Triglycerides in Very Small VLDL	596	1.11	1.01	1.21	2.3E-02	2.4E-01
Triglycerides in Medium HDL	596	1.11	1.01	1.21	2.4E-02	2.4E-01
Triglycerides in HDL	596	1.10	1.01	1.20	2.7E-02	2.4E-01
Free Cholesterol in Large HDL	596	1.12	1.01	1.24	2.7E-02	2.4E-01
Triglycerides in Very Large HDL	596	1.10	1.01	1.20	2.8E-02	2.4E-01
Concentration of Large HDL Particles	596	1.11	1.01	1.22	3.0E-02	2.4E-01
Free Cholesterol in Very Small VLDL	596	1.09	1.01	1.19	3.1E-02	2.4E-01
Concentration of Very Small VLDL Particles	596	1.09	1.01	1.19	3.1E-02	2.4E-01
Triglycerides in Large LDL	596	1.10	1.01	1.20	3.2E-02	2.4E-01
Total Lipids in Medium HDL	596	1.10	1.01	1.20	3.3E-02	2.4E-01

Appendix 5.6 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1]

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Omega-6 Fatty Acids to Total Fatty Acids percentage	596	0.91	0.84	0.99	3.3E-02	2.4E-01
Phosphoglycerides	596	1.09	1.01	1.19	3.8E-02	2.7E-01
Polyunsaturated Fatty Acids to Total Fatty Acids percentage	596	0.91	0.84	1.00	3.9E-02	2.7E-01
Free Cholesterol in Medium HDL	596	1.10	1.01	1.21	3.9E-02	2.7E-01
Triglycerides to Total Lipids in Large LDL percentage	596	1.09	1.00	1.19	4.2E-02	2.7E-01
Phosphatidylcholines	596	1.09	1.00	1.19	4.3E-02	2.7E-01
Saturated Fatty Acids	596	1.09	1.00	1.19	4.3E-02	2.7E-01
Sphingomyelins	596	1.09	1.00	1.18	4.7E-02	2.7E-01
Glutamine	595	0.92	0.85	1.00	4.9E-02	2.7E-01
Free Cholesterol to Total Lipids in Very Large HDL percentage	596	0.93	0.87	1.00	4.9E-02	2.7E-01
Apolipoprotein A1	596	1.09	1.00	1.19	4.9E-02	2.7E-01
Phospholipids to Total Lipids in Large LDL percentage	596	0.92	0.85	1.00	5.1E-02	2.7E-01
Cholesterol to Total Lipids in Very Large HDL percentage	596	0.96	0.92	1.00	5.1E-02	2.7E-01
Total Cholines	596	1.09	1.00	1.18	5.8E-02	3.1E-01
Creatinine	584	0.92	0.85	1.00	6.0E-02	3.1E-01
Concentration of Medium HDL Particles	596	1.09	1.00	1.19	6.1E-02	3.1E-01
Cholesterol in Large HDL	596	1.09	1.00	1.20	6.4E-02	3.1E-01
Cholesterol in Very Small VLDL	596	1.08	1.00	1.17	6.4E-02	3.1E-01
Triglycerides in LDL	596	1.08	0.99	1.18	7.0E-02	3.3E-01
Acetone	596	0.93	0.85	1.01	7.5E-02	3.5E-01
Concentration of Very Large HDL Particles	596	1.08	0.99	1.17	8.6E-02	3.8E-01
Total Phospholipids in Lipoprotein Particles	596	1.08	0.99	1.17	8.6E-02	3.8E-01
Cholesteryl Esters in Large HDL	596	1.09	0.99	1.19	8.6E-02	3.8E-01
Tyrosine	596	1.08	0.99	1.17	8.7E-02	3.8E-01
Free Cholesterol to Total Lipids in Small HDL percentage	596	1.07	0.99	1.16	8.7E-02	3.8E-01
Phospholipids in Very Large HDL	596	1.09	0.99	1.21	9.2E-02	3.9E-01
Cholesteryl Esters to Total Lipids in Medium HDL percentage	596	0.97	0.93	1.01	9.5E-02	3.9E-01
Cholesteryl Esters in Very Small VLDL	596	1.07	0.99	1.16	9.6E-02	3.9E-01
Cholesterol to Total Lipids in IDL percentage	596	0.94	0.86	1.01	9.8E-02	3.9E-01
Total Fatty Acids	596	1.07	0.99	1.17	1.0E-01	4.0E-01
Total Lipids in Very Large HDL	596	1.07	0.99	1.17	1.1E-01	4.2E-01
Triglycerides to Total Lipids in Small LDL percentage	596	1.07	0.98	1.17	1.1E-01	4.4E-01
Monounsaturated Fatty Acids	596	1.07	0.98	1.17	1.2E-01	4.4E-01

Appendix 5.6 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1]

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Polyunsaturated Fatty Acids to Monounsaturated Fatty Acids ratio	596	0.93	0.86	1.02	1.2E-01	4.5E-01
HDL Cholesterol	596	1.07	0.98	1.17	1.3E-01	4.6E-01
Triglycerides to Total Lipids in IDL percentage	596	1.07	0.98	1.16	1.3E-01	4.6E-01
Cholesterol in Medium HDL	596	1.07	0.98	1.17	1.4E-01	4.8E-01
Phospholipids to Total Lipids in Small LDL percentage	596	1.08	0.98	1.20	1.4E-01	4.8E-01
Free Cholesterol to Total Lipids in Medium HDL percentage	596	1.07	0.98	1.18	1.5E-01	5.0E-01
Cholesteryl Esters to Total Lipids in IDL percentage	596	0.94	0.87	1.02	1.5E-01	5.0E-01
Phospholipids to Total Lipids in Very Small VLDL percentage	596	1.06	0.98	1.16	1.6E-01	5.4E-01
Phospholipids to Total Lipids in Medium HDL percentage	596	1.08	0.97	1.20	1.7E-01	5.8E-01
Triglycerides in Medium LDL	596	1.06	0.97	1.16	1.8E-01	5.8E-01
Cholesteryl Esters in Very Large HDL	596	1.07	0.97	1.19	1.8E-01	5.8E-01
Cholesteryl Esters in Medium HDL	596	1.06	0.97	1.16	1.9E-01	6.0E-01
Free Cholesterol in Small HDL	596	1.06	0.97	1.15	1.9E-01	6.0E-01
Free Cholesterol to Total Lipids in Large HDL percentage	596	1.09	0.96	1.25	1.9E-01	6.0E-01
Phospholipids to Total Lipids in Very Large HDL percentage	596	1.08	0.96	1.22	2.0E-01	6.2E-01
Cholesteryl Esters in HDL	596	1.06	0.97	1.16	2.1E-01	6.4E-01
Acetoacetate	592	1.06	0.97	1.15	2.1E-01	6.4E-01
Average Diameter for VLDL Particles	596	0.95	0.87	1.03	2.2E-01	6.4E-01
Alanine	596	1.05	0.97	1.14	2.2E-01	6.4E-01
Phospholipids in Small HDL	596	1.06	0.97	1.15	2.2E-01	6.4E-01
Free Cholesterol to Total Lipids in IDL percentage	596	0.95	0.88	1.03	2.3E-01	6.5E-01
Cholesterol in Very Large HDL	596	1.05	0.97	1.15	2.4E-01	6.8E-01
Concentration of HDL Particles	596	1.05	0.97	1.15	2.5E-01	6.8E-01
Cholesterol to Total Lipids in Medium HDL percentage	596	0.97	0.93	1.02	2.5E-01	6.8E-01
Average Diameter for LDL Particles	596	1.05	0.97	1.14	2.5E-01	6.8E-01
Triglycerides in Small HDL	596	1.05	0.96	1.15	2.6E-01	7.0E-01
Total Concentration of Lipoprotein Particles	596	1.05	0.96	1.15	2.7E-01	7.2E-01
Triglycerides in Small VLDL	596	1.05	0.96	1.15	2.7E-01	7.3E-01
Phospholipids in IDL	596	1.05	0.97	1.13	2.8E-01	7.3E-01
Degree of Unsaturation	596	0.96	0.88	1.04	2.9E-01	7.5E-01

Appendix 5.6 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1]

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Triglycerides to Total Lipids in Medium HDL percentage	596	1.05	0.96	1.15	2.9E-01	7.5E-01
Total Lipids in Lipoprotein Particles	596	1.04	0.96	1.13	3.0E-01	7.6E-01
Free Cholesterol in Very Large HDL	596	1.05	0.96	1.14	3.0E-01	7.6E-01
Concentration of VLDL Particles	596	1.04	0.96	1.13	3.0E-01	7.6E-01
Triglycerides in Small LDL	596	1.04	0.96	1.14	3.3E-01	8.0E-01
Monounsaturated Fatty Acids to Total Fatty Acids percentage	596	1.05	0.96	1.14	3.3E-01	8.1E-01
Total Lipids in Small VLDL	596	1.04	0.96	1.13	3.4E-01	8.2E-01
Total Lipids in IDL	596	1.04	0.96	1.13	3.6E-01	8.5E-01
Triglycerides to Total Lipids in Small HDL percentage	596	1.04	0.95	1.14	3.6E-01	8.6E-01
Concentration of Small VLDL Particles	596	1.04	0.96	1.13	3.7E-01	8.7E-01
Apolipoprotein B to Apolipoprotein A1 ratio	596	0.96	0.89	1.05	3.8E-01	8.7E-01
Free Cholesterol in Small LDL	596	0.97	0.92	1.03	3.8E-01	8.7E-01
3-Hydroxybutyrate	584	1.04	0.95	1.13	3.9E-01	8.7E-01
Omega-3 Fatty Acids	596	1.04	0.95	1.13	3.9E-01	8.7E-01
Phospholipids to Total Lipids in Large HDL percentage	596	1.06	0.92	1.23	3.9E-01	8.7E-01
Phospholipids to Total Lipids in Small VLDL percentage	596	0.96	0.89	1.05	3.9E-01	8.7E-01
Polyunsaturated Fatty Acids	596	1.04	0.95	1.13	4.0E-01	8.8E-01
Cholesteryl Esters in Small HDL	596	0.96	0.88	1.05	4.2E-01	9.2E-01
Leucine	596	0.97	0.89	1.05	4.3E-01	9.2E-01
Free Cholesterol to Total Lipids in Small VLDL percentage	596	0.97	0.89	1.05	4.3E-01	9.2E-01
Concentration of IDL Particles	596	1.03	0.95	1.12	4.4E-01	9.2E-01
Total Lipids in Small HDL	596	1.04	0.95	1.13	4.4E-01	9.2E-01
Free Cholesterol to Total Lipids in Small LDL percentage	596	0.98	0.93	1.03	4.4E-01	9.3E-01
Cholesteryl Esters in Small VLDL	596	1.03	0.95	1.12	4.5E-01	9.3E-01
Docosahexaenoic Acid to Total Fatty Acids percentage	596	0.97	0.89	1.05	4.6E-01	9.3E-01
Omega-6 Fatty Acids	596	1.03	0.95	1.12	4.7E-01	9.3E-01
Cholesterol in Small LDL	596	0.98	0.92	1.04	4.8E-01	9.3E-01
Total Triglycerides	596	1.03	0.95	1.13	4.8E-01	9.3E-01
Cholesterol in Medium LDL	596	0.98	0.92	1.04	4.8E-01	9.3E-01
Free Cholesterol in Medium LDL	596	0.98	0.92	1.04	4.8E-01	9.3E-01
Cholesteryl Esters to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	583	0.97	0.89	1.06	4.9E-01	9.3E-01
Phospholipids in Small VLDL	596	1.03	0.95	1.12	4.9E-01	9.3E-01

Appendix 5.6 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1]

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Phospholipids to Total Lipids in IDL percentage	596	1.03	0.95	1.12	4.9E-01	9.3E-01
Remnant Cholesterol (Non-HDL, Non-LDL -Cholesterol)	596	1.03	0.95	1.11	5.0E-01	9.3E-01
Phospholipids in VLDL	596	1.03	0.95	1.12	5.0E-01	9.3E-01
Cholesterol in Small VLDL	596	1.03	0.95	1.11	5.1E-01	9.3E-01
Cholesteryl Esters in VLDL	596	1.03	0.95	1.11	5.1E-01	9.3E-01
Concentration of Medium LDL Particles	596	0.97	0.90	1.06	5.1E-01	9.3E-01
Phospholipids to Total Lipids in Very Large VLDL percentage	594	0.98	0.91	1.05	5.1E-01	9.3E-01
Cholesteryl Esters in Medium LDL	596	0.98	0.92	1.04	5.1E-01	9.3E-01
Omega-6 Fatty Acids to Omega-3 Fatty Acids ratio	596	0.97	0.89	1.06	5.2E-01	9.3E-01
Phospholipids in Medium LDL	596	0.97	0.90	1.06	5.2E-01	9.3E-01
Total Lipids in Medium LDL	596	0.97	0.90	1.06	5.2E-01	9.3E-01
Cholesteryl Esters to Total Lipids in Very Small VLDL percentage	596	0.97	0.90	1.06	5.3E-01	9.3E-01
Phenylalanine	596	1.03	0.95	1.12	5.3E-01	9.3E-01
Total Free Cholesterol	596	1.03	0.95	1.11	5.3E-01	9.3E-01
Cholesterol to Total Lipids in Very Small VLDL percentage	596	0.97	0.90	1.06	5.4E-01	9.3E-01
VLDL Cholesterol	596	1.03	0.95	1.11	5.4E-01	9.3E-01
Phospholipids to Total Lipids in Medium VLDL percentage	596	1.03	0.94	1.13	5.5E-01	9.4E-01
Triglycerides to Total Lipids in Large VLDL percentage	596	0.98	0.92	1.05	5.6E-01	9.4E-01
Free Cholesterol to Total Lipids in Very Small VLDL percentage	596	0.98	0.90	1.06	5.6E-01	9.4E-01
Cholesteryl Esters in Small LDL	596	0.98	0.92	1.05	5.7E-01	9.4E-01
Free Cholesterol to Total Lipids in Large VLDL percentage	596	0.98	0.91	1.06	5.7E-01	9.4E-01

Abbreviations: SD, standard deviation; HR, hazard ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit; LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol; IDL, intermediate-density lipoprotein cholesterol; VLDL, very low-density lipoprotein cholesterol.

1 Hazard ratios (95% confidence interval) for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features. Analyses were stratified by region and Townsend Deprivation Index with age as the underlying time variable. Adjusted for fasting status, body mass index, height, alcohol intake, family history of prostate cancer, ethnicity, level of education, smoking status, and weekly physical activity.

2 Padj < 0.05 are statistically significant after allowing for multiple testing using a false discovery rate controlling procedure at $\alpha = 0.05$ (Benjamini-Hochberg). Padj are sorted in ascending order. Associations with Padj < 0.05 are bolded.

Appendix 5.7 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	≤ 5 years						> 5 years					
Metabolites	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj
Glycoprotein acetyls	**290**	**1.40**	**1.24**	**1.57**	**6.4E-08**	**1.6E-05**	**306**	**1.11**	**0.99**	**1.25**	**8.4E-02**	**1.0E+00**
Albumin	**290**	**0.79**	**0.71**	**0.88**	**2.6E-05**	**2.2E-03**	**306**	**0.93**	**0.83**	**1.04**	**1.9E-01**	**1.0E+00**
Cholesteryl Esters to Total Lipids in Small HDL percentage	290	0.78	0.69	0.87	2.0E-05	2.2E-03	306	0.98	0.86	1.13	8.2E-01	1.0E+00
Glucose	289	1.21	1.09	1.34	4.0E-04	2.3E-02	306	1.07	0.95	1.20	2.6E-01	1.0E+00
Triglycerides in IDL	290	1.24	1.10	1.40	4.7E-04	2.3E-02	306	1.01	0.90	1.15	8.2E-01	1.0E+00
Cholesterol to Total Lipids in Small HDL percentage	290	0.82	0.73	0.92	5.6E-04	2.3E-02	306	0.98	0.87	1.11	7.7E-01	1.0E+00
Triglycerides in Very Small VLDL	290	1.22	1.08	1.38	1.3E-03	4.8E-02	306	1.00	0.88	1.13	9.8E-01	1.0E+00
Triglycerides in Large HDL	290	1.23	1.08	1.39	1.6E-03	5.0E-02	306	1.03	0.91	1.17	6.3E-01	1.0E+00
Triglycerides in HDL	290	1.21	1.07	1.37	2.5E-03	5.8E-02	306	1.01	0.89	1.14	9.3E-01	1.0E+00
Triglycerides in Large LDL	290	1.20	1.07	1.35	2.7E-03	5.8E-02	306	1.00	0.89	1.13	9.9E-01	1.0E+00
Triglycerides in Medium HDL	290	1.21	1.07	1.38	3.0E-03	5.8E-02	306	1.01	0.90	1.15	8.3E-01	1.0E+00
Phospholipids to Total Lipids in Large LDL percentage	290	0.84	0.74	0.94	2.9E-03	5.8E-02	306	1.01	0.90	1.13	8.4E-01	1.0E+00
Phospholipids to Total Lipids in Small HDL percentage	290	1.20	1.07	1.36	2.5E-03	5.8E-02	306	1.07	0.95	1.21	2.4E-01	1.0E+00
Triglycerides in Very Large HDL	290	1.20	1.06	1.35	4.1E-03	7.3E-02	306	1.01	0.90	1.14	8.3E-01	1.0E+00
Saturated Fatty Acids to Total Fatty Acids percentage	290	1.19	1.05	1.34	4.9E-03	7.8E-02	306	1.03	0.91	1.16	6.1E-01	1.0E+00

Appendix 5.7 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	≤ 5 years						> 5 years					
Triglycerides to Total Lipids in Medium LDL percentage	290	1.18	1.05	1.32	5.0E-03	7.8E-02	306	1.03	0.91	1.16	6.8E-01	1.0E+00
Triglycerides in LDL	290	1.18	1.05	1.33	7.3E-03	9.0E-02	306	0.99	0.88	1.12	8.8E-01	1.0E+00
Saturated Fatty Acids	290	1.17	1.04	1.32	8.3E-03	9.0E-02	306	1.01	0.90	1.14	8.4E-01	1.0E+00
Omega-6 Fatty Acids to Total Fatty Acids percentage	290	0.85	0.76	0.96	7.6E-03	9.0E-02	306	0.98	0.87	1.11	7.5E-01	1.0E+00
Polyunsaturated Fatty Acids to Total Fatty Acids percentage	290	0.85	0.75	0.95	6.3E-03	9.0E-02	306	0.99	0.87	1.12	8.6E-01	1.0E+00
Linoleic Acid to Total Fatty Acids percentage	290	0.85	0.76	0.96	8.1E-03	9.0E-02	306	0.96	0.85	1.09	5.2E-01	1.0E+00
Phospholipids in Very Small VLDL	290	1.18	1.05	1.32	6.8E-03	9.0E-02	306	1.05	0.94	1.18	4.0E-01	1.0E+00
Triglycerides to Total Lipids in Large LDL percentage	290	1.17	1.05	1.32	6.9E-03	9.0E-02	306	1.01	0.89	1.14	8.6E-01	1.0E+00
Citrate	290	1.18	1.04	1.33	8.9E-03	9.1E-02	306	1.18	1.05	1.33	5.9E-03	1.0E+00
Free Cholesterol to Total Lipids in Small HDL percentage	290	1.15	1.04	1.29	9.1E-03	9.1E-02	306	0.99	0.88	1.11	8.6E-01	1.0E+00
Phospholipids in HDL	290	1.18	1.04	1.33	9.8E-03	9.4E-02	306	1.07	0.95	1.21	2.9E-01	1.0E+00
Total Lipids in Very Small VLDL	290	1.16	1.04	1.31	1.0E-02	9.6E-02	306	1.05	0.93	1.17	4.3E-01	1.0E+00
Phosphoglycerides	290	1.17	1.03	1.32	1.3E-02	1.2E-01	306	1.03	0.91	1.16	6.6E-01	1.0E+00
Free Cholesterol in HDL	290	1.16	1.03	1.32	1.5E-02	1.3E-01	306	1.05	0.93	1.19	3.9E-01	1.0E+00
Triglycerides to Total Lipids in IDL percentage	290	1.16	1.03	1.30	1.6E-02	1.3E-01	306	0.98	0.87	1.11	7.9E-01	1.0E+00

Appendix 5.7 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	≤ 5 years						> 5 years					
Phosphatidylcholines	290	1.16	1.03	1.31	1.7E-02	1.3E-01	306	1.03	0.91	1.16	6.6E-01	1.0E+00
Phospholipids in Medium HDL	290	1.17	1.03	1.32	1.8E-02	1.4E-01	306	1.07	0.94	1.21	2.9E-01	1.0E+00
Total Lipids in HDL	290	1.15	1.02	1.31	2.4E-02	1.4E-01	306	1.06	0.94	1.20	3.2E-01	1.0E+00
Average Diameter for HDL Particles	290	1.15	1.02	1.29	2.3E-02	1.4E-01	306	1.07	0.95	1.20	2.7E-01	1.0E+00
Total Cholines	290	1.15	1.02	1.30	2.2E-02	1.4E-01	306	1.02	0.91	1.15	7.1E-01	1.0E+00
Tyrosine	290	1.15	1.02	1.29	2.1E-02	1.4E-01	306	1.01	0.90	1.14	8.8E-01	1.0E+00
Free Cholesterol in Very Small VLDL	290	1.14	1.02	1.28	2.4E-02	1.4E-01	306	1.05	0.93	1.17	4.4E-01	1.0E+00
Triglycerides in Medium LDL	290	1.15	1.02	1.30	2.5E-02	1.4E-01	306	0.98	0.87	1.10	7.1E-01	1.0E+00
Phospholipids in Large HDL	290	1.17	1.03	1.34	1.9E-02	1.4E-01	306	1.10	0.96	1.25	1.7E-01	1.0E+00
Free Cholesterol to Total Lipids in IDL percentage	290	0.88	0.80	0.98	2.5E-02	1.4E-01	306	1.03	0.92	1.16	5.9E-01	1.0E+00
Triglycerides to Total Lipids in Small LDL percentage	290	1.15	1.02	1.29	2.4E-02	1.4E-01	306	1.00	0.88	1.13	9.8E-01	1.0E+00
Cholesterol to Total Lipids in Very Large HDL percentage	290	0.94	0.88	0.99	2.2E-02	1.4E-01	306	0.98	0.89	1.08	7.0E-01	1.0E+00
Free Cholesterol to Total Lipids in Very Large HDL percentage	290	0.90	0.82	0.99	2.4E-02	1.4E-01	306	0.97	0.86	1.08	5.4E-01	1.0E+00
Total Fatty Acids	290	1.15	1.02	1.29	2.6E-02	1.5E-01	306	1.00	0.89	1.13	9.4E-01	1.0E+00
Concentration of Very Small VLDL Particles	290	1.14	1.02	1.28	2.6E-02	1.5E-01	306	1.05	0.94	1.17	4.3E-01	1.0E+00

Appendix 5.7 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	≤ 5 years						> 5 years					
Cholesterol to Total Lipids in IDL percentage	290	0.88	0.79	0.99	2.7E-02	1.5E-01	306	0.99	0.88	1.12	9.1E-01	1.0E+00
Monounsaturated Fatty Acids	290	1.15	1.01	1.29	2.8E-02	1.5E-01	306	1.00	0.89	1.13	9.9E-01	1.0E+00
Polyunsaturated Fatty Acids to Monounsaturated Fatty Acids ratio	290	0.87	0.77	0.99	3.0E-02	1.5E-01	306	1.00	0.88	1.13	9.9E-01	1.0E+00
Triglycerides in Small VLDL	290	1.15	1.01	1.30	3.0E-02	1.5E-01	306	0.96	0.85	1.08	5.1E-01	1.0E+00
Sphingomyelins	290	1.14	1.01	1.28	3.3E-02	1.6E-01	306	1.04	0.93	1.17	5.0E-01	1.0E+00
Cholesteryl Esters to Total Lipids in Medium HDL percentage	290	0.96	0.92	1.00	3.4E-02	1.7E-01	306	1.02	0.89	1.17	7.8E-01	1.0E+00
Total Phospholipids in Lipoprotein Particles	290	1.14	1.01	1.28	3.5E-02	1.7E-01	306	1.02	0.91	1.14	7.6E-01	1.0E+00
Free Cholesterol in Medium HDL	290	1.15	1.01	1.31	3.7E-02	1.7E-01	306	1.06	0.93	1.20	4.0E-01	1.0E+00
Histidine	290	0.89	0.79	0.99	4.1E-02	1.8E-01	306	0.90	0.81	1.01	7.5E-02	1.0E+00
Total Lipids in Large HDL	290	1.14	1.01	1.30	4.1E-02	1.8E-01	306	1.09	0.96	1.23	2.0E-01	1.0E+00
Total Lipids in Medium HDL	290	1.14	1.00	1.29	4.2E-02	1.8E-01	306	1.07	0.94	1.21	3.2E-01	1.0E+00
Triglycerides in Small HDL	290	1.14	1.01	1.29	4.1E-02	1.8E-01	306	0.97	0.86	1.10	6.5E-01	1.0E+00
Apolipoprotein A1	290	1.13	1.00	1.28	5.0E-02	2.1E-01	306	1.05	0.93	1.19	4.1E-01	1.0E+00
Free Cholesterol in Large HDL	290	1.15	1.00	1.34	5.3E-02	2.2E-01	306	1.09	0.95	1.25	2.2E-01	1.0E+00

Appendix 5.7 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	≤ 5 years						> 5 years					
Concentration of Very Large HDL Particles	290	1.13	1.00	1.28	5.6E-02	2.3E-01	306	1.03	0.92	1.16	5.9E-01	1.0E+00
Concentration of Large HDL Particles	290	1.14	0.99	1.30	6.0E-02	2.4E-01	306	1.08	0.95	1.23	2.3E-01	1.0E+00
Phenylalanine	290	1.12	0.99	1.26	6.2E-02	2.5E-01	306	0.95	0.84	1.06	3.5E-01	1.0E+00
Free Cholesterol in Small HDL	290	1.12	0.99	1.27	6.4E-02	2.5E-01	306	1.00	0.89	1.13	9.9E-01	1.0E+00
Triglycerides to Total Lipids in Medium HDL percentage	290	1.13	0.99	1.29	6.3E-02	2.5E-01	306	0.98	0.86	1.10	7.0E-01	1.0E+00
Triglycerides to Total Lipids in Small HDL percentage	290	1.13	0.99	1.28	6.8E-02	2.6E-01	306	0.97	0.86	1.09	5.8E-01	1.0E+00
Free Cholesterol to Total Lipids in Medium HDL percentage	290	1.14	0.99	1.31	7.1E-02	2.7E-01	306	1.02	0.89	1.16	7.8E-01	1.0E+00
Concentration of Medium HDL Particles	290	1.12	0.99	1.28	8.0E-02	2.9E-01	306	1.06	0.93	1.21	3.6E-01	1.0E+00
Phospholipids to Total Lipids in Medium HDL percentage	290	1.14	0.98	1.32	8.1E-02	2.9E-01	306	1.01	0.87	1.18	8.7E-01	1.0E+00
Triglycerides in Small LDL	290	1.11	0.99	1.26	8.2E-02	3.0E-01	306	0.98	0.86	1.10	7.1E-01	1.0E+00
Triglycerides to Total Lipids in Small VLDL percentage	290	1.12	0.98	1.26	8.6E-02	3.1E-01	306	0.95	0.84	1.07	3.6E-01	1.0E+00
Cholesterol in Very Small VLDL	290	1.10	0.99	1.24	8.9E-02	3.1E-01	306	1.05	0.94	1.18	3.7E-01	1.0E+00

Appendix 5.7 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	≤ 5 years						> 5 years					
Cholesteryl Esters to Total Lipids in IDL percentage	290	0.91	0.81	1.02	9.4E-02	3.2E-01	306	0.98	0.87	1.10	7.2E-01	1.0E+00
Degree of Unsaturation	290	0.90	0.80	1.02	1.0E-01	3.3E-01	306	1.01	0.90	1.14	8.9E-01	1.0E+00
Pyruvate	290	1.11	0.98	1.25	1.0E-01	3.3E-01	305	1.11	0.98	1.25	8.9E-02	1.0E+00
Total Lipids in Very Large HDL	290	1.11	0.98	1.25	1.0E-01	3.3E-01	306	1.04	0.92	1.17	5.1E-01	1.0E+00
Free Cholesterol to Total Lipids in Very Large VLDL percentage	289	0.91	0.82	1.02	1.0E-01	3.3E-01	305	1.08	0.96	1.22	1.9E-01	1.0E+00
Phospholipids to Total Lipids in Small VLDL percentage	290	0.91	0.80	1.02	9.9E-02	3.3E-01	306	1.03	0.91	1.16	6.6E-01	1.0E+00
Concentration of Small VLDL Particles	290	1.10	0.98	1.24	1.1E-01	3.5E-01	306	0.98	0.87	1.10	7.2E-01	1.0E+00
Cholesterol to Total Lipids in Medium HDL percentage	290	0.97	0.93	1.01	1.1E-01	3.5E-01	306	1.02	0.89	1.18	7.5E-01	1.0E+00
Phospholipids in Very Large HDL	290	1.14	0.97	1.34	1.2E-01	3.6E-01	306	1.06	0.93	1.21	4.0E-01	1.0E+00
Cholesterol to Total Lipids in Small VLDL percentage	290	0.91	0.81	1.02	1.2E-01	3.6E-01	306	1.05	0.93	1.18	4.1E-01	1.0E+00
Total Lipids in Small VLDL	290	1.10	0.98	1.24	1.2E-01	3.6E-01	306	0.99	0.88	1.11	8.0E-01	1.0E+00
Total Lipids in Lipoprotein Particles	290	1.10	0.98	1.23	1.2E-01	3.6E-01	306	0.99	0.89	1.11	9.2E-01	1.0E+00
Total Triglycerides	290	1.10	0.97	1.25	1.2E-01	3.7E-01	306	0.97	0.86	1.09	5.7E-01	1.0E+00
Cholesteryl Esters to Total Lipids in Small VLDL percentage	290	0.92	0.82	1.02	1.2E-01	3.7E-01	306	1.07	0.95	1.20	2.7E-01	1.0E+00

Appendix 5.7 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	≤ 5 years						> 5 years					
Phospholipids to Total Lipids in Very Small VLDL percentage	290	1.10	0.97	1.24	1.3E-01	3.7E-01	306	1.03	0.91	1.16	6.5E-01	1.0E+00
Lactate	289	0.91	0.81	1.03	1.3E-01	3.7E-01	306	1.04	0.93	1.17	4.8E-01	1.0E+00
Monounsaturated Fatty Acids to Total Fatty Acids percentage	290	1.10	0.97	1.25	1.4E-01	3.9E-01	306	0.99	0.88	1.12	9.0E-01	1.0E+00
Phospholipids in Small HDL	290	1.10	0.97	1.24	1.4E-01	3.9E-01	306	1.01	0.90	1.15	8.2E-01	1.0E+00
Concentration of VLDL Particles	290	1.09	0.97	1.23	1.4E-01	3.9E-01	306	1.00	0.89	1.12	9.6E-01	1.0E+00
Glutamine	290	0.92	0.82	1.03	1.4E-01	3.9E-01	305	0.93	0.83	1.04	1.9E-01	1.0E+00
Cholesteryl Esters to Total Lipids in Very Small VLDL percentage	290	0.92	0.82	1.03	1.4E-01	3.9E-01	306	1.04	0.92	1.17	5.6E-01	1.0E+00
Free Cholesterol to Total Lipids in Small VLDL percentage	290	0.92	0.82	1.03	1.5E-01	3.9E-01	306	1.02	0.91	1.15	7.3E-01	1.0E+00
Cholesterol to Total Lipids in Very Small VLDL percentage	290	0.92	0.82	1.03	1.5E-01	4.0E-01	306	1.04	0.92	1.17	5.7E-01	1.0E+00
Cholesteryl Esters in Very Small VLDL	290	1.09	0.97	1.22	1.5E-01	4.0E-01	306	1.05	0.94	1.18	3.6E-01	1.0E+00
Phospholipids to Total Lipids in Medium LDL percentage	290	0.92	0.82	1.03	1.5E-01	4.0E-01	306	1.10	0.99	1.23	8.4E-02	1.0E+00
Cholesterol in Large HDL	290	1.10	0.96	1.26	1.7E-01	4.4E-01	306	1.09	0.95	1.24	2.0E-01	1.0E+00
Valine	290	1.09	0.96	1.22	1.8E-01	4.6E-01	306	0.94	0.83	1.05	2.7E-01	1.0E+00

Appendix 5.7 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	≤ 5 years						> 5 years					
Triglycerides to Total Lipids in Very Small VLDL percentage	290	1.09	0.96	1.23	1.8E-01	4.6E-01	306	0.94	0.83	1.06	3.1E-01	1.0E+00
HDL Cholesterol	290	1.08	0.96	1.23	2.1E-01	5.2E-01	306	1.06	0.94	1.20	3.6E-01	1.0E+00
Free Cholesterol to Total Lipids in Large VLDL percentage	290	0.94	0.85	1.04	2.2E-01	5.5E-01	306	1.02	0.91	1.15	7.0E-01	1.0E+00
Total Concentration of Lipoprotein Particles	290	1.08	0.95	1.22	2.4E-01	5.7E-01	306	1.02	0.91	1.16	7.0E-01	1.0E+00
Concentration of HDL Particles	290	1.08	0.95	1.22	2.5E-01	5.7E-01	306	1.03	0.91	1.17	6.3E-01	1.0E+00
Polyunsaturated Fatty Acids	290	1.07	0.95	1.21	2.4E-01	5.7E-01	306	1.00	0.89	1.12	1.0E+00	1.0E+00
Docosahexaenoic Acid to Total Fatty Acids percentage	290	0.93	0.83	1.05	2.5E-01	5.7E-01	306	1.01	0.89	1.14	9.3E-01	1.0E+00
Isoleucine	290	1.07	0.95	1.20	2.5E-01	5.7E-01	306	0.94	0.84	1.06	3.3E-01	1.0E+00
Acetoacetate	288	1.08	0.95	1.22	2.3E-01	5.7E-01	304	1.04	0.92	1.17	5.5E-01	1.0E+00
Cholesteryl Esters in Large HDL	290	1.08	0.95	1.24	2.5E-01	5.7E-01	306	1.09	0.96	1.25	2.0E-01	1.0E+00
Cholesterol in Medium HDL	290	1.08	0.95	1.23	2.5E-01	5.7E-01	306	1.07	0.94	1.21	3.3E-01	1.0E+00
Phospholipids to Total Lipids in Large HDL percentage	290	1.12	0.92	1.38	2.6E-01	5.8E-01	306	1.01	0.86	1.20	8.7E-01	1.0E+00
Free Cholesterol to Total Lipids in Large HDL percentage	290	1.13	0.91	1.40	2.6E-01	5.8E-01	306	1.07	0.90	1.26	4.7E-01	1.0E+00
Omega-6 Fatty Acids	290	1.07	0.95	1.20	2.6E-01	5.9E-01	306	0.99	0.89	1.11	9.1E-01	1.0E+00
Triglycerides in Medium VLDL	290	1.07	0.95	1.21	2.7E-01	5.9E-01	306	0.95	0.84	1.07	3.8E-01	1.0E+00

Appendix 5.7 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	≤ 5 years						> 5 years					
Phospholipids to Total Lipids in Very Large HDL percentage	290	1.12	0.92	1.36	2.7E-01	5.9E-01	306	1.06	0.91	1.23	4.8E-01	1.0E+00
Average Diameter for LDL Particles	290	1.07	0.94	1.20	3.0E-01	6.1E-01	306	1.04	0.92	1.17	5.4E-01	1.0E+00
Phospholipids in Small VLDL	290	1.07	0.95	1.20	2.9E-01	6.1E-01	306	0.99	0.89	1.11	9.0E-01	1.0E+00
Cholesterol in Very Large HDL	290	1.07	0.94	1.21	3.0E-01	6.1E-01	306	1.04	0.92	1.17	5.3E-01	1.0E+00
Cholesteryl Esters in Very Large HDL	290	1.08	0.93	1.26	2.9E-01	6.1E-01	306	1.06	0.92	1.22	3.9E-01	1.0E+00
Total Lipids in Small HDL	290	1.07	0.94	1.21	3.0E-01	6.1E-01	306	1.00	0.89	1.13	9.6E-01	1.0E+00
Cholesteryl Esters in Small HDL	290	0.93	0.83	1.06	2.8E-01	6.1E-01	306	1.00	0.88	1.13	9.5E-01	1.0E+00
Free Cholesterol to Total Lipids in Small LDL percentage	290	0.97	0.93	1.02	2.9E-01	6.1E-01	306	1.02	0.87	1.19	8.4E-01	1.0E+00
Triglycerides in VLDL	290	1.07	0.94	1.21	3.0E-01	6.1E-01	306	0.96	0.85	1.08	5.3E-01	1.0E+00
Phospholipids in VLDL	290	1.06	0.94	1.20	3.1E-01	6.3E-01	306	0.99	0.89	1.12	9.2E-01	1.0E+00
BCAA	290	1.06	0.94	1.20	3.1E-01	6.3E-01	306	0.93	0.83	1.05	2.4E-01	1.0E+00
Total Lipids in VLDL	290	1.06	0.94	1.20	3.2E-01	6.4E-01	306	0.98	0.87	1.10	7.5E-01	1.0E+00
Phospholipids in IDL	290	1.06	0.94	1.19	3.3E-01	6.5E-01	306	1.03	0.92	1.15	5.9E-01	1.0E+00
Free Cholesterol in Small LDL	290	0.97	0.91	1.03	3.4E-01	6.5E-01	306	0.99	0.88	1.10	8.2E-01	1.0E+00
Free Cholesterol to Total Lipids in Very Small VLDL percentage	290	0.95	0.84	1.06	3.3E-01	6.5E-01	306	1.01	0.90	1.13	8.8E-01	1.0E+00

Appendix 5.7 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	≤ 5 years						> 5 years					
Phospholipids to Total Lipids in Small LDL percentage	290	1.07	0.93	1.24	3.4E-01	6.5E-01	306	1.09	0.94	1.26	2.3E-01	1.0E+00
Phospholipids to Total Lipids in Very Large VLDL percentage	289	0.96	0.88	1.04	3.4E-01	6.5E-01	305	1.00	0.90	1.12	9.5E-01	1.0E+00
Omega-3 Fatty Acids	290	1.06	0.94	1.20	3.5E-01	6.6E-01	306	1.01	0.90	1.14	8.1E-01	1.0E+00
Triglycerides to Total Lipids in Very Large VLDL percentage	289	1.08	0.92	1.27	3.5E-01	6.7E-01	305	0.97	0.89	1.06	5.2E-01	1.0E+00
Free Cholesterol in Very Large HDL	290	1.06	0.93	1.20	3.6E-01	6.7E-01	306	1.04	0.92	1.17	5.8E-01	1.0E+00
Acetate	290	1.06	0.94	1.19	3.7E-01	6.8E-01	304	0.92	0.82	1.04	1.9E-01	1.0E+00
Cholesteryl Esters in Medium HDL	290	1.06	0.93	1.21	3.7E-01	6.8E-01	306	1.07	0.94	1.21	3.3E-01	1.0E+00
Cholesteryl Esters in HDL	290	1.06	0.93	1.20	3.9E-01	7.0E-01	306	1.06	0.94	1.20	3.5E-01	1.0E+00
Total Free Cholesterol	290	1.05	0.94	1.18	3.9E-01	7.0E-01	306	1.00	0.89	1.12	1.0E+00	1.0E+00
Cholesteryl Esters to Total Lipids in Medium VLDL percentage	290	0.96	0.87	1.06	3.9E-01	7.0E-01	306	1.03	0.91	1.15	6.8E-01	1.0E+00
Triglycerides to Total Lipids in Medium VLDL percentage	290	1.06	0.93	1.19	3.9E-01	7.0E-01	306	0.94	0.85	1.04	2.5E-01	1.0E+00
Phospholipids to Total Lipids in IDL percentage	290	1.05	0.94	1.18	3.9E-01	7.0E-01	306	1.01	0.90	1.13	9.1E-01	1.0E+00
Cholesteryl Esters in Small VLDL	290	1.05	0.94	1.18	4.0E-01	7.1E-01	306	1.01	0.90	1.13	8.5E-01	1.0E+00

Appendix 5.7 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	≤ 5 years						> 5 years					
Total Lipids in Medium VLDL	290	1.05	0.93	1.18	4.2E-01	7.3E-01	306	0.97	0.87	1.09	6.4E-01	1.0E+00
Phospholipids to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	282	1.06	0.92	1.23	4.2E-01	7.3E-01	301	0.95	0.86	1.04	2.9E-01	1.0E+00
Free Cholesterol in VLDL	290	1.05	0.93	1.18	4.2E-01	7.3E-01	306	0.99	0.89	1.12	9.3E-01	1.0E+00
Apolipoprotein B to Apolipoprotein A1 ratio	290	0.95	0.85	1.07	4.4E-01	7.3E-01	306	0.97	0.87	1.09	6.2E-01	1.0E+00
Phospholipids in Large VLDL	290	1.05	0.92	1.20	4.3E-01	7.3E-01	306	0.98	0.88	1.09	6.7E-01	1.0E+00
Cholesterol in Small VLDL	290	1.05	0.93	1.18	4.3E-01	7.3E-01	306	1.01	0.90	1.13	9.2E-01	1.0E+00
Total Lipids in IDL	290	1.05	0.93	1.17	4.3E-01	7.3E-01	306	1.03	0.92	1.15	6.2E-01	1.0E+00
Cholesterol to Total Lipids in Very Large VLDL percentage	289	0.95	0.84	1.08	4.4E-01	7.3E-01	305	1.05	0.93	1.18	4.1E-01	1.0E+00
Cholesterol to Total Lipids in Medium VLDL percentage	290	0.96	0.85	1.07	4.3E-01	7.3E-01	306	1.04	0.92	1.17	5.2E-01	1.0E+00
VLDL Cholesterol	290	1.05	0.93	1.17	4.5E-01	7.4E-01	306	1.00	0.90	1.12	9.4E-01	1.0E+00
Concentration of Large VLDL Particles	290	1.05	0.93	1.19	4.5E-01	7.4E-01	306	0.97	0.86	1.09	5.8E-01	1.0E+00
Average Diameter for VLDL Particles	290	0.96	0.85	1.08	4.5E-01	7.4E-01	306	0.94	0.83	1.06	3.1E-01	1.0E+00
Triglycerides to Total Lipids in Very Large HDL percentage	290	1.05	0.93	1.18	4.6E-01	7.4E-01	306	0.98	0.87	1.10	7.2E-01	1.0E+00

Appendix 5.7 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	≤ 5 years						> 5 years					
Triglycerides to Total Lipids in Large HDL percentage	290	1.05	0.92	1.19	4.6E-01	7.4E-01	306	0.95	0.84	1.08	4.5E-01	1.0E+00
Free Cholesterol to Total Lipids in Medium LDL percentage	290	0.98	0.93	1.03	4.7E-01	7.5E-01	306	1.07	0.86	1.32	5.5E-01	1.0E+00
Total Cholesterol	290	1.04	0.93	1.17	4.9E-01	7.5E-01	306	1.00	0.89	1.12	9.7E-01	1.0E+00
Remnant Cholesterol (Non-HDL, Non-LDL -Cholesterol)	290	1.04	0.93	1.17	4.8E-01	7.5E-01	306	1.01	0.91	1.13	8.2E-01	1.0E+00
Cholesteryl Esters in VLDL	290	1.04	0.93	1.17	4.9E-01	7.5E-01	306	1.01	0.90	1.13	8.5E-01	1.0E+00
Triglycerides in Large VLDL	290	1.05	0.92	1.18	4.8E-01	7.5E-01	306	0.95	0.85	1.07	3.8E-01	1.0E+00
Phospholipids in Medium VLDL	290	1.04	0.93	1.17	4.9E-01	7.5E-01	306	0.99	0.89	1.11	8.6E-01	1.0E+00
Free Cholesterol in Small VLDL	290	1.04	0.93	1.17	4.9E-01	7.5E-01	306	1.00	0.89	1.11	9.5E-01	1.0E+00
Total Lipids in Large VLDL	290	1.04	0.92	1.18	5.1E-01	7.7E-01	306	0.96	0.86	1.08	5.1E-01	1.0E+00
Triglycerides in Very Large VLDL	290	1.04	0.92	1.18	5.2E-01	7.9E-01	306	0.98	0.87	1.10	7.6E-01	1.0E+00
Free Cholesterol in Medium LDL	290	0.98	0.91	1.05	5.2E-01	7.9E-01	306	0.98	0.88	1.09	7.4E-01	1.0E+00
Phospholipids to Total Lipids in Large VLDL percentage	290	1.05	0.90	1.23	5.2E-01	7.9E-01	306	1.00	0.90	1.12	1.0E+00	1.0E+00
Cholesteryl Esters in Large VLDL	290	1.04	0.92	1.17	5.3E-01	7.9E-01	306	0.98	0.88	1.10	7.6E-01	1.0E+00
Total Esterified Cholesterol	290	1.04	0.92	1.17	5.4E-01	8.0E-01	306	1.00	0.90	1.12	9.6E-01	1.0E+00
Creatinine	280	0.96	0.86	1.08	5.4E-01	8.0E-01	304	0.89	0.79	0.99	4.0E-02	1.0E+00

Appendix 5.7 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	≤ 5 years						> 5 years					
Free Cholesterol in Medium VLDL	290	1.04	0.92	1.16	5.5E-01	8.1E-01	306	0.99	0.89	1.11	9.2E-01	1.0E+00
Concentration of Very Large VLDL Particles	290	1.04	0.92	1.18	5.6E-01	8.1E-01	306	0.97	0.86	1.09	6.2E-01	1.0E+00
Free Cholesterol to Total Lipids in Large LDL percentage	290	0.99	0.94	1.03	5.6E-01	8.1E-01	306	1.05	0.83	1.34	6.9E-01	1.0E+00
Linoleic Acid	290	1.03	0.92	1.16	5.6E-01	8.1E-01	306	0.98	0.88	1.10	7.6E-01	1.0E+00
Omega-6 Fatty Acids to Omega-3 Fatty Acids ratio	290	0.97	0.86	1.09	5.7E-01	8.2E-01	306	0.98	0.87	1.11	7.6E-01	1.0E+00
Triglycerides to Phosphoglycerides ratio	290	1.04	0.91	1.18	5.8E-01	8.2E-01	306	0.95	0.84	1.07	3.8E-01	1.0E+00
Cholesterol in Large VLDL	290	1.03	0.92	1.17	5.9E-01	8.2E-01	306	0.97	0.87	1.09	6.6E-01	1.0E+00
Concentration of Medium VLDL Particles	290	1.03	0.92	1.16	5.9E-01	8.2E-01	306	0.98	0.88	1.10	7.5E-01	1.0E+00
Cholesteryl Esters to Total Lipids in Very Large VLDL percentage	289	0.97	0.86	1.09	5.9E-01	8.2E-01	305	1.04	0.92	1.17	5.1E-01	1.0E+00
Cholesterol to Total Lipids in Large VLDL percentage	290	0.97	0.86	1.09	5.8E-01	8.2E-01	306	1.06	0.94	1.19	3.7E-01	1.0E+00
Acetone	290	0.97	0.86	1.09	6.2E-01	8.5E-01	306	0.88	0.78	1.00	4.6E-02	1.0E+00
Triglycerides in Chylomicrons and Extremely Large VLDL	290	0.97	0.87	1.09	6.4E-01	8.7E-01	306	0.99	0.89	1.11	8.8E-01	1.0E+00

Appendix 5.7 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], stratified by time since blood collection

	≤ 5 years						> 5 years					
Cholesterol in Small LDL	290	0.98	0.91	1.06	6.4E-01	8.7E-01	306	0.97	0.89	1.07	5.7E-01	1.0E+00
Total Lipids in Very Large VLDL	290	1.03	0.91	1.16	6.6E-01	8.9E-01	306	0.98	0.87	1.10	7.2E-01	1.0E+00
Free Cholesterol in Large VLDL	290	1.03	0.91	1.16	6.5E-01	8.9E-01	306	0.97	0.86	1.08	5.6E-01	1.0E+00
Cholesterol to Total Lipids in Small LDL percentage	290	0.99	0.95	1.03	6.7E-01	9.0E-01	306	0.99	0.92	1.06	7.4E-01	1.0E+00
Clinical LDL Cholesterol	290	0.98	0.87	1.09	6.8E-01	9.1E-01	306	0.99	0.88	1.10	8.1E-01	1.0E+00
Cholesteryl Esters to Total Lipids in Large HDL percentage	290	0.98	0.90	1.07	6.9E-01	9.2E-01	306	1.07	0.90	1.27	4.2E-01	1.0E+00
Free Cholesterol in LDL	290	0.98	0.90	1.07	7.0E-01	9.2E-01	306	0.99	0.88	1.12	9.1E-01	1.0E+00
Glycine	287	0.98	0.88	1.09	7.0E-01	9.2E-01	301	1.03	0.91	1.16	6.8E-01	1.0E+00
Cholesteryl Esters in IDL	290	1.02	0.91	1.15	7.0E-01	9.2E-01	306	1.02	0.91	1.14	7.1E-01	1.0E+00
Phospholipids in Medium LDL	290	0.98	0.87	1.10	7.0E-01	9.2E-01	306	0.97	0.87	1.08	5.8E-01	1.0E+00

Abbreviations: HR, hazard ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit; LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol; IDL, intermediate-density lipoprotein cholesterol; VLDL, very low-density lipoprotein cholesterol; BCAA, Total concentration of branched-chain amino acids (Leucine + Isoleucine + Valine).

1 Hazard ratios (95% confidence interval) for risk of prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features by time since blood collection. Analyses were stratified by region and Townsend Deprivation Index with age as the underlying time variable. Adjusted for fasting status, body mass index, height, alcohol intake, family history of prostate cancer, ethnicity, level of education, smoking status, and weekly physical activity.

2 Padj < 0.05 are statistically significant after allowing for multiple testing using a false discovery rate controlling procedure at α = 0.05 (Benjamini-Hochberg). Padj values in the ≤ 5 years group are sorted in ascending order. Association with Padj < 0.05 in both the overall risk of death dataset and this dataset are bolded.

Appendix 5.8 Hazard ratios for risk of prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes.

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Glycoprotein Acetyls	**596**	**1.24**	**1.14**	**1.35**	**6.8E-07**	**1.7E-04**
Albumin	**596**	**0.85**	**0.79**	**0.92**	**9.3E-05**	**1.2E-02**
Citrate	**596**	**1.17**	**1.08**	**1.28**	**2.8E-04**	**2.3E-02**
Cholesteryl Esters to Total Lipids in Small HDL percentage	596	0.87	0.79	0.95	2.0E-03	1.2E-01
Phospholipids in HDL	596	1.13	1.03	1.23	6.8E-03	1.3E-01
Histidine	596	0.90	0.83	0.97	6.8E-03	1.3E-01
Concentration of Very Small VLDL Particles	596	1.12	1.03	1.22	6.8E-03	1.3E-01
Total Lipids in Very Small VLDL	596	1.13	1.04	1.23	3.5E-03	1.3E-01
Phospholipids in Very Small VLDL	596	1.14	1.05	1.23	2.7E-03	1.3E-01
Free Cholesterol in Very Small VLDL	596	1.12	1.03	1.22	6.1E-03	1.3E-01
Triglycerides in IDL	596	1.13	1.04	1.23	5.3E-03	1.3E-01
Phospholipids in Large HDL	596	1.14	1.04	1.25	5.5E-03	1.3E-01
Phospholipids to Total Lipids in Small HDL percentage	596	1.13	1.04	1.23	4.9E-03	1.3E-01
Triglycerides in Large HDL	596	1.13	1.03	1.23	7.8E-03	1.4E-01
Free Cholesterol in HDL	596	1.12	1.03	1.22	9.3E-03	1.4E-01
Total Lipids in HDL	596	1.12	1.02	1.22	1.4E-02	1.4E-01
Average Diameter for HDL Particles	596	1.11	1.02	1.21	1.2E-02	1.4E-01
Sphingomyelins	596	1.12	1.03	1.22	1.1E-02	1.4E-01
Glucose	595	1.12	1.03	1.22	8.6E-03	1.4E-01
Cholesterol in Very Small VLDL	596	1.11	1.03	1.21	1.1E-02	1.4E-01
Concentration of Large HDL Particles	596	1.13	1.03	1.24	1.4E-02	1.4E-01
Total Lipids in Large HDL	596	1.12	1.03	1.23	1.0E-02	1.4E-01
Free Cholesterol in Large HDL	596	1.14	1.03	1.26	1.2E-02	1.4E-01
Phospholipids in Medium HDL	596	1.12	1.02	1.22	1.4E-02	1.4E-01
Cholesterol to Total Lipids in Small HDL percentage	596	0.90	0.83	0.98	1.2E-02	1.4E-01
Phosphoglycerides	596	1.11	1.02	1.21	1.6E-02	1.4E-01
Saturated Fatty Acids to Total Fatty Acids percentage	596	1.11	1.02	1.21	1.6E-02	1.4E-01
Cholesteryl Esters in Very Small VLDL	596	1.11	1.02	1.21	1.6E-02	1.4E-01
Phosphatidylcholines	596	1.11	1.02	1.21	1.7E-02	1.5E-01
Triglycerides in Very Large HDL	596	1.11	1.02	1.21	1.9E-02	1.6E-01
Triglycerides in Very Small VLDL	596	1.11	1.02	1.21	2.0E-02	1.6E-01
Total Cholines	596	1.11	1.02	1.20	2.1E-02	1.6E-01
Saturated Fatty Acids	596	1.10	1.01	1.20	2.4E-02	1.6E-01
Pyruvate	595	1.11	1.01	1.21	2.3E-02	1.6E-01

Appendix 5.8 Hazard ratios for risk of prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes.

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Triglycerides in Large LDL	596	1.11	1.02	1.20	2.2E-02	1.6E-01
Triglycerides in Medium HDL	596	1.11	1.01	1.21	2.4E-02	1.6E-01
Triglycerides in HDL	596	1.11	1.01	1.21	2.6E-02	1.7E-01
Total Phospholipids in Lipoprotein Particles	596	1.10	1.01	1.20	2.7E-02	1.7E-01
Total Lipids in Medium HDL	596	1.11	1.01	1.21	2.8E-02	1.7E-01
Free Cholesterol in Medium HDL	596	1.11	1.01	1.21	2.6E-02	1.7E-01
Free Cholesterol to Total Lipids in Very Large HDL percentage	596	0.93	0.87	0.99	2.9E-02	1.8E-01
Apolipoprotein A1	596	1.10	1.01	1.20	3.3E-02	1.9E-01
Concentration of Very Large HDL Particles	596	1.10	1.01	1.20	3.3E-02	1.9E-01
Cholesterol in Large HDL	596	1.11	1.01	1.22	3.2E-02	1.9E-01
Free Cholesterol to Total Lipids in Small HDL percentage	596	1.09	1.00	1.18	4.0E-02	2.2E-01
Linoleic Acid to Total Fatty Acids percentage	596	0.91	0.84	1.00	4.3E-02	2.3E-01
Omega-6 Fatty Acids to Total Fatty Acids percentage	596	0.92	0.84	1.00	4.7E-02	2.4E-01
Total Lipids in Very Large HDL	596	1.09	1.00	1.19	4.8E-02	2.4E-01
Cholesteryl Esters in Large HDL	596	1.10	1.00	1.21	4.6E-02	2.4E-01
Concentration of Medium HDL Particles	596	1.10	1.00	1.20	4.7E-02	2.4E-01
Triglycerides in LDL	596	1.09	1.00	1.19	5.3E-02	2.5E-01
Polyunsaturated Fatty Acids to Total Fatty Acids percentage	596	0.92	0.84	1.00	5.1E-02	2.5E-01
Phospholipids in Very Large HDL	596	1.11	1.00	1.24	5.2E-02	2.5E-01
Phospholipids to Total Lipids in Large LDL percentage	596	0.92	0.85	1.00	5.6E-02	2.6E-01
Total Fatty Acids	596	1.09	1.00	1.18	5.7E-02	2.6E-01
Triglycerides to Total Lipids in Medium LDL percentage	596	1.09	1.00	1.18	6.0E-02	2.7E-01
Cholesterol to Total Lipids in Very Large HDL percentage	596	0.96	0.92	1.00	6.1E-02	2.7E-01
Creatinine	584	0.93	0.85	1.00	6.3E-02	2.7E-01
Acetone	596	0.92	0.85	1.01	6.5E-02	2.8E-01
Glutamine	595	0.93	0.85	1.01	6.7E-02	2.8E-01
Free Cholesterol to Total Lipids in Medium HDL percentage	596	1.10	0.99	1.21	6.9E-02	2.8E-01
Phospholipids in IDL	596	1.08	0.99	1.17	7.4E-02	3.0E-01
HDL Cholesterol	596	1.08	0.99	1.18	8.0E-02	3.2E-01
Tyrosine	596	1.08	0.99	1.17	8.4E-02	3.3E-01
Monounsaturated Fatty Acids	596	1.08	0.99	1.17	9.2E-02	3.5E-01
Cholesteryl Esters in Very Large HDL	596	1.09	0.99	1.22	9.4E-02	3.6E-01
Total Lipids in IDL	596	1.07	0.99	1.17	1.0E-01	3.8E-01

Appendix 5.8 Hazard ratios for risk of prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes.

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Cholesteryl Esters to Total Lipids in Medium HDL percentage	596	0.97	0.93	1.01	1.0E-01	3.8E-01
Triglycerides to Total Lipids in Large LDL percentage	596	1.07	0.99	1.17	1.1E-01	3.8E-01
Free Cholesterol to Total Lipids in Large HDL percentage	596	1.13	0.97	1.30	1.1E-01	3.9E-01
Cholesterol in Medium HDL	596	1.08	0.98	1.18	1.1E-01	3.9E-01
Cholesterol in Very Large HDL	596	1.07	0.98	1.17	1.2E-01	4.0E-01
Total Lipids in Lipoprotein Particles	596	1.07	0.98	1.16	1.3E-01	4.3E-01
Free Cholesterol in Small HDL	596	1.07	0.98	1.17	1.3E-01	4.3E-01
Concentration of IDL Particles	596	1.06	0.98	1.16	1.4E-01	4.7E-01
Cholesteryl Esters in HDL	596	1.07	0.98	1.17	1.5E-01	4.8E-01
Triglycerides in Medium LDL	596	1.07	0.98	1.16	1.5E-01	4.8E-01
Total Concentration of Lipoprotein Particles	596	1.06	0.98	1.16	1.6E-01	4.9E-01
Concentration of VLDL Particles	596	1.06	0.98	1.15	1.6E-01	4.9E-01
Polyunsaturated Fatty Acids to Monounsaturated Fatty Acids ratio	596	0.94	0.86	1.03	1.6E-01	4.9E-01
Free Cholesterol in Very Large HDL	596	1.07	0.98	1.17	1.6E-01	4.9E-01
Cholesteryl Esters in Medium HDL	596	1.07	0.97	1.17	1.6E-01	4.9E-01
Phospholipids to Total Lipids in Very Large HDL percentage	596	1.09	0.97	1.23	1.6E-01	4.9E-01
Average Diameter for LDL Particles	596	1.06	0.98	1.16	1.7E-01	5.1E-01
Concentration of HDL Particles	596	1.06	0.97	1.16	1.7E-01	5.1E-01
Remnant Cholesterol (Non-HDL, Non-LDL -Cholesterol)	596	1.06	0.97	1.15	1.9E-01	5.4E-01
Average Diameter for VLDL Particles	596	0.95	0.87	1.03	1.9E-01	5.4E-01
Phospholipids in Small HDL	596	1.06	0.97	1.16	1.9E-01	5.4E-01
Phospholipids to Total Lipids in Small LDL percentage	596	1.07	0.97	1.19	1.9E-01	5.4E-01
Triglycerides to Total Lipids in Small LDL percentage	596	1.06	0.97	1.15	2.0E-01	5.6E-01
Total Free Cholesterol	596	1.06	0.97	1.15	2.1E-01	5.6E-01
Polyunsaturated Fatty Acids	596	1.05	0.97	1.15	2.2E-01	5.6E-01
Total Lipids in Small VLDL	596	1.06	0.97	1.15	2.1E-01	5.6E-01
Cholesterol in IDL	596	1.06	0.97	1.15	2.1E-01	5.6E-01
Cholesteryl Esters in IDL	596	1.06	0.97	1.15	2.1E-01	5.6E-01
Acetoacetate	592	1.06	0.97	1.15	2.2E-01	5.7E-01
Cholesteryl Esters in VLDL	596	1.05	0.97	1.14	2.3E-01	5.8E-01
Cholesteryl Esters in Small VLDL	596	1.05	0.97	1.14	2.3E-01	5.8E-01
Free Cholesterol in IDL	596	1.05	0.97	1.15	2.3E-01	5.8E-01

Appendix 5.8 Hazard ratios for risk of prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes.

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Phospholipids to Total Lipids in Medium VLDL percentage	596	1.06	0.96	1.18	2.3E-01	5.8E-01
Cholesterol to Total Lipids in IDL percentage	596	0.95	0.88	1.03	2.4E-01	5.9E-01
Phospholipids to Total Lipids in Medium HDL percentage	596	1.07	0.96	1.18	2.4E-01	5.9E-01
Omega-6 Fatty Acids	596	1.05	0.97	1.14	2.4E-01	5.9E-01
Concentration of Small VLDL Particles	596	1.05	0.97	1.14	2.5E-01	5.9E-01
Total Cholesterol	596	1.05	0.97	1.14	2.5E-01	5.9E-01
Cholesterol in Small VLDL	596	1.05	0.97	1.14	2.5E-01	5.9E-01
Phospholipids in Small VLDL	596	1.05	0.97	1.14	2.6E-01	6.0E-01
Triglycerides in Small VLDL	596	1.05	0.96	1.15	2.6E-01	6.0E-01
Phospholipids to Total Lipids in Very Small VLDL percentage	596	1.05	0.96	1.14	2.6E-01	6.0E-01
VLDL Cholesterol	596	1.05	0.96	1.14	2.8E-01	6.1E-01
Total Esterified Cholesterol	596	1.05	0.96	1.14	2.7E-01	6.1E-01
Triglycerides in Small HDL	596	1.05	0.96	1.15	2.7E-01	6.1E-01
Triglycerides in Small LDL	596	1.05	0.96	1.14	2.9E-01	6.3E-01
Triglycerides to Total Lipids in IDL percentage	596	1.05	0.96	1.14	2.9E-01	6.3E-01
Cholesterol to Total Lipids in Medium HDL percentage	596	0.98	0.93	1.02	2.9E-01	6.3E-01
Triglycerides to Total Lipids in Medium HDL percentage	596	1.05	0.96	1.15	3.1E-01	6.7E-01
Free Cholesterol in Small VLDL	596	1.04	0.96	1.13	3.2E-01	6.8E-01
Leucine	596	0.96	0.88	1.04	3.2E-01	6.8E-01
Cholesteryl Esters to Total Lipids in IDL percentage	596	0.96	0.88	1.04	3.3E-01	6.8E-01
Phospholipids in VLDL	596	1.04	0.96	1.13	3.3E-01	6.8E-01
Cholesteryl Esters to Total Lipids in Large VLDL percentage	596	1.04	0.96	1.14	3.4E-01	7.0E-01
Free Cholesterol to Total Lipids in Medium VLDL percentage	596	1.04	0.96	1.14	3.4E-01	7.0E-01
Degree of Unsaturation	596	0.96	0.88	1.05	3.5E-01	7.1E-01
Omega-3 Fatty Acids	596	1.04	0.96	1.14	3.6E-01	7.2E-01
Free Cholesterol in VLDL	596	1.04	0.96	1.13	3.7E-01	7.3E-01
Free Cholesterol in Medium VLDL	596	1.04	0.96	1.13	3.7E-01	7.3E-01
Free Cholesterol to Total Lipids in IDL percentage	596	0.96	0.89	1.04	3.7E-01	7.3E-01
Alanine	596	1.04	0.96	1.13	3.8E-01	7.3E-01
Total Lipids in Small HDL	596	1.04	0.95	1.13	3.8E-01	7.3E-01
Phospholipids in Medium VLDL	596	1.04	0.95	1.13	3.9E-01	7.5E-01
Triglycerides to Total Lipids in Large VLDL percentage	596	0.98	0.92	1.04	4.1E-01	7.7E-01

Appendix 5.8 Hazard ratios for risk of prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes.

Metabolite	Cases	HR	LL	UL	P-value[2]	Padj[2]
Docosahexaenoic Acid to Total Fatty Acids percentage	596	0.97	0.89	1.05	4.1E-01	7.7E-01
Triglycerides to Total Lipids in Small HDL percentage	596	1.04	0.95	1.13	4.1E-01	7.7E-01
Total Triglycerides	596	1.04	0.95	1.13	4.4E-01	8.0E-01
Monounsaturated Fatty Acids to Total Fatty Acids percentage	596	1.04	0.95	1.13	4.3E-01	8.0E-01
3-Hydroxybutyrate	584	1.03	0.95	1.13	4.4E-01	8.0E-01
Cholesterol in Medium VLDL	596	1.03	0.95	1.12	4.3E-01	8.0E-01
Total Cholesterol Minus HDL-C	596	1.03	0.95	1.12	4.5E-01	8.1E-01
Apolipoprotein B	596	1.03	0.95	1.12	4.6E-01	8.2E-01
Total Lipids in VLDL	596	1.03	0.95	1.12	4.7E-01	8.3E-01
Lactate	595	0.97	0.89	1.05	4.7E-01	8.3E-01
Cholesterol to Total Lipids in Large HDL percentage	596	1.04	0.94	1.15	4.9E-01	8.5E-01
Cholesteryl Esters in Small HDL	596	0.97	0.89	1.06	5.0E-01	8.7E-01
Linoleic Acid	596	1.03	0.95	1.12	5.1E-01	8.7E-01
Concentration of Medium VLDL Particles	596	1.03	0.95	1.12	5.3E-01	8.7E-01
Total Lipids in Medium VLDL	596	1.03	0.95	1.12	5.2E-01	8.7E-01
Concentration of Large LDL Particles	596	1.03	0.95	1.12	5.2E-01	8.7E-01
Total Lipids in Large LDL	596	1.03	0.95	1.12	5.2E-01	8.7E-01
Triglycerides to Total Lipids in Medium VLDL percentage	596	0.97	0.90	1.06	5.1E-01	8.7E-01

Abbreviations: SD, standard deviation; HR, hazard ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit; LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol; IDL, intermediate-density lipoprotein cholesterol; VLDL, very low-density lipoprotein cholesterol.

1 Hazard ratios (95% confidence interval) for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features, after additional adjustment for diabetes. Analyses were stratified by region and Townsend Deprivation Index with age as the underlying time variable. Adjusted for diabetes, fasting status, body mass index, height, alcohol intake, family history of prostate cancer, ethnicity, level of education, smoking status, and weekly physical activity.

2 Padj < 0.05 are statistically significant after allowing for multiple testing using a false discovery rate controlling procedure at $\alpha = 0.05$ (Benjamini-Hochberg). Padj values are sorted in ascending order. Associations with Padj < 0.05 are bolded.

Appendix 5.9 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes, stratified by time since blood collection

	≤ 5 years						> 5 years					
Metabolites	Cases	HR	LL	UL	P-value	Padj	Cases	HR	LL	UL	P-value	Padj
Glycoprotein acetyls	290	1.39	1.23	1.57	7.5E-08	1.9E-05	306	1.11	0.98	1.25	9.6E-02	1.0E+00
Albumin	290	0.79	0.71	0.88	2.3E-05	2.0E-03	306	0.92	0.82	1.04	1.8E-01	1.0E+00
Cholesteryl Esters to Total Lipids in Small HDL percentage	290	0.78	0.69	0.87	2.4E-05	2.0E-03	306	0.99	0.86	1.13	8.7E-01	1.0E+00
Triglycerides in IDL	290	1.24	1.10	1.40	3.6E-04	2.2E-02	306	1.02	0.91	1.15	7.2E-01	1.0E+00
Glucose	289	1.23	1.09	1.39	7.6E-04	3.2E-02	306	1.03	0.91	1.16	6.8E-01	1.0E+00
Cholesterol to Total Lipids in Small HDL percentage	290	0.82	0.73	0.92	7.4E-04	3.2E-02	306	0.99	0.88	1.12	8.8E-01	1.0E+00
Triglycerides in Very Small VLDL	290	1.22	1.08	1.38	1.2E-03	4.3E-02	306	1.00	0.89	1.13	9.8E-01	1.0E+00
Triglycerides in Large HDL	290	1.23	1.08	1.40	1.4E-03	4.3E-02	306	1.04	0.92	1.17	5.8E-01	1.0E+00
Triglycerides in HDL	290	1.21	1.07	1.37	2.4E-03	5.4E-02	306	1.01	0.89	1.14	9.1E-01	1.0E+00
Phospholipids in Very Small VLDL	290	1.20	1.06	1.35	2.9E-03	5.4E-02	306	1.08	0.96	1.21	2.2E-01	1.0E+00
Triglycerides in Large LDL	290	1.21	1.07	1.36	2.1E-03	5.4E-02	306	1.01	0.89	1.14	8.9E-01	1.0E+00
Triglycerides in Very Large HDL	290	1.20	1.06	1.36	3.2E-03	5.4E-02	306	1.02	0.91	1.15	7.3E-01	1.0E+00
Triglycerides in Medium HDL	290	1.21	1.07	1.38	3.0E-03	5.4E-02	306	1.01	0.90	1.15	8.3E-01	1.0E+00
Phospholipids to Total Lipids in Large LDL percentage	290	0.84	0.75	0.94	3.2E-03	5.4E-02	306	1.01	0.91	1.13	8.3E-01	1.0E+00
Phospholipids to Total Lipids in Small HDL percentage	290	1.20	1.06	1.35	3.1E-03	5.4E-02	306	1.07	0.95	1.20	2.9E-01	1.0E+00

Appendix 5.9 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes, stratified by time since blood collection

	≤ 5 years						> 5 years					
Saturated Fatty Acids to Total Fatty Acids percentage	290	1.19	1.06	1.34	4.4E-03	6.5E-02	306	1.04	0.92	1.17	5.6E-01	1.0E+00
Total Lipids in Very Small VLDL	290	1.19	1.06	1.34	4.2E-03	6.5E-02	306	1.08	0.96	1.21	2.2E-01	1.0E+00
Free Cholesterol to Total Lipids in Small HDL percentage	290	1.17	1.05	1.30	5.0E-03	7.0E-02	306	1.01	0.90	1.13	9.0E-01	1.0E+00
Saturated Fatty Acids	290	1.18	1.05	1.33	5.6E-03	7.4E-02	306	1.03	0.91	1.16	6.8E-01	1.0E+00
Triglycerides in LDL	290	1.18	1.05	1.34	6.0E-03	7.5E-02	306	1.00	0.88	1.13	9.7E-01	1.0E+00
Phospholipids in HDL	290	1.18	1.05	1.34	8.0E-03	8.6E-02	306	1.08	0.95	1.22	2.4E-01	1.0E+00
Phosphoglycerides	290	1.18	1.04	1.33	7.9E-03	8.6E-02	306	1.05	0.93	1.18	4.6E-01	1.0E+00
Polyunsaturated Fatty Acids to Total Fatty Acids percentage	290	0.85	0.76	0.96	7.7E-03	8.6E-02	306	0.99	0.88	1.13	9.3E-01	1.0E+00
Free Cholesterol in HDL	290	1.18	1.04	1.33	9.8E-03	8.7E-02	306	1.07	0.95	1.21	2.7E-01	1.0E+00
Phosphatidylcholines	290	1.18	1.04	1.33	9.4E-03	8.7E-02	306	1.05	0.93	1.18	4.5E-01	1.0E+00
Omega-6 Fatty Acids to Total Fatty Acids percentage	290	0.86	0.76	0.96	9.6E-03	8.7E-02	306	0.99	0.87	1.12	8.5E-01	1.0E+00
Free Cholesterol in Very Small VLDL	290	1.17	1.04	1.32	9.7E-03	8.7E-02	306	1.08	0.96	1.21	2.1E-01	1.0E+00
Triglycerides to Total Lipids in Medium LDL percentage	290	1.17	1.04	1.32	8.4E-03	8.7E-02	306	1.00	0.89	1.13	9.8E-01	1.0E+00
Total Cholines	290	1.17	1.04	1.32	1.2E-02	9.2E-02	306	1.05	0.93	1.18	4.7E-01	1.0E+00
Citrate	290	1.17	1.04	1.32	1.1E-02	9.2E-02	306	1.17	1.04	1.32	8.5E-03	1.0E+00

Appendix 5.9 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes, stratified by time since blood collection

	≤ 5 years						> 5 years					
Concentration of Very Small VLDL Particles	290	1.17	1.04	1.31	1.1E-02	9.2E-02	306	1.08	0.96	1.21	2.1E-01	1.0E+00
Triglycerides to Total Lipids in Large LDL percentage	290	1.17	1.03	1.31	1.2E-02	9.2E-02	306	0.99	0.87	1.12	8.2E-01	1.0E+00
Linoleic Acid to Total Fatty Acids percentage	290	0.86	0.76	0.97	1.2E-02	9.4E-02	306	0.98	0.86	1.11	7.2E-01	1.0E+00
Sphingomyelins	290	1.16	1.03	1.31	1.4E-02	1.1E-01	306	1.07	0.95	1.21	2.6E-01	1.0E+00
Phospholipids in Large HDL	290	1.18	1.03	1.35	1.5E-02	1.1E-01	306	1.11	0.97	1.26	1.3E-01	1.0E+00
Total Phospholipids in Lipoprotein Particles	290	1.16	1.03	1.31	1.7E-02	1.1E-01	306	1.04	0.93	1.18	4.8E-01	1.0E+00
Total Lipids in HDL	290	1.16	1.03	1.31	1.8E-02	1.1E-01	306	1.07	0.95	1.22	2.5E-01	1.0E+00
Average Diameter for HDL Particles	290	1.15	1.02	1.29	1.9E-02	1.1E-01	306	1.08	0.96	1.21	2.2E-01	1.0E+00
Total Fatty Acids	290	1.16	1.02	1.30	1.8E-02	1.1E-01	306	1.02	0.90	1.15	7.6E-01	1.0E+00
Phospholipids in Medium HDL	290	1.17	1.03	1.33	1.7E-02	1.1E-01	306	1.07	0.95	1.22	2.7E-01	1.0E+00
Free Cholesterol to Total Lipids in Very Large HDL percentage	290	0.90	0.82	0.98	1.9E-02	1.1E-01	306	0.96	0.86	1.06	4.1E-01	1.0E+00
Tyrosine	290	1.15	1.02	1.29	2.1E-02	1.2E-01	306	1.01	0.90	1.14	8.7E-01	1.0E+00
Triglycerides in Medium LDL	290	1.15	1.02	1.30	2.2E-02	1.3E-01	306	0.98	0.87	1.11	7.8E-01	1.0E+00
Monounsaturated Fatty Acids	290	1.15	1.02	1.30	2.5E-02	1.4E-01	306	1.01	0.89	1.14	9.1E-01	1.0E+00

Appendix 5.9 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes, stratified by time since blood collection

	≤ 5 years						> 5 years					
Triglycerides to Total Lipids in IDL percentage	290	1.15	1.02	1.30	2.7E-02	1.4E-01	306	0.95	0.84	1.08	4.7E-01	1.0E+00
Cholesterol to Total Lipids in Very Large HDL percentage	290	0.94	0.88	0.99	2.6E-02	1.4E-01	306	0.98	0.89	1.08	6.9E-01	1.0E+00
Triglycerides in Small VLDL	290	1.15	1.01	1.30	2.9E-02	1.5E-01	306	0.96	0.85	1.09	5.3E-01	1.0E+00
Total Lipids in Large HDL	290	1.15	1.01	1.31	3.0E-02	1.5E-01	306	1.10	0.97	1.25	1.4E-01	1.0E+00
Free Cholesterol in Medium HDL	290	1.15	1.01	1.31	3.0E-02	1.5E-01	306	1.07	0.94	1.21	3.3E-01	1.0E+00
Concentration of Very Large HDL Particles	290	1.15	1.01	1.30	3.2E-02	1.6E-01	306	1.06	0.94	1.19	3.7E-01	1.0E+00
Polyunsaturated Fatty Acids to Monounsaturated Fatty Acids ratio	290	0.88	0.77	0.99	3.7E-02	1.7E-01	306	1.01	0.89	1.14	9.1E-01	1.0E+00
Cholesterol in Very Small VLDL	290	1.14	1.01	1.28	3.6E-02	1.7E-01	306	1.09	0.97	1.23	1.4E-01	1.0E+00
Free Cholesterol in Large HDL	290	1.17	1.01	1.36	3.6E-02	1.7E-01	306	1.11	0.97	1.28	1.3E-01	1.0E+00
Free Cholesterol to Total Lipids in IDL percentage	290	0.89	0.80	0.99	3.7E-02	1.7E-01	306	1.05	0.93	1.18	4.2E-01	1.0E+00
Triglycerides to Total Lipids in Small LDL percentage	290	1.14	1.01	1.29	3.5E-02	1.7E-01	306	0.98	0.87	1.11	7.5E-01	1.0E+00
Cholesteryl Esters to Total Lipids in	290	0.96	0.92	1.00	3.5E-02	1.7E-01	306	1.02	0.89	1.18	7.6E-01	1.0E+00

Appendix 5.9 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes, stratified by time since blood collection

	≤ 5 years						> 5 years					
Medium HDL percentage												
Total Lipids in Medium HDL	290	1.14	1.01	1.30	3.8E-02	1.7E-01	306	1.07	0.94	1.21	2.9E-01	1.0E+00
Apolipoprotein A1	290	1.14	1.01	1.29	4.0E-02	1.7E-01	306	1.06	0.94	1.20	3.3E-01	1.0E+00
Histidine	290	0.89	0.79	0.99	4.0E-02	1.7E-01	306	0.90	0.80	1.01	7.2E-02	1.0E+00
Concentration of Large HDL Particles	290	1.15	1.01	1.32	4.1E-02	1.7E-01	306	1.10	0.97	1.26	1.4E-01	1.0E+00
Triglycerides in Small HDL	290	1.14	1.00	1.29	4.3E-02	1.8E-01	306	0.97	0.86	1.10	6.3E-01	1.0E+00
Free Cholesterol to Total Lipids in Medium HDL percentage	290	1.16	1.00	1.33	4.4E-02	1.8E-01	306	1.04	0.91	1.19	5.6E-01	1.0E+00
Cholesterol to Total Lipids in IDL percentage	290	0.89	0.80	1.00	4.7E-02	1.8E-01	306	1.02	0.91	1.15	7.3E-01	1.0E+00
Free Cholesterol in Small HDL	290	1.13	1.00	1.28	4.9E-02	1.9E-01	306	1.01	0.90	1.14	8.7E-01	1.0E+00
Total Lipids in Lipoprotein Particles	290	1.12	0.99	1.26	6.9E-02	2.5E-01	306	1.02	0.91	1.15	7.6E-01	1.0E+00
Phenylalanine	290	1.12	0.99	1.26	6.8E-02	2.5E-01	306	0.94	0.84	1.06	3.2E-01	1.0E+00
Cholesteryl Esters in Very Small VLDL	290	1.12	0.99	1.26	6.7E-02	2.5E-01	306	1.10	0.98	1.23	1.2E-01	1.0E+00
Total Lipids in Very Large HDL	290	1.12	0.99	1.27	6.5E-02	2.5E-01	306	1.06	0.94	1.20	3.3E-01	1.0E+00
Triglycerides to Total Lipids in Medium HDL percentage	290	1.13	0.99	1.29	6.7E-02	2.5E-01	306	0.97	0.86	1.10	6.7E-01	1.0E+00
Concentration of Medium HDL Particles	290	1.13	0.99	1.28	7.1E-02	2.5E-01	306	1.07	0.94	1.21	3.1E-01	1.0E+00

Appendix 5.9 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes, stratified by time since blood collection

	≤ 5 years						> 5 years					
Triglycerides in Small LDL	290	1.12	0.99	1.26	7.5E-02	2.6E-01	306	0.98	0.87	1.11	7.6E-01	1.0E+00
Triglycerides to Total Lipids in Small HDL percentage	290	1.12	0.99	1.28	7.6E-02	2.6E-01	306	0.96	0.85	1.09	5.3E-01	1.0E+00
Concentration of Small VLDL Particles	290	1.11	0.99	1.26	8.4E-02	2.9E-01	306	0.99	0.88	1.12	9.0E-01	1.0E+00
Total Lipids in Small VLDL	290	1.11	0.99	1.26	8.5E-02	2.9E-01	306	1.00	0.89	1.12	1.0E+00	1.0E+00
Phospholipids in Very Large HDL	290	1.15	0.98	1.36	8.8E-02	2.9E-01	306	1.08	0.94	1.24	2.8E-01	1.0E+00
Concentration of VLDL Particles	290	1.11	0.98	1.25	9.5E-02	3.1E-01	306	1.02	0.90	1.14	7.8E-01	1.0E+00
Phospholipids to Total Lipids in Medium HDL percentage	290	1.13	0.98	1.32	9.9E-02	3.2E-01	306	1.00	0.86	1.16	1.0E+00	1.0E+00
Lactate	289	0.91	0.81	1.02	1.1E-01	3.5E-01	306	1.03	0.92	1.16	5.7E-01	1.0E+00
Pyruvate	290	1.11	0.98	1.25	1.1E-01	3.5E-01	305	1.11	0.98	1.25	1.0E-01	1.0E+00
Total Triglycerides	290	1.10	0.98	1.25	1.2E-01	3.6E-01	306	0.97	0.86	1.09	6.2E-01	1.0E+00
Degree of Unsaturation	290	0.91	0.81	1.02	1.2E-01	3.6E-01	306	1.01	0.90	1.14	8.2E-01	1.0E+00
Cholesterol to Total Lipids in Medium HDL percentage	290	0.97	0.93	1.01	1.2E-01	3.7E-01	306	1.03	0.89	1.19	6.8E-01	1.0E+00
Cholesterol in Large HDL	290	1.11	0.97	1.27	1.3E-01	3.8E-01	306	1.11	0.97	1.27	1.3E-01	1.0E+00
Phospholipids in Small HDL	290	1.10	0.97	1.25	1.3E-01	3.8E-01	306	1.02	0.90	1.15	7.7E-01	1.0E+00
Phospholipids to Total Lipids in	290	0.91	0.81	1.03	1.3E-01	3.8E-01	306	1.09	0.98	1.23	1.2E-01	1.0E+00

Appendix 5.9 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes, stratified by time since blood collection

	≤ 5 years						> 5 years					
Medium LDL percentage												
Triglycerides to Total Lipids in Small VLDL percentage	290	1.10	0.97	1.25	1.3E-01	3.8E-01	306	0.92	0.81	1.04	1.7E-01	1.0E+00
Free Cholesterol to Total Lipids in Very Large VLDL percentage	289	0.92	0.82	1.03	1.4E-01	4.1E-01	305	1.10	0.98	1.24	1.1E-01	1.0E+00
Cholesteryl Esters to Total Lipids in IDL percentage	290	0.92	0.82	1.03	1.5E-01	4.2E-01	306	1.00	0.89	1.13	9.5E-01	1.0E+00
Phospholipids to Total Lipids in Small VLDL percentage	290	0.92	0.81	1.03	1.5E-01	4.3E-01	306	1.06	0.94	1.20	3.7E-01	1.0E+00
Polyunsaturated Fatty Acids	290	1.09	0.97	1.23	1.6E-01	4.4E-01	306	1.02	0.91	1.15	7.4E-01	1.0E+00
Monounsaturated Fatty Acids to Total Fatty Acids percentage	290	1.09	0.96	1.24	1.6E-01	4.4E-01	306	0.98	0.87	1.11	7.7E-01	1.0E+00
Glutamine	290	0.92	0.82	1.03	1.6E-01	4.4E-01	305	0.93	0.83	1.05	2.3E-01	1.0E+00
Phospholipids to Total Lipids in Very Small VLDL percentage	290	1.09	0.96	1.23	1.7E-01	4.5E-01	306	1.01	0.90	1.14	8.4E-01	1.0E+00
HDL Cholesterol	290	1.09	0.96	1.24	1.7E-01	4.5E-01	306	1.07	0.95	1.21	2.6E-01	1.0E+00
Omega-6 Fatty Acids	290	1.09	0.96	1.22	1.7E-01	4.6E-01	306	1.02	0.90	1.14	7.9E-01	1.0E+00
Total Concentration of Lipoprotein Particles	290	1.09	0.96	1.23	1.8E-01	4.6E-01	306	1.04	0.92	1.18	5.3E-01	1.0E+00
Phospholipids in IDL	290	1.09	0.96	1.22	1.8E-01	4.6E-01	306	1.07	0.95	1.20	2.5E-01	1.0E+00

Appendix 5.9 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes, stratified by time since blood collection

	≤ 5 years						> 5 years					
Cholesterol to Total Lipids in Small VLDL percentage	290	0.92	0.82	1.04	1.8E-01	4.6E-01	306	1.09	0.96	1.23	1.9E-01	1.0E+00
Cholesteryl Esters in Large HDL	290	1.09	0.96	1.25	1.9E-01	4.7E-01	306	1.11	0.97	1.27	1.2E-01	1.0E+00
Cholesteryl Esters to Total Lipids in Small VLDL percentage	290	0.93	0.83	1.04	1.9E-01	4.7E-01	306	1.10	0.97	1.24	1.2E-01	1.0E+00
Phospholipids in Small VLDL	290	1.08	0.96	1.22	2.0E-01	4.8E-01	306	1.01	0.90	1.14	8.0E-01	1.0E+00
Free Cholesterol to Total Lipids in Large HDL percentage	290	1.16	0.93	1.45	2.0E-01	4.8E-01	306	1.10	0.91	1.33	3.2E-01	1.0E+00
Concentration of HDL Particles	290	1.08	0.96	1.23	2.1E-01	5.0E-01	306	1.04	0.92	1.18	5.1E-01	1.0E+00
Cholesterol in Very Large HDL	290	1.09	0.96	1.23	2.1E-01	5.0E-01	306	1.06	0.94	1.20	3.3E-01	1.0E+00
Cholesteryl Esters in Very Large HDL	290	1.10	0.95	1.28	2.2E-01	5.1E-01	306	1.09	0.94	1.26	2.5E-01	1.0E+00
Phospholipids in VLDL	290	1.07	0.95	1.21	2.4E-01	5.2E-01	306	1.01	0.90	1.13	8.7E-01	1.0E+00
Total Free Cholesterol	290	1.08	0.95	1.21	2.3E-01	5.2E-01	306	1.03	0.92	1.16	6.0E-01	1.0E+00
Docosahexaenoic Acid to Total Fatty Acids percentage	290	0.93	0.83	1.05	2.3E-01	5.2E-01	306	1.00	0.89	1.13	9.9E-01	1.0E+00
Valine	290	1.08	0.95	1.22	2.2E-01	5.2E-01	306	0.92	0.82	1.04	1.9E-01	1.0E+00
Acetoacetate	288	1.08	0.95	1.22	2.4E-01	5.2E-01	304	1.04	0.92	1.17	5.6E-01	1.0E+00
Triglycerides in Medium VLDL	290	1.08	0.95	1.22	2.4E-01	5.2E-01	306	0.96	0.85	1.07	4.5E-01	1.0E+00
Total Lipids in IDL	290	1.07	0.95	1.21	2.4E-01	5.2E-01	306	1.07	0.95	1.20	2.7E-01	1.0E+00

Appendix 5.9 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes, stratified by time since blood collection

	≤ 5 years						> 5 years					
Cholesterol in Medium HDL	290	1.08	0.95	1.23	2.3E-01	5.2E-01	306	1.07	0.94	1.22	2.9E-01	1.0E+00
Free Cholesterol to Total Lipids in Small VLDL percentage	290	0.93	0.83	1.05	2.3E-01	5.2E-01	306	1.05	0.93	1.19	4.1E-01	1.0E+00
Cholesterol to Total Lipids in Very Small VLDL percentage	290	0.93	0.83	1.05	2.4E-01	5.2E-01	306	1.07	0.95	1.22	2.7E-01	1.0E+00
Cholesteryl Esters to Total Lipids in Very Small VLDL percentage	290	0.93	0.83	1.05	2.3E-01	5.2E-01	306	1.07	0.95	1.22	2.6E-01	1.0E+00
Phospholipids to Total Lipids in Very Large HDL percentage	290	1.13	0.92	1.38	2.4E-01	5.2E-01	306	1.07	0.92	1.24	4.0E-01	1.0E+00
Average Diameter for LDL Particles	290	1.07	0.95	1.21	2.5E-01	5.2E-01	306	1.05	0.93	1.18	4.3E-01	1.0E+00
Free Cholesterol to Total Lipids in Large VLDL percentage	290	0.94	0.85	1.04	2.6E-01	5.4E-01	306	1.03	0.92	1.16	6.0E-01	1.0E+00
Total Lipids in VLDL	290	1.07	0.95	1.21	2.7E-01	5.5E-01	306	0.99	0.88	1.12	9.1E-01	1.0E+00
Free Cholesterol in Very Large HDL	290	1.08	0.95	1.22	2.7E-01	5.5E-01	306	1.06	0.93	1.20	3.8E-01	1.0E+00
Triglycerides to Total Lipids in Very Small VLDL percentage	290	1.07	0.95	1.22	2.7E-01	5.5E-01	306	0.91	0.80	1.03	1.3E-01	1.0E+00
Total Lipids in Small HDL	290	1.07	0.95	1.21	2.7E-01	5.6E-01	306	1.01	0.89	1.14	8.9E-01	1.0E+00
Cholesteryl Esters in Small VLDL	290	1.07	0.95	1.20	2.8E-01	5.6E-01	306	1.03	0.92	1.16	5.6E-01	1.0E+00
Isoleucine	290	1.07	0.95	1.20	2.8E-01	5.7E-01	306	0.94	0.83	1.05	2.6E-01	1.0E+00

Appendix 5.9 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes, stratified by time since blood collection

	≤ 5 years						> 5 years					
Triglycerides in VLDL	290	1.07	0.95	1.21	2.9E-01	5.7E-01	306	0.97	0.86	1.09	5.7E-01	1.0E+00
Remnant Cholesterol (Non-HDL, Non-LDL -Cholesterol)	290	1.07	0.95	1.20	2.9E-01	5.7E-01	306	1.05	0.93	1.18	4.4E-01	1.0E+00
Cholesterol in Small VLDL	290	1.07	0.95	1.20	2.9E-01	5.7E-01	306	1.03	0.92	1.16	6.0E-01	1.0E+00
Total Cholesterol	290	1.07	0.94	1.20	3.0E-01	5.9E-01	306	1.03	0.92	1.16	5.7E-01	1.0E+00
VLDL Cholesterol	290	1.06	0.94	1.20	3.1E-01	5.9E-01	306	1.03	0.92	1.16	6.3E-01	1.0E+00
Total Lipids in Medium VLDL	290	1.06	0.94	1.20	3.1E-01	5.9E-01	306	0.99	0.88	1.11	8.8E-01	1.0E+00
Free Cholesterol to Total Lipids in Small LDL percentage	290	0.97	0.93	1.02	3.1E-01	5.9E-01	306	1.03	0.87	1.22	7.5E-01	1.0E+00
Cholesteryl Esters in Small HDL	290	0.94	0.83	1.06	3.2E-01	5.9E-01	306	1.00	0.88	1.14	9.6E-01	1.0E+00
Free Cholesterol in VLDL	290	1.06	0.94	1.20	3.2E-01	5.9E-01	306	1.01	0.90	1.14	8.2E-01	1.0E+00
Cholesteryl Esters in VLDL	290	1.06	0.94	1.20	3.2E-01	5.9E-01	306	1.04	0.93	1.17	5.0E-01	1.0E+00
Omega-3 Fatty Acids	290	1.06	0.94	1.20	3.3E-01	6.0E-01	306	1.02	0.90	1.15	7.7E-01	1.0E+00
Free Cholesterol in Small VLDL	290	1.06	0.94	1.19	3.3E-01	6.0E-01	306	1.02	0.91	1.15	6.9E-01	1.0E+00
Cholesteryl Esters in HDL	290	1.06	0.94	1.21	3.4E-01	6.0E-01	306	1.07	0.95	1.21	2.7E-01	1.0E+00
Total Esterified Cholesterol	290	1.06	0.94	1.20	3.5E-01	6.2E-01	306	1.04	0.92	1.17	5.6E-01	1.0E+00
Phospholipids in Medium VLDL	290	1.06	0.94	1.19	3.4E-01	6.2E-01	306	1.01	0.90	1.14	8.1E-01	1.0E+00
Cholesteryl Esters in Medium HDL	290	1.06	0.93	1.21	3.5E-01	6.2E-01	306	1.07	0.94	1.22	2.9E-01	1.0E+00

Appendix 5.9 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes, stratified by time since blood collection

	≤ 5 years						> 5 years					
Phospholipids to Total Lipids in Large HDL percentage	290	1.10	0.90	1.35	3.6E-01	6.3E-01	306	0.99	0.88	1.12	9.2E-01	1.0E+00
Acetate	290	1.06	0.94	1.19	3.7E-01	6.4E-01	304	0.92	0.82	1.04	2.0E-01	1.0E+00
BCAA	290	1.06	0.94	1.19	3.7E-01	6.4E-01	306	0.92	0.82	1.04	1.7E-01	1.0E+00
Free Cholesterol in Medium VLDL	290	1.05	0.94	1.19	3.8E-01	6.5E-01	306	1.02	0.91	1.15	7.2E-01	1.0E+00
Phospholipids to Total Lipids in Very Large VLDL percentage	289	0.96	0.88	1.05	3.8E-01	6.5E-01	305	1.01	0.90	1.13	8.6E-01	1.0E+00
Phospholipids to Total Lipids in Small LDL percentage	290	1.07	0.92	1.23	3.9E-01	6.6E-01	306	1.08	0.93	1.25	3.1E-01	1.0E+00
Phospholipids in Large VLDL	290	1.06	0.93	1.21	4.0E-01	6.7E-01	306	0.98	0.88	1.09	7.5E-01	1.0E+00
Concentration of Large VLDL Particles	290	1.05	0.93	1.19	4.1E-01	6.8E-01	306	0.97	0.87	1.10	6.7E-01	1.0E+00
Linoleic Acid	290	1.05	0.93	1.18	4.1E-01	6.9E-01	306	1.00	0.89	1.13	9.4E-01	1.0E+00
Average Diameter for VLDL Particles	290	0.95	0.85	1.07	4.3E-01	7.0E-01	306	0.94	0.83	1.05	2.8E-01	1.0E+00
Cholesteryl Esters in Large VLDL	290	1.05	0.93	1.18	4.3E-01	7.0E-01	306	1.00	0.89	1.12	9.7E-01	1.0E+00
Concentration of Medium VLDL Particles	290	1.05	0.93	1.18	4.3E-01	7.0E-01	306	1.00	0.90	1.13	9.4E-01	1.0E+00
Phospholipids to Total Lipids in Chylomicrons and Extremely Large VLDL percentage	282	1.06	0.92	1.23	4.3E-01	7.0E-01	301	0.95	0.86	1.04	2.9E-01	1.0E+00

Appendix 5.9 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes, stratified by time since blood collection

	≤ 5 years						> 5 years					
Triglycerides to Total Lipids in Very Large VLDL percentage	289	1.07	0.91	1.25	4.3E-01	7.0E-01	305	0.97	0.89	1.04	3.9E-01	1.0E+00
Free Cholesterol to Total Lipids in Very Small VLDL percentage	290	0.96	0.85	1.07	4.4E-01	7.0E-01	306	1.03	0.91	1.16	6.2E-01	1.0E+00
Free Cholesterol in Small LDL	290	0.97	0.90	1.05	4.4E-01	7.1E-01	306	1.01	0.89	1.15	8.7E-01	1.0E+00
Cholesteryl Esters in IDL	290	1.05	0.93	1.18	4.5E-01	7.1E-01	306	1.06	0.94	1.20	3.2E-01	1.0E+00
Triglycerides in Large VLDL	290	1.05	0.93	1.19	4.6E-01	7.2E-01	306	0.95	0.85	1.07	4.2E-01	1.0E+00
Phospholipids to Total Lipids in IDL percentage	290	1.05	0.93	1.18	4.6E-01	7.2E-01	306	0.99	0.89	1.11	9.3E-01	1.0E+00
Total Lipids in Large VLDL	290	1.05	0.93	1.18	4.7E-01	7.2E-01	306	0.97	0.86	1.09	5.9E-01	1.0E+00
Cholesterol in IDL	290	1.04	0.92	1.18	4.9E-01	7.6E-01	306	1.07	0.95	1.20	2.9E-01	1.0E+00
Free Cholesterol to Total Lipids in Medium LDL percentage	290	0.98	0.93	1.03	5.0E-01	7.6E-01	306	1.09	0.88	1.35	4.3E-01	1.0E+00
Triglycerides in Very Large VLDL	290	1.04	0.92	1.18	5.1E-01	7.6E-01	306	0.98	0.88	1.10	7.8E-01	1.0E+00
Cholesterol in Large VLDL	290	1.04	0.92	1.18	5.1E-01	7.6E-01	306	0.99	0.88	1.11	8.1E-01	1.0E+00
Phospholipids to Total Lipids in Large VLDL percentage	290	1.05	0.90	1.23	5.1E-01	7.6E-01	306	1.00	0.90	1.12	9.7E-01	1.0E+00

Appendix 5.9 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes, stratified by time since blood collection

	≤ 5 years						> 5 years					
Triglycerides to Total Lipids in Large HDL percentage	290	1.04	0.92	1.19	5.1E-01	7.6E-01	306	0.95	0.84	1.07	3.8E-01	1.0E+00
Concentration of Very Large VLDL Particles	290	1.04	0.92	1.18	5.3E-01	7.7E-01	306	0.98	0.87	1.10	6.8E-01	1.0E+00
Triglycerides to Total Lipids in Very Large HDL percentage	290	1.04	0.92	1.18	5.2E-01	7.7E-01	306	0.97	0.86	1.09	6.0E-01	1.0E+00
Total Cholesterol Minus HDL-C	290	1.04	0.92	1.17	5.3E-01	7.8E-01	306	1.03	0.91	1.15	6.7E-01	1.0E+00
Triglycerides to Total Lipids in Medium VLDL percentage	290	1.04	0.92	1.18	5.3E-01	7.8E-01	306	0.92	0.84	1.01	8.4E-02	1.0E+00
Creatinine	280	0.96	0.86	1.08	5.5E-01	7.9E-01	304	0.89	0.79	1.00	4.3E-02	1.0E+00
Cholesterol to Total Lipids in Very Large VLDL percentage	289	0.96	0.85	1.09	5.5E-01	7.9E-01	305	1.07	0.95	1.20	2.6E-01	1.0E+00
Apolipoprotein B	290	1.04	0.92	1.17	5.6E-01	8.0E-01	306	1.03	0.91	1.15	6.7E-01	1.0E+00
Cholesteryl Esters to Total Lipids in Medium VLDL percentage	290	0.97	0.87	1.08	5.6E-01	8.0E-01	306	1.06	0.93	1.20	3.9E-01	1.0E+00
Apolipoprotein B to Apolipoprotein A1 ratio	290	0.97	0.86	1.09	5.7E-01	8.0E-01	306	0.99	0.88	1.12	9.0E-01	1.0E+00
Concentration of Large LDL Particles	290	1.03	0.92	1.16	5.8E-01	8.2E-01	306	1.02	0.91	1.15	7.5E-01	1.0E+00
Cholesteryl Esters in Large LDL	290	1.04	0.91	1.18	5.8E-01	8.2E-01	306	1.02	0.90	1.15	7.6E-01	1.0E+00

Appendix 5.9 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes, stratified by time since blood collection

	≤ 5 years						> 5 years					
Triglycerides to Phosphoglycerides ratio	290	1.03	0.91	1.17	6.1E-01	8.2E-01	306	0.94	0.83	1.07	3.4E-01	1.0E+00
Omega-6 Fatty Acids to Omega-3 Fatty Acids ratio	290	0.97	0.86	1.09	6.1E-01	8.2E-01	306	0.99	0.88	1.11	8.3E-01	1.0E+00
Acetone	290	0.97	0.86	1.09	6.0E-01	8.2E-01	306	0.88	0.78	0.99	4.0E-02	1.0E+00
Free Cholesterol in Large VLDL	290	1.03	0.91	1.17	6.0E-01	8.2E-01	306	0.97	0.87	1.09	6.6E-01	1.0E+00
Total Lipids in Large LDL	290	1.03	0.92	1.16	5.9E-01	8.2E-01	306	1.02	0.91	1.15	7.3E-01	1.0E+00
Cholesterol to Total Lipids in Medium VLDL percentage	290	0.97	0.86	1.09	6.1E-01	8.2E-01	306	1.07	0.95	1.21	2.5E-01	1.0E+00
Free Cholesterol to Total Lipids in Large LDL percentage	290	0.99	0.94	1.04	6.1E-01	8.2E-01	306	1.11	0.85	1.44	4.5E-01	1.0E+00
Total Lipids in Very Large VLDL	290	1.03	0.91	1.17	6.2E-01	8.2E-01	306	0.98	0.88	1.11	7.8E-01	1.0E+00
Cholesterol in Medium VLDL	290	1.03	0.92	1.16	6.2E-01	8.2E-01	306	1.03	0.92	1.16	5.7E-01	1.0E+00
Free Cholesterol in IDL	290	1.03	0.91	1.16	6.3E-01	8.4E-01	306	1.07	0.95	1.21	2.3E-01	1.0E+00
Triglycerides in Chylomicrons and Extremely Large VLDL	290	0.97	0.87	1.09	6.4E-01	8.5E-01	306	0.99	0.89	1.11	8.9E-01	1.0E+00
Cholesteryl Esters in Very Large VLDL	290	1.03	0.91	1.16	6.5E-01	8.5E-01	306	1.00	0.89	1.12	9.7E-01	1.0E+00
Concentration of IDL Particles	290	1.03	0.91	1.16	6.5E-01	8.5E-01	306	1.10	0.98	1.24	1.1E-01	1.0E+00

Appendix 5.9 Hazard ratios for prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features in the UK Biobank[1], after additional adjustment for diabetes, stratified by time since blood collection		
	≤ 5 years	> 5 years

Abbreviations: HR, hazard ratio; LL, 95% confidence interval lower limit; UL, 95% confidence interval upper limit; LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol; IDL, intermediate-density lipoprotein cholesterol; VLDL, very low-density lipoprotein cholesterol; BCAA, Total concentration of branched-chain amino acids (Leucine + Isoleucine + Valine).

1 Hazard ratios (95% confidence interval) for risk of prostate cancer death associated with a one standard deviation increase in concentration, percentage, or ratio of metabolite features by time since blood collection, with additional adjustment for diabetes. Analyses were stratified by region and Townsend Deprivation Index with age as the underlying time variable. Adjusted for diabetes, fasting status, body mass index, height, alcohol intake, family history of prostate cancer, ethnicity, level of education, smoking status, and weekly physical activity.

2 Padj < 0.05 are statistically significant after allowing for multiple testing using a false discovery rate controlling procedure at α = 0.05 (Benjamini-Hochberg). Padj values in the ≤ 5 years group are sorted in ascending order. Association with Padj < 0.05 in both the overall risk of death dataset and this dataset are bolded.