# This electronic thesis or dissertation has been downloaded from Explore Bristol Research, http://research-information.bristol.ac.uk 

Author:
Sandu, Meda R
Title:
Molecular epidemiology of Prostate Cancer progression

## General rights

Access to the thesis is subject to the Creative Commons Attribution - NonCommercial-No Derivatives 4.0 International Public License. A copy of this may be found at https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode This license sets out your rights and the restrictions that apply to your access to the thesis so it is important you read this before proceeding.

## Take down policy

Some pages of this thesis may have been removed for copyright restrictions prior to having it been deposited in Explore Bristol Research. However, if you have discovered material within the thesis that you consider to be unlawful e.g. breaches of copyright (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please contact collections-metadata@bristol.ac.uk and include the following information in your message:

[^0]Your claim will be investigated and, where appropriate, the item in question will be removed from public view as soon as possible.

# Molecular epidemiology of Prostate Cancer progression 

Meda Ramona Sandu

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Health Sciences.

Bristol Medical School: Population Health Sciences.

January 2023

Word count: 50,649


#### Abstract

Localised Prostate Cancer ( PCa ) is a slow-growing and heterogenous disease, with both indolent and aggressive phenotypes, which are very difficult to differentiate at diagnosis. Circulating biomarkers measured at diagnosis that distinguish indolent from lethal PCa are yet to be established and could be used clinically to avoid both over- (for indolent disease) or under- (for aggressive cancer) treatment.

This thesis investigated whether circulating metabolites, measured by nuclear magnetic resonance (NMR), are observationally associated with PCa progression (clinical progression, PCa metastases or PCa death) in a cohort of patients with localised disease and eligible to be randomised into a controlled trial (RCT). Using genetic data (Mendelian randomization), the thesis also examined if circulating metabolites are causally linked to PCa mortality. The thesis next explored whether the circulating metabolome of men with localised PCa, who had been randomised to lifestyle (brisk walking) and dietary (lycopene supplementation or increased fruit and vegetable and reduced dairy milk intake) interventions, was altered by these interventions and if, in turn, the altered metabolites were causally linked to PCa death. Finally, the thesis also evaluated the performance of two metabolomic risk scores, previously developed in the general population, at predicting PCa death in men with prostate specific antigen (PSA) detected localised PCa, in a clinical setting.

In the observational analyses, there was suggestive evidence that some metabolites (lipoproteins, cholesterol, glycolysis, fluid balance and inflammation markers) were associated with PCa progression. In the Mendelian randomization genetic analysis, there was evidence that the following metabolites causally increased PCa mortality: total-, freeand esterified-cholesterol, some measures of intermediate-, low- and very low-density lipoprotein cholesterol, sphingomyelins, apolipoprotein B, omega-3 fatty acids, docosahexaenoic acid and valine. Some high-density and very low-density lipoprotein related measures, histidine and the ratio of omega- 6 to omega- 3 fatty acids were causally associated with decreased PCa mortality. There was some evidence that the randomised lifestyle and dietary interventions may alter lipids, alanine, lactate and pyruvate. In the Mendelian randomization causal analysis, there was evidence that some of the altered metabolites in the brisk walking intervention (cholesterol esters in medium VLDL, phospholipids in chylomicrons and extremely large VLDL) may be linked to increased PCa


mortality risk. Lastly, there was evidence that two metabolomic risk scores developed in the general population did not predict PCa or all-cause mortality in a clinical setting This thesis supports a role for circulating metabolites in PCa progression, with implications for diagnostic and interventional research. Further studies of localised PCa with larger sample sizes and staging information are however needed to establish the role that metabolites, particular lipids, play in PCa progression and to allow for clinical translation.

## Acknowledgements

To all my supervisors, thank you. I am still in awe of how supportive, understanding and empathic you have been. You have gone over and beyond what would be expected and your constant support on so many levels has been the greatest of motivator and what allowed me to get to this day. Thank you to University of Bristol medical and dental school staff, who have been incredibly understanding of my bereavement, allowed me to extend my studies in a very supportive and guilt free manner, you all propped me up throughout the hardest of times (particularly huge thanks to the star Sharen Hockey-O'Keefe!).

Richard Martin- thank you for showing so much enthusiasm about my projects and for teaching me how to think of the bigger picture. Even when I thought there was no way to fix something, you always managed to come up with a creative solution.

Athene Lane - thank you for your pragmatism, your words would resonate in my mind while I was writing and guided me to not get lost.

Lucy McGeagh - thank you for your incredible dedication, reading my work even when you were on maternity leave. Shared grief is halved; thank you for your astonishing empathy.

Paul Yousefi - first, thank you for agreeing to join an already 5-person supervisory team. Thank you for your immense input, I cannot believe how much I learnt from you and how easy you make complex topics sound. Thank you for always being available and finding the patience to go through the smallest of details.

Rebecca Richmond - Thank you for your incredible support, for helping me from the beginning to find my feet and having the unbelievable patience to teach me MR from scratch.

Diana Santos Ferreira - thank you for teaching me everything I know about metabolomics. I cannot imagine anyone better to do this.

It is impossible to acknowledge everyone who has help/picked me up in these last 5 years. I am forever grateful for all that you have done, and I hope throughout my life to be able to show it.

Osama Mahmoud - thank you for your incredible support in the Slope-hunter correction, it was incredibly helpful and you made it all sound so simple.

Rhona Beynon - thank you for answering all my questions on IV and MR and Cox and just pretty much everything and for co-authoring the MR methodology paper. Thanks for listening for hours how amazing my dogs are.

Fiona Kinnear - I do not know where to start or end, you have been a wonderful friend and colleague and I really think you were the one person who could talk some sense into me. Thank you for it all, including the ridiculous amount of sugary deliveries.

Natasha Hakimi - I cannot believe you actually listened to all my voice notes.

Richard Townley - Thanks for giving up all our weekends and evenings to this cause.
Nancy McBride - Thank you for always having codes for everything and most importantly for the banter.

Sofia Leadbetter - Thank you for always sorting all my admin issues, of which we know I had many. But most importantly thank you for looking after "our" dogs.

Vanessa Marshall - Thank you for always having my back. The sense of security, support and belonging that brough to me was quite unbelievable, particularly at times where things were really tough.

Thank you to all my friends and family who supported and kept me sane though this period, with the best hug awards going to Polly, Rusa, my protectors and therapy mutts and Lars, my perfect nephew.

Lastly, thank you Tim, for believing in me more than I will ever believe in myself, I am incredibly grateful for having you in my life even if it was for such a cruelly short amount of time.

This thesis is dedicated to Major Tim Addison RM, my beloved late husband
"Take risks, deliver more, fear nothing"

## COVID-19 statement

The purpose of this statement is to provide some detail on how my thesis has been impacted by the COVID-19 pandemic. It presents personal and professional circumstances which changed as a result of the pandemic and how these have impacted the final thesis content and timeline. This statement is not suggesting that the quality of the thesis has been affected or would have been different if the pandemic had not occurred.

In 2020, the UK saw multiple lockdowns which restricted the capacity of the laboratories of the University of Bristol as well as the types of work they were able to conduct. This affected one of my chapters, which I have no longer included in my thesis. During the first year of my PhD, I helped set-up and run one clinical trial and continued to support the physical activity intervention administrative tasks and metabolomic analysis (liaising with the laboratory and producing all required documentation) throughout my PhD. Samples collected from the trial would be quantified for metabolomic analysis at baseline and followup to assess the impact of the randomised intervention (metformin and physical activity) in men with localised Prostate cancer, as well as investigate how the timing of blood sample processing (the aim was 1.5 hours) impacted on the glycolysis metabolites. Due to the COVID-19 pandemic, the laboratory at the University of Bristol was unable to process a large portion of the samples, particularly the follow-up ones, since it was uncertain whether the samples were contaminated with COVID-19. As such, it was impossible to include the work I had done to date without the metabolomic quantification. Instead, Chapter 7 of this thesis was expanded to account for the loss, for which a 3-month extension was granted.

On a personal level, after the first lockdown, I very unexpectedly lost my husband in 2020. I was unable to work and had to take a leave of absence for 6 months. The resulting trauma and suffering impacted my cognitive ability, to such extent that my productivity was very low and I had to have a phased returned to my studies. I changed from full-time to part time study and another extension of 6 -months was granted. Since then, I have been slowly recovering. I completed a five-month fellowship at University of Bristol, and I now work privately as a research scientist whilst finalising my thesis.

I have been incredibly lucky to be a student at the University of Bristol, who have been unbelievable supportive. Everyone has shown so much understanding, compassion and encouragement and I never felt pressured or too overwhelmed. Despite the terrible circumstances, I believe this thesis kept me engaged and throughout I put the maximum of
efforts to get it done. I am very proud of it and of what I have manged to endure in these last few years.

Thank you for being my examiners.

## Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of others, is indicated as such. Any views expressed in the dissertation are those of the author.

Signed:


Date: 03/01/2023

## Research outputs included in this thesis

At the time of submission of this thesis, one methodological manuscript has been published as pre-print and submitted for publication and relates to Chapters 4,5 and 6 . Chapters 5 and 6 have each been written up and Chapter 7 will be written up as individual manuscripts and will be submitted for publication in peer reviewed journals.

Sandu* ${ }^{*}$ M.R.; Beynon*, R.; Richmond, R.; Santos Ferreira, D.L.; Hackshaw-McGeagh, L.; Davey Smith, G.; Metcalfe, C.; Lane, J.A.; Martin, R. Two-step Randomisation: Applying the Results of Small Feasibility Studies of Interventions to Large-scale Mendelian Randomisation Studies to Robustly Infer Causal Effects on Clinical Endpoints. Preprints 2019, 2019100276 (doi:10.20944/ preprints201910.0276.v1).<br>*Denotes authors with equal contributions

## Research outputs associated but not included in this thesis

Wade, K. H., Yarmolinsky, J., Giovannucci, E., Lewis, S. J., Millwood, I. Y., Munafò, M. R., Meddens, F., Burrows, K., Bell, J. A., Davies, N. M., Mariosa, D., Kanerva, N., Vincent, E. E., Smith-Byrne, K., Guida, F., Gunter, M. J., Sanderson, E., Dudbridge, F., Burgess, S., Cornelis, M. C., Richardson, T.G., Borges, M. C., Bowden J., Hemani, G., Cho, Y., Spiller, W., Richmond, R. C., Carter, A. R., Langdon, R., Lawlor D.A., Walters, R. G., Vimaleswaran, K. S., Anderson, A., Sandu M.R., Tiling, K., Davey-Smith, G., Martin, R. M., Relton, C. L., with the M. R. in Nutrition, Cancer working group (2022). Applying Mendelian randomization to appraise causality in relationships between nutrition and cancer. Cancer causes $\mathcal{E}$ control : CCC, 33(5), 631-652. https://doi.org/10.1007/s10552-022-01562-1

McGeagh, L., Robles, L. A., Persad, R., Rowe, E., Bahl, A., Aning, J., Koupparis, A., Abrams, P., Perks, C., Holly, J., Johnson, L., Shiridzinomwa, C., Challapalli, A., Shingler, E., Taylor, H., Oxley, J., Sandu, M., Martin, R. M., \& Lane, J. A. (2022). Prostate cancer-Exercise and Metformin Trial (Pre-EMpT): study protocol for a feasibility factorial randomized controlled trial in men with localised or locally advanced prostate cancer. Pilot and feasibility studies, 8(1), 179. https:// doi.org/10.1186/s40814-022-01136-7

## Table of Contents

Chapter 1. Overview of thesis ..... 1
1.1. Thesis rationale ..... 1
1.2. Aims and Objectives ..... 3
1.3. Thesis outline ..... 4
1.4. Datasets used ..... 8
1.5. Statistical Methods ..... 9
Chapter 2. Prostate Cancer epidemiology and lifestyle factors. ..... 10
2.1. Introduction ..... 10
2.2. Prostate Cancer epidemiology ..... 10
2.2.1. PCa prevalence ..... 10
2.2.2. Prostate Cancer incidence rates ..... 11
2.2.3. Prostate Cancer mortality rates and survival. ..... 13
2.3. Clinical aspects of Prostate Cancer ..... 18
2.4. Diagnosis of Prostate Cancer in the UK ..... 19
2.4.1. Treatment of Prostate Cancer ..... 22
2.4.2. Side effects of prostate biopsy and of treatment of localised prostate cancer ..... 23
2.5. Risk and protective factors for Prostate Cancer progression ..... 24
2.5.1. Dietary Factors. ..... 25
2.5.2. Prostate Cancer and physical activity ..... 28
2.6. Conclusions ..... 30
Chapter 3. Introduction to metabolomics ..... 32
3.1. Introduction ..... 32
3.2. The field of metabolomics ..... 32
3.2.1. Untargeted metabolomics ..... 33
3.2.2. Targeted metabolomics ..... 34
3.3. Metabolomics in cancer ..... 36
3.4. Metabolomics, lifestyle and diet ..... 37
3.5. Metabolomics and Prostate Cancer progression ..... 38
3.5.1. Circulating metabolites. ..... 38
3.5.2. Non-circulating metabolites ..... 52
3.6. Chapter summary ..... 57
Chapter 4. Methodology ..... 58
4.1. Chapter introduction ..... 58
4.2. Study populations ..... 58
4.2.1. The Cluster randomised triAl of PSA testing for Prostate cancer and The Prostate Testing for Cancer and Treatment trials ..... 58
4.2.2. The Prostate cancer: Evidence of Exercise and Nutrition Trial ..... 65
4.2.3. United Kingdom Biobank (UKBB) ..... 68
4.2.4. The Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL) consortium ..... 68
4.3. Observational statistical methods ..... 69
4.3.1. Metabolomic methods ..... 69
4.3.2. Survival analysis ..... 70
4.3.3. Predictive biomarker validation analysis ..... 75
4.4. Causality statistical methods ..... 81
4.4.1. Intention to treat analysis ..... 81
4.4.2. Introduction to instrumental variable analysis ..... 82
4.4.3. Instrumental variable analysis in RCTs ..... 83
4.4.4. Introduction to Mendelian Randomisation (MR) ..... 84
4.4.5. Mendelian Randomisation techniques used in this thesis ..... 88
4.4.6. Mendelian Randomisation in feasibility trials. ..... 91
4.5. Chapter summary ..... 96
Chapter 5. Assessment of systemic metabolic biomarkers to detect PCa progression in the Prostate Testing for Cancer and Treatment (ProtecT) trial. ..... 97
5.1. Chapter overview ..... 97
5.2. Research questions ..... 98
5.3. Methods ..... 98
5.3.1. Study population ..... 98
5.3.2. Measures ..... 99
5.3.3. Statistical analyses ..... 100
5.4. Results ..... 102
5.4.1. Baseline descriptive analysis ..... 102
5.4.2. Survival analysis ..... 104
5.4.3. Causal analysis using Mendelian randomisation ..... 112
5.5. Discussion. ..... 121
5.5.1. Main findings ..... 121
5.5.2. Strengths and limitations ..... 127
5.6. Chapter Summary ..... 128
Chapter 6. Effects of exercise and nutrition on the metabolome of men with localised Prostate Cancer: the Prostate Cancer Evidence of Exercise and Nutrition trial (PrEvENT) randomised controlled trial and a Mendelian randomisation analysis ..... 130
6.1. Chapter overview ..... 130
6.2. Research questions ..... 131
6.3. Methods ..... 132
6.3.1. Study population ..... 132
6.3.2. Measures ..... 132
6.3.3. Statistical analyses ..... 133
6.4. Results ..... 139
6.4.1. Descriptive analysis ..... 139
6.4.2. Intention-to-treat analysis ..... 143
6.4.3. Instrumental variable analysis ..... 150
6.4.4. Mendelian Randomisation analysis ..... 151
6.5. Discussion ..... 155
6.5.1. Main findings ..... 156
6.5.2. Strengths and Limitations ..... 159
6.5.3. Implications for future research ..... 162
6.6. Chapter Summary ..... 164
Chapter 7. Evaluating the performance of multi-metabolomic risk scores to predict PCa mortality ..... 165
7.1. Chapter overview. ..... 165
7.2. Research questions ..... 166
7.3. Methods ..... 166
7.3.1. Study population ..... 166
7.3.2. Measures ..... 167
7.3.3. Statistical analyses ..... 168
7.4. Results ..... 169
7.4.1. Baseline descriptive analyses ..... 169
7.4.2. Individual metabolites ..... 171
7.4.3. Multi-metabolite risk scores ..... 176
7.5. Discussion ..... 184
7.5.1. Summary of findings ..... 184
7.5.2. Findings in context ..... 186
7.5.3. Strengths and weaknesses ..... 188
7.6. Chapter summary ..... 188
Chapter 8. Discussion ..... 190
8.1. Chapter introduction ..... 190
8.2. Summary of key findings ..... 190
8.3. Strengths and weaknesses of this research ..... 195
8.3.1. Data sources ..... 195
8.3.2. Randomised controlled trial design ..... 195
8.3.3. Mendelian randomisation ..... 196
8.3.4. Interventional feasibility randomised controlled trial ..... 197
8.3.5. Triangulation of evidence ..... 197
8.3.6. Missing data ..... 198
8.3.7. Bias and confounding ..... 198
8.4. Future work ..... 201
8.4.1. Prognostic biomarker development ..... 201
8.4.2. Metformin and Prostate Cancer progression ..... 201
8.5. Concluding remarks ..... 202
8.6. References ..... 203
Chapter 9. Appendices. ..... 240
9.1. Appendix A ..... 240
9.2. Appendix B ..... 250
9.3. Appendix C ..... 385
9.4. Appendix D ..... 519

## Abbreviations used in this thesis

| Abbreviation | Term |
| :---: | :---: |
| All-cause mRS | All-cause mortality metabolomic risk score |
| ASAP | Atypical small acinar proliferation |
| AUROC | Area under the receiver operating curve |
| BMI | Body mass index |
| CAP | Cluster randomised triAl of PSA testing for Prostate cancer |
| CONSORT | Consolidated Standards of Reporting Trial |
| Da | Daltons |
| FA | Fatty acid |
| FN | False negative |
| FP | False positive |
| FPF | False positive fraction |
| GP | General practitioner |
| GWAS | Genetic wide association study |
| HDL | High-density lipoprotein cholesterol |
| HR | Hazard ratio |
| IDL | Intermediate-density lipoprotein |
| IGF-1 | Insulin-like growth factor-1 |
| ITT | Intention to treat analysis |
| IV | Instrumental variable analysis |


| IVW | Inverse variance weighted |
| :--- | :--- |
| KM | Kaplan Meier |
| LDL | Low-density lipoprotein cholesterol |
| MR | Mendelian randomisation |
| MREC | Multicentre Research Ethics Committee |
| MRI | Magnetic Resonance Imaging |
| MS | Mass spectrometry |
| NHS | National Health Service |
| NICE | National Institute for Health and Care |
| NMR | Excellence |
| NPV | Nuclear magnetic resonance |
| OR | Negative predictive value |
| PIN | Prostatic intraepithelial neoplasia |
| PPCA | Prostate Cancer Association Group to |
| PRACTICAL | Prostate cancer mortality metabolomic |
| PCa mRS | Investigate Cancer Associated Alterations |
| in the Genome |  |
|  | Prostate Cancer |
| Pritition trial |  |


| ProMPT | Prostate Cancer: Mechanisms of |
| :--- | :--- |
|  | Progression and Treatment |
| ProtecT | Prostate Testing for Cancer and |
|  | Treatment |
| PSA | Prostate Specific Antigen |
| RCT | Randomised controlled trial |
| ROC | Receiver operating curve |
| SD | Standard deviation |
| SNP | Single nucleotide polymorphism |
| TN | True negative |
| TNM | Tumour node metastasis |
| TP | True positive |
| UKBB | UK Biobank |
| VLDL | Very low-density lipoprotein |
| WCRF | World Cancer Research Fund |

## List of Figures

Figure 1.1: Diagram of Thesis research questions ..... 6
Figure 1.2 Schematic Thesis summary ..... 7
Figure 2.1: Heatmap of the estimated proportion of prevalent Prostate Cancer cases globally. ..... 11
Figure 2.2: Prostate Cancer age-standardized incidence rates globally, 2020. ..... 13
Figure 2.3: Map of Prostate Cancer age-standardized mortality rates globally, 2020 ..... 14
Figure 2.4: Prostate Cancer age-standardized mortality rates globally, 2020. ..... 15
Figure 2.5: Bar chart of Prostate Cancer net survival by age, 2009-2013, England ..... 16
Figure 2.6: Average number of deaths per year, by age group, per 100,000 population, UK, 2016-2018 ..... 17
Figure 2.7: European age-standardised mortality rates per 100,000 population, by age group, in the UK, 1971-2018. ..... 17
Figure 2.8: One-year net survival for Prostate Cancer in England at local authority level, 2019. ..... 18
Figure 3.1: Definitions of measures of test performance in the context of Gas chromatography or liquid chromatography in metabolomics ..... 34
Figure 3.2: Selection process for inclusion of publications into the literature search ..... 40
Figure 3.3: Net citrate production metabolic pathway ..... 53
Figure 4.1: The overlap between the CAP and embedded ProtecT trial. ..... 60
Figure 4.2: Flow diagram of the Prostate Testing for Cancer and Treatment and the Cluster randomised triAl of PSA testing for Prostate cancer trial from invitation to diagnosis ..... 62
Figure 4.3: The Consolidated Standards of Reporting Trials diagram of the The Prostate cancer: Evidence of Exercise and Nutrition Trial. ..... 66
Figure 4.4: A receiver operating curve for Test 1 and 2 for outcome $X$ ..... 78
Figure 4.5: Theoretical calibration curves examples ..... 79
Figure 4.6: Diagram of the three IV assumptions ..... 83
Figure 4.7: Two-stage randomisation in feasibility trials diagram ..... 95
Figure 5.1: Kaplan Meier plot of Prostate Cancer clinical progression by categorical confounding variables in The Prostate Testing for Cancer and Treatment trial (Gleason score grouping, weight category, Tumour, Node andMetastasis stage, age group) ..... 105

Figure 5.2: Kaplan Meier plots of Prostate Cancer metastases and death by categorical confounding variables in the The Prostate Testing for Cancer and Treatment trial (Gleason score grouping, weight category, Tumour, Node and Metastasis stage, age group)

Figure 5.3: Forest plot of hazard ratios for clinical progression and metastases or death, in
the minimally adjusted Cox models..................................................................................... 108
Figure 5.4: Forest plot of MR estimates (odds ratios) using IVW or Wald ratio and their $95 \% \mathrm{CI}$ for the effects of metabolites on PCa mortality among 67,758 PCa cases in the PRACTICAL consortium, using instruments developed in the UKBB GWAS

Figure 5.5: Scatter plot of the logged odds ratios of developing metastases or Prostate Cancer death in the observational (Prostate Testing for Cancer and Treatment) and logged odds ratios for Prostate Cancer mortality in the MR (Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome) studies.

Figure 5.6: Forest plot of odds ratio for metastases or death in the Prostate Testing for Cancer and Treatment trial and odds ratio of Prostate Cancer mortality in the Mendelian Randomisation analysis in the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome consortium.

Figure 6.1: The two-stage randomisation process in the Prostate Cancer Evidence of Exercise and Nutrition trial and Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome consortium. 138

Figure 6.2: Forest plots of the overall effects (Intention-to-treat) of the intervention vs control
on serum metabolites, by intervention arm in the unadjusted and adjusted analyses........ 145 on serum metabolites, by intervention arm in the unadjusted and adjusted analyses.......... 145
Figure 6.3: Comparison of the overall effects of the interventions on metabolites from the
Intention-to-treat analysis intervention arm vs control................................................... 149
Figure 6.4: Causal effects (odds ratio) of the altered metabolites in the Prostate Cancer Evidence of Exercise and Nutrition trial on Prostate Cancer mortality using data from the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome consortium.

Figure 7.1: Hazard ratios (per 1 standard deviation) and their $95 \%$ confidence intervals for associations between individual metabolites and Prostate Cancer specific mortality (2,059 survivors, 34 deaths) in minimally* and fully adjusted** Cox regression models in the Prostate Testing for Cancer and Treatment cohort.
Figure 7.2: Hazard ratios (per 1SD) and their $95 \%$ confidence intervals for the associations between individual biomarkers and all-cause mortality ( 1,864 survivors and 229 deaths) in minimally* and fully adjusted** Cox regression models

Figure 7.3: Hazard ratios (per 1 standard deviation) and $95 \%$ confidence intervals for the association of Prostate Cancer mortality metabolomic risk score model and all-cause mortality multi-metabolite risk score in cases with Prostate Cancer specific (A) and all-cause mortality (B) in the minimally* and fully adjusted** models and in controls and pooled participants (C).

Figure 7.4: Receiver operating curves for the Prostate Cancer mortality metabolomic risk score model and all-cause mortality multi-metabolite risk score at discriminating PCaspecific and all-cause mortality in all eligible to be randomised participants and participants with localised disease only.

Figure 7.5: Calibration slopes for the Prostate Cancer mortality metabolomic risk score model and all-cause mortality multi-metabolite risk score models and Prostate Canceer and all-cause death.

Figure 7.6: Receiver operating curves for the Prostate Cancer mortality metabolomic risk score model and all-cause mortality multi-metabolite risk score at discriminating all-cause mortality in all Prostate Testing for Cancer and Treatment controls $(\mathrm{N}=2,150)$ and Prostate Testing for Cancer and Treatment cases and controls ( $\mathrm{N}=4,225$ ).

## List of tables

Table 2.1: Prostate Cancer age-standardised survival at one, five and ten years after diagnosis, 2013-2017, England. ..... 16
Table 2.2: Five-year net survival for Prostate Cancer, by stage, 2013-2017, England ..... 18
Table 2.3: Definition and clinical implications of the Gleason score. ..... 20
Table 2.4: Description of the TNM staging classification for Prostate Cancer ..... 21
Table 2.5: The 3- and 5-tier risk stratification classifications according to the National Institute for Care and Excellence guidelines $(28,34,35)$. ..... 22
Table 2.6: Risk Stratification and treatment options for patients with localised and locally advanced Prostate Cancer, according to the National Institute for Health and Care Excellence (NICE) guidelines. ..... 23
Table 3.1: The eight hallmarks of Cancer according to Hanahan and Weinberg $(152,153)$. ..... 37
Table 3.2: An overview of studies which examined metabolomics in relation to Prostate Cancer progression ..... 43
Table 4.1: Censoring in survival analysis. ..... 71
Table 4.2: Confusion matrix table model for an outcome A ..... 75
Table 4.3: Confusion matrix measures ..... 75
Table 4.4: Strength and consideration for inference of ITT analyses ..... 82
Table 4.5: Potential issues in feasibility RCTs ..... 92
Table 5.1: Baseline characteristics, by progression category in the The Prostate Testing for Cancer and Treatment trial ..... 103
Table 6.1: Baseline characteristics of the men by intervention group. ..... 140
Table 6.2: Effects of the interventions on lycopene serum levels, self-reported fruit and vegetable, and dairy milk intake and step count. ..... 142
Table 6.3: Linear regression results of exposures and intervention arms (Instrumental variable instruments) from the first step of the Instrumental variable analysis. ..... 151
Table 6.4: Causal effect estimates of altered metabolites from the Prostate Cancer Evidence of Exercise and Nutrition trial on Prostate Cancer mortality in the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome consortium, in the inverse variance weighted or Wald ratio methods. ..... 153
Table 7.1: Baseline characteristics for: i) men with a Prostate Cancer specific death vs. survivors in cases; and ii) all-cause deaths vs survivors in controls in the Prostate Testing for Cancer and Treatment cohort. ..... 170
Table 7.2: Model discrimination statistics for the Prostate Cancer mortality metabolomic risk score model and all-cause mortality multi-metabolite risk score for Prostate Cancer and allcause mortality in the ProtecT cohort.180
Table 8.1: Summary of the main findings of this thesis ..... 192
List of appendix tables
Appendix A
Appendix Table A 1: Baseline assessment variable in UKBB. ..... 240
Appendix Table A 2: Current and future enhanced phenotyping variables in UKBB ..... 241
Appendix Table A 3: List of metabolites investigated in UKB ..... 242
Appendix Table A 4: List of studies included in the PRACTICAL consortium ..... 245
Appendix Table A 5: Threats to robust inference when conducting two-sample MR. ..... 247
Appendix B
Appendix B Table B 1: Cox regression results for clinical progression minimally adjusted model, fully adjusted model, censored minimally adjusted model ..... 250
Appendix B Table B 2: Cox regression results for metastases or PCa death- minimally adjusted model, fully adjusted model, censored minimally adjusted model ..... 256
Appendix B Table B 3: Proportional hazard assumption test in the minimally adjusted cox model for clinical progression, metastases or death and any progression ..... 262
Appendix B Table B 4: F-statistic and r-square measures for the instruments of individual metabolites in UKBB ..... 268
Appendix B Table B 5: Mendelian Randomisation results for Prostate Cancer survival in the PRACTICAL consortium with instruments derived from UKBB ..... 274
Appendix B Table B 6:Odds ratios for metastses and PCa death, in the minimally adjusted model, in the ProtecT trial and PCa death in the MR analysis, in the PRACTICAL consortium. ..... 364
Appendix C
Appendix C Table C 1 Linear regression results of baseline metabolic trait levels in the ..... 385
Appendix C Table C 2: Linear regression results of baseline metabolic trait levels in the lycopene arm, compared to control. ..... 393
Appendix C Table C 3: Linear regression results of baseline metabolic trait levels in the brisk walking arm, compared to control ..... 401
Appendix C Table C 4:Linear regression results of metabolic traits in the lycopene arm compared to control (ITT) ..... 408
Appendix C Table C 5:Linear regression results of metabolic traits in the plant-based diet arm compared to control (ITT) ..... 416
Appendix C Table C 6:Linear regression results of metabolic traits in the lycopene arm compared to control, adjusted for baseline metabolic trait levels and smoking status (ITT).424
Appendix C Table C 7:Linear regression results of metabolic traits in the plant-based diet arm compared to control, adjusted for baseline metabolic trait levels (ITT) ..... 432
Appendix C Table C 8:Linear regression results of metabolic traits in the brisk walking arm compared to control (ITT) ..... 440
Appendix C Table C 9:Linear regression results of metabolic traits in the brisk walking arm compared to control, adjusted for metabolic trait levels and age (ITT) ..... 448
Appendix C Table C 10: Linear regression results of the instrumented exposure (dairy milk intake) on metabolic traits in the plant-based diet arm compared to control (IV). ..... 456
Appendix C Table C 11:Linear regression results of the instrumented exposure (dairy milk) on metabolic traits in the plant-based diet arm compared to control, adjusted for baseline metabolic traits (IV) ..... 464
Appendix C Table C 12: Linear regression results of the instrumented exposure (daily fruitand veg portions) on metabolic traits in the plant-based diet arm compared to control (IV)472
Appendix C Table C 13: Linear regression results of the instrumented exposure (daily fruit and veg portions) on metabolic traits in the plant-based diet arm compared to control (IV), adjusted for baseline metabolic traits ..... 480
Appendix C Table C 14: Linear regression results of the instrumented exposure (percentage of days with 12,500 steps or more) on metabolic traits in the brisk walking arm compared to control (IV) ..... 487Appendix C Table C 15: Linear regression results of the instrumented exposure (percentageof days with 12,500 steps or more) on metabolic traits in the brisk walking arm compared tocontrol (IV), adjusted for baseline metabolic traits and age495
Appendix C Table C 16: List of metabolic traits taken forward to MR stage and instrument availability which had a p-value $<0.05$ in the unadjusted ITT ..... 503
Appendix C Table C 17: F-statistic and r-square measures for instruments of individual metabolites (metabolites with p-value <0.05 in the unadjusted ITT) ..... 505
Appendix C Table C 18: Mendelian Randomisation results for Prostate Cancer survival in the PRACTICAL consortium, using altered metabolites in the PrEvENT trial ( $\mathrm{p}<0.05$ ) ..... 507

# Appendix C Table C 19: F-statistic and r-square measures for instruments of individual 6 extra metabolites from the dietary intervention arm (top $10 \%$ metabolites in the unadjusted ITT) <br> 516 

Appendix C Table C 20: Mendelian Randomization results for Prostate Cancer survival in
the PRACTICAL consortium of select metabolites (sensitivity analysis, metabolites taken
forward from PrEvENT trial, top 10\% effect sizes) ..... 517

## Appendix D

Appendix D Table D 1: Cox regression results for individual metabolites and PCa death in
the minimally and fully adjusted model, cases only. ......................................................................... 519
Appendix D Table D 2: Cox regression results for individual metabolites and all-cause death in the minimally and fully adjusted model, case only................................................................ 542

Appendix D Table D 3: Cox regression results for PCa-mRS and All-cause mRS, for all-cause
and PCa specific mortality, in the minimally and fully adjusted regression models, cases
only............................................................................................................................................... 566
Appendix D Table D 4: Correlation between individual metabolites and PCa mRS and all-
cause mRS in the ProtecT trial, cases only................................................................................... 567
Appendix D Table D 5: Cox regression results for individual metabolites and all-cause death in the minimally adjusted model, in controls and pooled568

Appendix D Table D 6:Cox regression results for PCa mRS and All-cause mRS for all-cause mortality, in controls and cases and controls (pooled)602

Appendix D Table D 7: Linear regression results for the associations between individual biomarkers and age in cases, controls and pooled cases and controls with follow-up data in the CAP trial (4,260 participants: 2,167 controls, 2,093 cases)

Appendix D Table D 8: Linear regression results for the associations between individual biomarkers and BMI in cases, controls and pooled cases and controls with follow-up data in the CAP trial (4,260 participants: 2,167 controls, 2,093 cases) 625

## List of appendix figures

## Appendix B

Appendix B Figure B 1: Forest plot of hazard ratios for clinical progression and metastases or death, in the fully adjusted Cox models ..... 378
Appendix B Figure B 2: Forest plot of hazard ratios for clinical progression, in the minimally and fully adjusted Cox models ..... 379
Appendix B Figure B 3: Forest plot of hazard ratios for metastases or PCa death, in the minimally and fully adjusted Cox models ..... 380
Appendix B Figure B 4: Forest plot of hazard ratios for clinical progression and metastases ..... 381or death, in the minimally adjusted Cox models, censored at 10 years of follow-up.
Appendix B Figure B 5: Forest plot of hazard ratios for clinical progression, in the minimally and fully adjusted Cox models, censored at 10 years of follow-up. ..... 383
Appendix B Figure B 6: Forest plot of hazard ratios for metastases or PCa death, in the minimally and fully adjusted Cox models, censored at 10 years of follow-up ..... 384
Appendix D
Appendix D Figure D 1: Hazard ratios (per 1SD) for the associations between individualbiomarkers and all-cause mortality in controls, cases and pooled cases and controls) withfollow-up data in the ProtecT trial ( $\mathrm{N}=4,260$ participants ( 2,167 controls and 2,093 cases)with 436 deaths)648Appendix D Figure D 2: Beta coefficients (per 1SD) for the associations between individualbiomarkers and age in cases, controls and pooled cases and controls with follow-up data inthe CAP trial (4,260 participants: 2,167 controls, 2,093 cases)649
Appendix D Figure D 3: Beta coefficients for the associations between individual biomarkers and BMI in cases, controls and pooled cases and controls with follow-up data in the CAP trial ( $\mathrm{N}=2,820$ participants: 1,393 controls and 1,425 cases) ..... 650

## Chapter 1. Overview of thesis

### 1.1. Thesis rationale

Prostate cancer (PCa) is a leading cause of mortality in men worldwide, with more than 336,000 deaths each year (1). While overall PCa 5 -year survival is $87 \%$ in the UK, stage at diagnosis plays a crucial role, and there is a marked difference in survival between those with the lowest stage ( $96 \%$ ) and those in the highest stage (49\%) (2). Prostate Specific Antigen (PSA) is an enzyme which can be quantified in blood and is used to help detect the presence of PCa and track disease progression. However, most cancers that are detected via PSA testing do not progress and patients may be subject to overdiagnosis and treatment. Multiple trials have looked at the effectiveness of PSA testing in relation to survival, and there is still weak evidence to suggest that PSA screening leads to better PCa survival (3-5). In order to establish the aggressiveness of a tumour, biopsies are performed, and staging is attributed to the cancer. Gleason score is generated following the biopsy by comparing the prostate tumour cells to normal cells and analysing the two most common patterns observed. A score between 2 and 10 is generated, with 10 suggesting the most abundance of abnormal cells and most aggressive disease. Establishing the aggressiveness of disease is done using a combination of measures, such as PSA, Gleason scoring, and more recently using magnetic resonance imaging (MRI)(6). The staging assignment is therefore dependent on the method and quality of performing biopsies, and while recent advances have allowed for this process to be more precise, there are still over $20 \%$ of PCa cases that are missed and $14 \%$ that are under-graded $(7,8)$. Therefore, missed diagnoses of aggressive localised PCa cases are still a significant problem in the clinical management of PCa. More accurate tools that distinguish between localised PCa cancers that progress and those that are indolent are required.

Scientific fields that quantify a large number of molecular measurements, called 'omics, have been able to provide comprehensive snapshots of the underlying biological systems and have been used in instances where standard clinical approaches did not address the need, such as predictive or prognostic models of a particular disease.

Metabolomics, is an 'omics field, and represents the simultaneous study of many metabolites which are small molecule end products of many genetic and proteomic functions (9). Small changes at genetic, transcriptomic and proteomic level can lead to considerable changes at metabolite level which can provide an overall picture of the biological status of health, disease risk and progression (9-11). Metabolomics can also explore disease aetiology through pathway assessment, as well as providing a new approach to identification of predictive biomarkers for prognosis and treatment response (12-15). Recent technological advances have allowed rapid, high throughput, and cheaper quantification of metabolites. Metabolomic lipid profiling has shown that there is an association between lipids and prostate cancer death (16).

Identifying risk and protective factors for progression in localised PCa is challenging. Multiple aspects contribute to this. The factors mentioned above, such as the slow growing nature of localised PCa and the limited indicators of progression contribute to the lack of identified risk and protective factors for localised PCa. The variable incidence and survival around the globe have inspired many studies to look at identifying lifestyle factors that could explain the geographic variation (17). In England, overall PCa survival is $80 \%$ at 10 years, while survival from localised PCa is high, nearly $100 \%$ at 5 years (2). However, some men experience clinical progression of disease and develop castration-resistant prostate cancer which is associated with worse PCa outcomes (18). The Prostate Testing for Cancer and Treatment (ProtecT) trial has found that nearly $20 \%$ of men in the active surveillance group developed clinical progression over the 10-year follow-up (19). Therefore, dietary and lifestyle factors that could help prevent the progression of PCa need more thorough investigation. This could reduce the need for radical treatment and improve outcomes in men with PCa who are undergoing active surveillance.

Evidence of a link between diet and lifestyle in PCa progression has yet to be established. Fruit and vegetable consumption, dairy intake and physical activity levels have been suggested to play a role in PCa progression (20-26). Lycopene, a carotenoid mainly found in tomatoes, and increased fruit and vegetable intake, has been suggested to reduce PCa progression (20-22,27). Increased high-fat dairy milk intake was observed to have a detrimental effect on PCa progression in men diagnosed with localised disease $(22,28)$. However, recent systematic reviews suggest the evidence is inconsistent and many of the interventional studies that follow an RCT design are poorly conducted and have a high risk of bias (20,27,29-31). Physical activity, particularly moderate and vigorous exercise and
brisk walking, has also been suggested to delay the progression of localised PCa disease $(22,32)$. However, the evidence was mainly observational and the results were not replicated when an RCT design was used (20).

Causal metabolites, which are causally linked to PCa progression, can further the understanding of disease aetiology and mechanisms of progression, and provide therapeutic targets to reduce PCa progression. Metabolites which are non-causally linked to PCa progression, can serve as markers of disease progression and can be used in developing clinical prognostic biomarkers. In this thesis, I aimed to better understand the causal and predictive role (non-causal) that circulating metabolites may play in PCa progression and assess whether randomised dietary and physical activity interventions act on these metabolites to impact on the progression of PCa to metastases and death.

### 1.2. Aims and Objectives

The aims of my thesis are two-fold: i) to better understand molecular mechanisms underlying PCa progression, so as to identify new potential therapeutic or behavioural targets for interventions to reduce PCa progression; and ii) investigate the performance of metabolomic risk scores which could be used to distinguish indolent from lethal localised PCa in a clinical setting. These aims will be achieved by the following objectives:
(1) Identify metabolic traits which are associated with localised PCa progression (clinical progression, PCa metastases or PCa death) in men with screen detected localised PCa
(2) Estimate the causal effect of individual metabolic traits on PCa survival using Mendelian Randomisation (MR)
(3) Assess the effects of a randomised intervention (lycopene supplementation, reduced dairy milk consumption or brisk walking) on metabolic traits of men with localised PCa, and then estimate their causal effect on PCa survival using MR
(4) Investigate the performance of two metabolomic risk scores previously developed to predict progression in localised PCa cases in clinical settings

### 1.3. Thesis outline

This PhD thesis starts with a description of the clinical aspects of PCa, its progression, epidemiology, as well as a review of known lifestyle risk and protective factors in Chapter 2. Chapter 3 introduces general concepts in the field of metabolomics and the link of metabolomics to PCa, lifestyle and physical activity. The methodological Chapter 4 introduces the data sources and methodological approaches used in answering the research questions of the thesis. This chapter includes statistical methods used in analysing metabolomic data, and methodologies of observational (survival and prediction analyses) and causal (intention to treat, instrumental variable and MR analyses) epidemiology. The visual representation of research questions for the results chapters ( 5,6 and 7 ) are presented in Figure 1.1. Chapter 5 presents the results of an observational and a causal (MR) study which assessed the link between individual metabolites in relation to PCa progression. The observational analysis used metabolomic data from the ProtecT trial, an embedded trial in the Cluster randomised triAl of PSA testing for Prostate cancer (CAP) trial while the MR analysis used genetic data from the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL) and UK Biobank (UKBB). Chapter 6, discusses the findings from the Prostate Cancer Evidence of Exercise and Nutrition trial (PrEvENT) feasibility trial, assessing if interventions of lycopene supplementation, increased fruit and vegetable and reduced dairy milk intake and physical activity altered the metabolome of men with localised PCa. The second part of Chapter 6 investigates whether the metabolites altered by the intervention are causally linked to PCa mortality, by using results from the MR analysis of metabolites in relation to PCa mortality from Chapter 5. Chapter 7 uses two already developed metabolomic risk scores in the general population and assesses their predictive performance in a clinical setting within the ProtecT trial. In Chapter 8, I bring together the results of the three results chapters, and discuss how they fit within the overall PCa progression picture, and the implications of the findings in the clinical and public health context. In addition, I present how metabolomics via methodologies employed in this thesis allowed the investigation of new potential mechanisms of progression in PCa. Chapter 8 aims to also to identify intermediate
biomarkers potentially causally linked to PCa progression, which could be altered by lifestyle interventions. Lastly, I present the strength and weaknesses of my studies and methodologies used and suggest future research that may help tackle the limitations.

Figure 1.1: Diagram of Thesis research questions


Are circulating metabolites associated with Prostate Cancer specific and all-cause mortality in men with screen detected disease?
2) Do Prostate Cancer and all-cause multi metabolite rsk scores predict Prostate Cancer or all-cause mortality in men with screen detected disease?

Are circulating metabolite levels a baseline associated of clinical
progression in progression in men Prostate Cancer?

Do circulating metabolite levels at metabolite levels at
baseline have a baseline have
causal role in causal role in
subsequent Prostate Cancer specific mortality?

1) Is the metabolome of men with PCa altered following lifestyle interventions of : i) lycopene
supplementation; or ii) dietary advice to increase fruit and vegetable consumption and reduce dairy milk consumption; or iii) a brisk walking physical activity intervention? 2) What is the causal effect of the metabolites altered by these lifestyle interventions on PCa mortality?

Figure 1.2 Schematic Thesis summary


## Suggestive observational evidence that

 lipoproteins, cholesterol, glycolysis, fluid balance and inflammation markers were associated with PCa progression. Causal evidence that total-, free- and esterifiedcholesterol, some measures of intermediate-, low- and very low-density lipoprotein cholesterol, sphingomyelins, apolipoprotein B, omega-3 fatty acids, docosahexaenoic acid and valine increased and high-density and very low-density lipoprotein related measures, histidine and the ratio of omega-6 to omega-3 fatty acids decreased PCa mortality.

> Weak evidence that two metabolomic risk scores developed in the general population did predict PCa or all-

ProtecT = Prostate Testing for Cancer and Treatment
PrEvENT = Prostate Cancer Evidence of Exercise and Nutrition trial PRACTICAL $=$ Prostate Cancer Association Group to Investigate

Cancer Associated Alterations in the Genome
PCa $=$ Prostate Cancer
VLDL=very low-density lipoprotein

### 1.4. Datasets used

The datasets used in this thesis are described in detail in methodology Chapter 4. Four main datasets were used within this thesis.

The ProtecT trial is a large randomised controlled trial (RCT) in which 82,429 men aged 50-69 years were PSA tested and those men who were subsequently diagnosed with localised PCa ( $\mathrm{n}=2,417$ ) were invited to be randomised to three possible treatments: radical prostatectomy, radical radiotherapy or active surveillance (33). Data on circulating metabolites measured by NMR were available for men at baseline (their first visit). Men were followed-up for a median of 10 years for PCa progression outcomes: clinical progression (evidence of metastases, diagnosis of clinical T3 or T4 disease, longterm androgen-deprivation therapy, ureteric obstruction, rectal fistula, the need for a urinary catheter owing to local tumour growth or PCa death), metastases and PCa specific death.

The second dataset used was the PrEvENT trial (34). This is a $3 \times 2$ factorial RCT, in which 81 men were randomised twice to both a dietary (lycopene supplementation, increased fruit and vegetable and reduced dairy milk intake or control) and a physical activity intervention (brisk walking or control). The active interventions were applied over a period of 6 months. Blood samples and patient reported outcomes were collected at baseline and follow-up using questionnaires that captured information on diet, physical activity, mental health and urinary symptoms.

The third and fourth datasets consist of genetic wide association studies (GWAS) data which were used in the MR analyses. Genetic data from a GWAS of metabolites in UKB was used, which contained genetic data from 115,078 participants (35). Genetic data from the PCa mortality GWAS was obtained from the PRACTICAL consortium which holds data from over 133 study groups and over 120,000 PCa cases (23).

### 1.5. Statistical Methods

The analyses employed in this thesis focus on metabolomic data. Chapter 4 describes in detail all the statistical methods used throughout this Thesis to answer the research questions. Briefly, Cox proportional hazard models were used to assess the association of the individual metabolic traits in relation to PCa progression in the ProtecT trial in Chapter 5. Mendelian Randomisation (MR) was employed to estimate a causal effect of metabolites on PCa survival, using the PRACTICAL mortality data and UKB genetic data. MR employs a form of instrumental variable (IV) analysis using genetic variants, which are randomly assigned at birth and thus less likely to be subject to be confounded by environmental factors, to proxy exposures of interest, in this case metabolite levels.

Chapter 6 investigated the effects of multiple lifestyle and dietary interventions on metabolite levels in the PrEvENT feasibility RCT. The RCT was analysed using both intention to treat (ITT) and IV analyses to estimate the causal effect and complier-adjusted causal effect of multiple lifestyle and dietary interventions on metabolic levels. I then employed a novel extension of MR which aimed to estimate the causal effect of the interventions on PCa survival, acting via metabolic trait levels. Chapter 7 evaluated the performance of two previously developed metabolomic prediction risk scores for PCa and all-cause mortality, at predicting PCa death in the ProtecT study using confusion matrix measures (negative and positive predictive values, accuracy, Cohen's Kappa), area under the receiver operating curve (AUROC).

# Chapter 2. Prostate Cancer epidemiology and lifestyle factors 

### 2.1. Introduction

Prostate Cancer (PCa) affects men worldwide, particularly in the developed world, with over 1.2 million new diagnoses and 359,000 deaths recorded globally in 2018 (1,36). Large variations in PCa incidence and survival across the world has led to many studies investigating modifiable factors that could explain these observed patterns $(17,18)$.

### 2.2. Prostate Cancer epidemiology

### 2.2.1. PCa prevalence

PCa is one of the top five cancers for incidence and mortality, with 1.2 million newly diagnosed cases globally every year (37). The prevalence of PCa, i.e. number of men living with the disease, is generally high globally, with western and developed income countries having the highest rates (Figure 2.1) (38). 1 in 8 Caucasian men and 1 in 4 Black men develop PCa during their lifetime (39). In the UK, there are around 475,000 men living with PCa and PCa has an estimated 5-year prevalence rate of 705/100,000 population (40). Approximately $80 \%$ of those living with PCa in the UK are aged 65 and over, with similar patterns observed worldwide (41).

Figure 2.1: Heatmap of the estimated proportion of prevalent Prostate Cancer cases globally.


Credit: Global Cancer Statistics and International Agency for Research on Cancer (41)

### 2.2.2. Prostate Cancer incidence rates

Cancer incidence is a standard measure of frequency of primary cancer diagnoses made within a specified time period and population. Generally, incidence is expressed as a rate per 100,000 or per million in the case of a rare disease. This crude measure does not take into account the age distribution of the populations and therefore incorrect conclusions can be drawn when comparisons are made between multiple populations or regions that have different age distributions. For example, if country A has an older population and a higher incidence of a particular disease which is more common in the elderly, compared to country B, then it would be misleading to state that country A has more incident cases. This is because the likelihood of developing the disease is higher in older people and country A has more elderly people, thus the incident rates cannot be reasonably compared. To address this issue, incident rates are commonly age standardised, or adjusted to take into account the age distribution within each population, allowing inter-population comparisons to be made $(42,43)$.

In the UK, PCa is the most common cancer with approximately 48,500 new diagnoses made every year, accounting for $26 \%$ of all new cancer cases in males (2). Similar to the prevalence figure, $76 \%$ of all new cases are diagnosed in men over 65 years of age (2). PCa age-adjusted
incidence has increased over the last 40 years worldwide, mainly due to the rise of PSA testing. However there is substantial global variation (Figure 2.2) (1). Many factors have been attributed as contributing to the variation, such as the extent of PSA screening programmes, genetics, environmental and lifestyle factors (1). For example, the odds of someone being diagnosed with PCa before 79 years of age is 1 in 47 in low-middle income countries, and 1 in 6 in high-income countries $(1,44)$. Asian countries have lower incidence of PCa, and it is unclear whether this is a result of genetics, lifestyle or underdiagnosis. However, studies that looked at the pattern change in incidence for people migrating from a lower incident country (e.g., Asian countries) to a higher incidence country (e.g., United States) noticed that the incidence within this group was higher than that of their native country, but lower than that of the host country $(1,45,46)$. Whilst PSA screen programmes may have been the main contributor to the observed increased incidence rates in western countries, by detecting indolent cases which would otherwise not be detected (overdiagnosis of 23-42\%), there is evidence that even before the introduction of PSA testing there was still variation in rates globally $(1,17,47,48)$. This would suggest while there may be a genetic aspect and healthcare or screening differences, lifestyle factors such as western diets, sedentary behaviour and obesity may also contribute to the observed patterns $(17,49,50)$.

Figure 2.2: Prostate Cancer age-standardized incidence rates globally, 2020.


The rates are computed using estimated number of cases and are per 100,000 population.

Credit: Global Cancer Statistics and International Agency for Research on Cancer (41)

### 2.2.3. Prostate Cancer mortality rates and survival

PCa specific mortality rates vary less than incidence rates, and the highest mortality estimates are observed in low to middle income countries of the Caribbean, southern and western Africa and South America. The lowest PCa mortality rates are observed across Asia, north Africa and North America $(17,38)$. Although PCa is the most common cancer in males in many countries, mortality from PCa is mainly concentrated in areas of the Caribbean, South America and Sub-Saharan Africa (Figure 2.3) (17). Age standardized mortality rates worldwide vary from 41.7 per 100,000 population in Zimbabwe to 0.54 per 100,000 population in Bhutan (Figure2.4). In the UK, the estimated mortality rate is 12.4 per 100,000, three times less than that of Zimbabwe and nearly 20 times higher than the lowest global mortality rate of Bhutan. Some countries have likely been over diagnosing PCa for over 20 years, particularly the type of indolent disease that would very rarely lead to death $(17,47,51)$. This may have introduced lead time bias in survival statistics, which occurs when detection happens before symptoms arise and does not affect the time of death (17). In the US, men of African ancestry are at higher risk of death from PCa, with African American men being 2.5 times more likely to die from PCa than Caucasians in the US (52). However, the reasons for this are believed to be influenced more by inequality rather than genes, since
compared to white men, African American men have less access to healthcare, present with later stages of disease at diagnosis are more likely to be undertreated due to institutionalised racism $(52,53)$. A literature review found no differences in PCa specific mortality between races in equal access healthcare systems (Veterans Affairs in the United States, National Health Service (NHS) in the UK, Health Canada) while a systematic review found consistent higher risk of PCa mortality in black men after taking into account the impact of PSA screening and access to free healthcare $(54,55)$. A large study conducted in routinely collected cancer and mortality data in the UK, found that black men are twice as likely to die from PCa compared to white men during their lifetime (39). A recent study of metastatic castration-resistant PCa which used real world data from the largest healthcare trust in England found that black men may have better PCa specific mortality compared to white men (56). The limited ethnic data and multiple sources of social factors that affect the link between PCa mortality and race make the current evidence hard to disentangle.

Figure 2.3: Map of Prostate Cancer age-standardized mortality rates globally, 2020.


The rates are computed using estimated number of deaths and are per 100,000 population.

Credit: Global Cancer Statistics and International Agency for Research on Cancer (41)

Figure 2.4: Prostate Cancer age-standardized mortality rates globally, 2020.


Credit: Global Cancer Statistics and International Agency for Research on Cancer (41)
Cancer survival is the percentage of people who are alive after a given period of time following primary diagnosis, generally 1,5 or 10 years. PCa survival has been increasing over time, but this could (at least partially) be an artefact of lead time bias, due to an increased number of diagnoses of indolent disease been made through PSA testing $(47,57)$. Thus, when assessing PCa survival patterns over time, regional screening and detection practices should be taken into account.

According to estimates from CRUK, between 2013-2017 in England overall PCa survival was $97 \%$ at one year, $87 \%$ at five years and $78 \%$ at 10 years (Table 2.1). Five-year survival was lowest in the 80-99 years age-group (66\%) and highest in the 60-69 age group (94\%), most likely due to higher rates of PSA testing within this age group (Figure 2.6). PCa mortality is strongly associated with age, and in the UK 75\% of all PCa deaths occurred in those aged 75 and over. Age specific mortality rates increase markedly from 55-59 and steeply from approximately 70-74, with the highest mortality rates in those aged over 90 (Figure 2.6) (2).

Although over time the age standardised mortality rate has decreased by $10 \%$, the rates vary by age groups, with those between the ages of 25-49, 50-59 and 60-69 remaining stable, those aged $70-79$ decreasing by $13 \%$, and the over 80 s increasing by $45 \%$ (Figure 2.7). In England, net survival at one year varies from $95.9 \%$ in South Yorkshire and Bassetlaw to $98.7 \%$ in

Suffolk and North East Essex (Figure 2.8). Five-year PCa survival is highest in the localised disease groups ( $100 \%$ for stages 1 and 2 ) and lowest in the metastatic stage 4 disease ( $49 \%$ ) (Table 2.2).

Table 2.1: Prostate Cancer age-standardised survival at one, five and ten years after diagnosis, 2013-2017, England.

| Years after diagnosis | Number of cases | Net survival | Lower Confidence | Upper <br> Confidence |
| :---: | :---: | :---: | :---: | :---: |
|  |  |  | Interval | Interval |
| one year | 204,175.0 | 96.6 | 96.5 | 96.7 |
| five years | 204,175.0 | 86.6 | 86.2 | 87.0 |
| ten years | 365,274.0 | 77.6 | 76.5 | 78.7 |

Credit: Cancer Research UK (2)

Figure 2.5: Bar chart of Prostate Cancer net survival by age, 2009-2013, England


Credit: Cancer Research UK (2)

Figure 2.6: Average number of deaths per year, by age group, per 100,000 population, UK, 2016-2018.


Credit: Cancer Research UK (2)

Figure 2.7: European age-standardised mortality rates per 100,000 population, by age group, in the UK, 1971-2018.


[^1]Figure 2.8: One-year net survival for Prostate Cancer in England at local authority level, 2019.


Source: National Cancer Registration and Analyses Service (58)

Table 2.2: Five-year net survival for Prostate Cancer, by stage, 2013-2017, England.

| Stage | Number of <br> cases | Net <br> survival | Lower <br> ConfidenceConfidence <br> Interval | Upper <br> Interval |  |
| ---: | ---: | ---: | ---: | ---: | ---: |
| 1 |  |  |  |  |  |
| 2 | $3,158.0$ | 100.1 | 99.5 | 100.8 |  |
| 3 | $40,086.0$ | 95.6 | 94.7 | 96.6 |  |
| 4 | $39,141.0$ | 49.0 | 47.9 | 50.2 |  |
| Unstageable | 512.0 |  |  |  | 101.1 |
| Unknown / | $24,336.0$ | 82.1 | 81.2 | 83.0 |  |
| missing |  |  |  |  |  |

Credit: Cancer Research UK (2)

### 2.3. Clinical aspects of Prostate Cancer

Despite being the second commonest cause of death due to cancer in men, in the UK, PCa that is localised to the gland tends to progress slowly, with most tumours not spreading from the prostate. PCa can be indolent and before symptoms appear death could occur from other causes. Men with PCa do not usually exhibit symptoms of PCa until the tumour has
grown large enough to put pressure on the urethra, as this can lead to urinary symptoms such as frequent urination, hesitancy and straining while urination, haematuria (blood in urine) or haematospermia (blood in semen). In addition, PCa that has spread beyond the gland will include symptoms such as back, neck and testicle pain, loss of appetite and unexplained weight loss. The following sections have been written with reference to the latest update of the 2019 NICE guidelines on PCa diagnosis and management (59).

### 2.4. Diagnosis of Prostate Cancer in the UK

The first steps in assessing and diagnosing PCa are the clinical history assessment, digital rectal examination and a PSA test which are conducted in general practice. These investigations allow the clinician to assign a suspected localised PCa diagnostic. The firstline investigation for suspected localised PCa is a multiparametric MRI scan of the prostate. The scan allows for a prostate Imagining-Reporting and Data Systems or a Likert score to be generated, which helps identify more high-risk PCa and inform what further investigations should be performed. The Likert score is generated by combing the findings from multiple images generated from different perspectives. A Likert score of 1 or 2 suggests there is low risk of clinically significant PCa presence (i.e., the cancer may affect the man during his lifetime) and the clinician may decide against a biopsy but recommend regular PSA tests. A Likert score of 3 or more suggests the man is at increased risk of having clinically significant PCa.

For men with a Likert score of 3 or more, a multiparametric MRI prostate biopsy should be offered. Of those diagnosed with low-risk MRI, 11-28\% will have clinically significant cancer which represents a problem of underdetection. However, prostate biopsies detect less than half the missed cases by MRI scans, mainly due to the incorrect acquisition of tissue following the MRI image, a common problem when performing MRI targeted transrectal ultrasound-guided biopsies $(60,61)$. The multiparametric MRI-influenced transrectal ultrasound-guided biopsy is the most common type of biopsy used in the NHS for PCa diagnosis. It uses the information generated from the multiparametric MRI images to determine the biopsy needle placement, which is guided via ultrasound. Following the biopsy, overdetection occurs in 18-23\% of those with a low-risk MRI, who are diagnosed with clinically insignificant disease, which is unlikely to be life-threating, but could lead to further tests and treatment (61). The pathologist will examine the biopsy and assign two
histological scores from 1 to 5 , which are used in calculating the Gleason score. The Gleason score is made up of the most common (first number) and the second most common (second number) pattern in the carcinoma cells, by assessing the glandular differentiation and the pattern of growth of the tumour (62). The Gleason score is calculated by adding the two histological scores. The grade group is a system based on Gleason score which was developed to help clinicians grade PCa in relation to the severity of cancer (Table 2.3) (63).

Table 2.3: Definition and clinical implications of the Gleason score.

| Gleason score | Grade <br> Group | What it means |
| :--- | :--- | :--- |
| Gleason score 6 (or $3+3=6$ ) | Grade <br> Group 1 | The cells look similar to normal prostate cells. The <br> cancer is likely to grow very slowly, if at all |
| Gleason score 7 (or 3+4=7) | Grade <br> Group 2 | Most cells still look similar to normal prostate cells. The <br> cancer is likely to grow slowly |
| Gleason score 7 (or 4+3=7) | Grade <br> Group 3 | The cells look less like normal prostate cells. The <br> cancer is likely to grow at a moderate rate |
| Gleason score 8 (or 4+4=8) | Grade <br> Group 4 | Some cells look abnormal. The cancer might grow <br> quickly or at a moderate rate |
|  | Grade | The cells look very abnormal. The cancer is likely to <br> grow quickly |
| Gleason score 9 or 10 (or $4+5=9$, <br> $5+4=9$ or $5+5=10$ ) | Group 5 |  |

Credit: Cancer Research UK(2)

TNM staging is the most common system of assessing cancer severity available in clinical practice (64). The scoring is based on tumour growth (T), lymph node spread (N) and metastasis presence (M). Table 2.4 provides the breakdown of stages within each category for PCa. The most common sites for metastases to develop in PCa are in the bones and lymph nodes. Localised PCa is defined as PCa of T1 or T2 stage, while locally advanced disease is generally defined as having T3 or T4 stage.

Table 2.4: Description of the TNM staging classification for Prostate Cancer

| Stage | Subtype | Description of stage |
| :--- | :--- | :--- |
| T | Tumour cannot be evaluated |  |

Credit: Cancer Research UK (2)
Until 2019, NICE guidelines recommended a three-tier risk stratification system for risk stratification of PCa severity. In 2019, the 3-tier system was replaced by the 5-tier Cambridge prognostic group classification (65) (Table 2.5).

Table 2.5: The 3- and 5-tier risk stratification classifications according to the National Institute for Care and Excellence guidelines $(28,34,35)$.

| NICE risk group | Criteria | CPG category | Criteria |
| :---: | :---: | :---: | :---: |
| Low-risk disease | $\begin{aligned} & \text { Gleason score } \leq 6 \\ & \text { AND PSA }<10 \mathrm{ng} / \mathrm{ml} \\ & \text { AND stages } 11-\mathrm{T} 2 \mathrm{a} \end{aligned}$ | 1 | Gleason score 6 (Grade Group 1) AND PSA $<10 \mathrm{ng} / \mathrm{ml}$ <br> AND stages T1-T2 |
| Intermediate- <br> risk disease | Gleason score 7 <br> OR <br> PSA $10-20 \mathrm{ng} / \mathrm{ml}$ OR <br> Stage T2b | 2 | Gleason score 3+4=7(Grade Group 2) OR <br> PSA $10-20 \mathrm{ng} / \mathrm{ml}$ <br> AND stages T1-T2 |
|  |  | 3 | Gleason score 3+4=7 (Grade Group 2) <br> AND PSA $10-20 \mathrm{ng} / \mathrm{ml}$ <br> AND stages T1-T2 <br> OR <br> Gleason score $4+3=7$ (Grade Group 3) <br> AND stages T1-T2 |
| High-risk or locally advanced disease | Gleason score 8-10 <br> OR $\mathrm{PSA}>20 \mathrm{ng} / \mathrm{ml}$ <br> OR <br> Stage $\geq$ T2c | 4 | One of: <br> Gleason score 8 (Grade Group 4) <br> OR $\text { PSA }>20 \mathrm{ng} / \mathrm{ml}$ <br> OR <br> Stage T3 |
|  |  | 5 | Any combination of: <br> Gleason score 8 (Grade Group 4), PSA > $20 \mathrm{ng} / \mathrm{ml}$ or Stage T3 <br> OR <br> Gleason score 9-10 (Grade Group 5) <br> OR <br> Stage T4 |

PSA $=$ Prostate specific antigen

Reproduced from: National Prostate Cancer Audit Report 2021. Copyright © 2022 Healthcare Quality Improvement Partnership. Reproduced with permission (66)

### 2.4.1. Treatment of Prostate Cancer

There are three main treatment options for localised disease, active surveillance, radical prostatectomy and radiotherapy. Active surveillance is an option for patients with low- and intermediate-risk PCa where the patient is monitored long term with the option to have radical treatment or hormone therapy on signs of disease progression. Radical prostatectomy removes the whole of the prostate gland, while radical radiotherapy uses external beam radiotherapy, either hypofractionated or conventional, to destroy the cancerous area of the prostate. Men with intermediate- and high-risk PCa are offered an
androgen deprivation therapy before, during and after external beam therapy. Treatment choices for PCa are made based on the Cambridge risk stratification model (Table 2.6).

Table 2.6: Risk Stratification and treatment options for patients with localised and locally advanced Prostate Cancer, according to the National Institute for Health and Care Excellence (NICE) guidelines.

| Risk Stratification <br> (Cambridge prognostic <br> group classification) |  | Treatment recommendations |
| :--- | :--- | :--- |
| 1-low risk | -Active surveillance <br> Radical prostatectomy and radiotherapy (if active <br> surveillance is not suitable/acceptable to person) |  |
| 2- favourable intermediate risk | -Active surveillance <br> Radical prostatectomy or radiotherapy (if suitable) |  |
| 3-unfavourable intermediate risk | -Radical prostatectomy or radiotherapy <br> Active surveillance (if person chooses not to have radical <br> treatment immediately) |  |
| 4-high risk | $-\quad$ Radical prostatectomy or radiotherapy |  |
| 5- very high risk |  |  |

### 2.4.2. Side effects of prostate biopsy and of treatment of localised prostate cancer

Transrectal biopsy is the most common type of biopsy used in the diagnosis of PCa. They are reasonably well tolerated. However nearly $7 \%$ of men rate the pain following the biopsy as a major moderate problem and 7 days after the biopsy, nearly $20 \%$ would consider a follow-up biopsy to be a major or moderate problem (67). In rare cases (1 in 100) it can lead to sepsis which is a serious condition that can require hospitalization (61).

All treatments for PCa have side effects. Men who undergo radical prostatectomy for localised PCa are more likely to have poorer urinary and sexual function compared to men receiving active surveillance or radiotherapy (19). The ProtecT trial, which randomised men to active surveillance, radical radiotherapy, and radical prostatectomy found that men in the prostatectomy group had the worst outcomes in terms of sexual side effects and urinary continence compared to the other groups, and that even though some recovery was made,
the negative effects remained throughout the 6 years' follow-up of the trial (68). Absorbent pads, for urinary incontinence, were used by $46 \%$ of men at 6 months after surgery in the prostatectomy group compared to $4 \%$ and $5 \%$ after starting active-monitoring and radical radiotherapy, respectively. Furthermore, the prostatectomy group maintained the highest rate of use of pads 6 years on (68). Erectile function decreased in all men regardless of treatment over the 6 years, with the active surveillance group experiencing the least negative effects on erectile function and the prostatectomy group the greatest (68). Men who underwent radiotherapy seemed to have worse bowel function than men on other treatment groups; however, the bowel function side effects were generally short term only (68).

### 2.5. Risk and protective factors for Prostate Cancer progression

Due to the slow rate of progression of PCa, particularly in localised disease (stages T 1 and T2), research into identifying potential risk factors for disease progression is challenging. Proxy measures which do not have a direct link to progression, such as rate of change of PSA, are generally used in progression studies. This makes findings difficult to interpret since the findings observed in the proxy measures may not extrapolate well to observed disease progression outcomes. This means that while studies may find risk factors for proxy measures, these may not be risk factors of actual PCa progression. Research into diet and physical activity is often complicated by the complex interventions required to assess their impact on disease in the absence of good proxies for clinical outcomes (22).

In addition to the above aspects, methodology variability, poor study design and inconsistency of results of these studies contribute to the lack of evidence around risk factors for disease progression for PCa (20). General recommendations from the World Cancer Research Fund (WCRF) suggest eating a diet rich in wholegrains, vegetables, fruit and beans, not smoking and being at least moderately physically active should be undertaken to prevent the development and progression of many cancers, including PCa.

The next section presents the existing evidence in relation to PCa progression for the most investigated dietary and physical activity factors. Systematic reviews on the impact of dietary, nutritional and physical activity interventions on PCa progression have come to no reliable conclusions on the impact of such interventions on disease progression
$(20,27,29,69,70)$. Some observational epidemiological studies have reported a range of potential risk and protective factors such oily fish intake, dairy products, fruits and vegetables, tomato and lycopene products, selenium, isoflavones, vigorous physical activity on PCa development and progression $(22,25,26,28,32,71)$. However, these findings were not replicated when interventions using an RCT design, the gold standard in assessing causality, were conducted $(22,25,28,32,72)$. Observational studies are prone to confounding, particularly when the disease is slow growing and exposure to environmental and lifestyle factors is accumulated over many years and poorly documented. On the other hand, existing RCTs may have been underpowered and at high or unclear risk of bias, may not have intervened long enough, or used outcomes of dubious link to clinical outcomes of progression (20,72,73).

### 2.5.1. Dietary Factors

### 2.5.1.1. Lycopene

Lycopene is a natural compound found in plants, particularly in tomatoes, which acts as an antioxidant, and has been suggested to reduce inflammation (74). Because inflammation has been suggested to be a precursor of PCa , compounds targeted at reducing inflammation have been investigated as a tool to delay PCa progression. Lycopene has been linked to reduced PCa risk, although the evidence is inconsistent, with high quality observational studies with longer follow-up showing the greatest protective effect (75-77). A recent metaanalysis of RCTs found that lycopene did not decrease PSA levels in men with nonmetastatic cancer, except in men with higher baseline PSA (70). However, it is unclear what the implications of these findings are on the long-term effects of lycopene on PCa progression since the outcome investigated was changes in PSA levels rather than PCa progression.

Multiple mechanisms of action for lycopene's impact on PCa risk or progression have been proposed. One plausible mechanism links lycopene supplementation to the upregulation of the connexin 43 expression and gap junctional intercellular communication, which have been showed to decrease with disease severity and have been suggested to be useful intermediate biomarkers in PCa chemoprevention clinical trials (78-80). Another mechanism through which lycopene could affect PCa progression is the peroxisome proliferatoractivated receptor gamma, a nuclear receptor responsible for cell growth, differentiation,
and homeostasis which has been shown to exhibit anti-tumour apoptotic effects in human PCa cells $(81,82)$. Lycopene has also been postulated to affect insulin-like growth factor-1 (IGF-1) inhibition which has been linked with increased risk of development of PCa $(83,84)$. However, none of these mechanisms have been established and there is still much uncertainty around the bioavailability of lycopene and its effects on PCa risk and progression.

### 2.5.1.2. Dairy/ calcium

Dairy intake, particularly in the form of milk, has been suggested to increase risk of developing and progression of PCa (85-93) . Meta-analyses linking dairy intake and increased PCa risk have given inconsistent results, with different values of risk estimations and high uncertainty, possibly due to the inaccurate and inconsistent ways of measuring dairy intake $(27,94)$. This led to the WCRF to conclude that the evidence is limited and suggestive that dairy products and high calcium intake may increase the risk of $\mathrm{PCa}(24)$. There is limited epidemiological evidence linking PCa progression and dairy intake, with most studies using proxies for survival such as non-localised and high-grade disease status. The few studies which used PCa mortality data were restricted to three cohorts. Two were US health professional based cohorts which were similar in cohort characteristics and had cases diagnosed via PSA screening and the third was a Swedish cohort in which cases were not diagnosed via screening but following attendance for prostate-related symptoms $(24,28,85,95)$. Dairy intake has been shown to be linked to increased risk of death from PCa, however it is still unclear which subsets of cases and what type and quantity of dairy most influence this association. For instance, Yang et al. found that in men with non-metastatic PCa cancer, dairy intake was associated with a higher risk of dying from PCa and the association was stronger for high-fat dairy (95). Other studies found that high-fat, but not total dairy intake was associated with increased risk of PCa mortality $(28,71,85)$. The evidence between calcium intake and PCa progression is not well established. Studies that investigated the association between calcium intake and PCa progression found a positive relationship between calcium intake for high grade disease or advanced PCa status, but no link to PCa mortality $(85,96)$.

Mechanisms of actions linking dairy and calcium intake to PCa progression have been proposed, however none fully explain the epidemiological evidence. One hypothesis is that dairy calcium inhibits the concentration of plasma calciferol, the active form of vitamin D ,
which has been associated with pro-differentiating, antiproliferative and anti-metastatic effects in PCa cells. Calciferol has been hypothesised to slow down metastatic disease through multiple mechanisms, such as by down regulating the parathyroid hormone related protein expression, leading to reduced bone resorption and deceleration of PCa metastases $(95,97)$. Another hypothesis that has been proposed to explain the link between dairy intake and PCa progression is through IGF-1, a hormone molecularly similar to insulin, which is an anti-apoptotic and mitogenic hormone involved in prostate carcinogenesis and which has been shown to increase due to dairy consumption $(83,85,98,99)$. Despite the many proposed mechanisms, none has been consistently linked to PCa progression.

### 2.5.1.3. Fruit and vegetables

The WCRF, along with clinical guidelines, suggest that eating a diet rich in fruit and vegetable could be beneficial in reducing PCa risk and progression $(24,100,101)$. A study carried out in the Health Professionals' Follow-up Study looked at the pattern of consumption of vegetables after a diagnosis of PCa in relation to PCa mortality and found that intake of cruciferous vegetables (e.g broccoli, cauliflower, Brussels sprouts etc.) was associated with decreased PCa progression (25). However, the study was observational, and had no pre-diagnostic diet exposure data and a limited number of events (25). A recent review of diet and lifestyle for patients with PCa has suggested that multimodality studies, which combine various factors, may be a better approach to investigating the effects of dietary and lifestyle factors on PCa outcomes (102). An example of this is the Men's Eating and Living study, a phase III RCT, which randomised patients with early stage PCa to a telephone counselling intervention that aimed to encourage participants to eat at least 7 servings of fruit or vegetables per day over a period of 24 months $(103,104)$. This behavioural intervention which increased carotenoid, cruciferous and leafy green vegetables intake over a period of two years did not find evidence of a decrease in the risk of PCa progression in the intervention arm when compared to the control group. The study emphasizes the lack of reliable evidence generated via observational epidemiology of risk and protective dietary factors in PCa progression.

One mechanism that has been proposed that could link fruit and vegetable intake and better PCa progression is the molecular mechanism for the proposed anti-cancerous agent apigenin (105). Apigenin which is a flavone and is commonly found in many
fruits and vegetables, has been suggested to have anti-cancerous properties, through multiple mechanisms of actions such as cell growth arrest, apoptosis of tumorous cells and signalling pathways (105). The evidence linking apigenin and PCa progression is based on studies of cell lines and animal models. Apigenin has been shown to decrease levels of antiapoptotic protein, which led to increased rate of apoptosis, and inhibited expression of multiple genes which are involved in PCa progression and are regulated by histone decetylase enzymes $(105,106)$. Histone deacetylases are being proposed as molecular targets for anti-cancerous drugs, however inhibitors of histone deacetylases that have been investigated in clinical trials have shown considerable side-effects, such as atrial fibrillation (106). Apigenin has also been found to supress PCa progression by altering the IGF-1 signalling(107). Apigenin, a plant derived histone decetylase and widely present in fruit and vegetables, is showing to be a promising agent in PCa given its potential antiproliferative and apoptotic properties (106).

### 2.5.1.4. Other dietary factors

Pomegranate is a fruit which contains a large amount of flavonoids and polyphenols, both of which have been suggested to have anti-cancerous properties (108-110). Pomegranate has been shown to have antiproliferative effects in PCa cells and to inhibit factor kB which is linked to PCa progression and biochemical relapse post prostatectomy (108,111-113). Interventions which randomised men to pomegranate supplementation have found better outcomes in PSA levels, PSA doubling time and reduced indicators of oxidative stress and androgen signalling when compared to the control groups $(108,114,115)$. A trial in men with PCa undergoing active surveillance found that a supplement containing pomegranate, green tea, broccoli and turmeric slowed down the PSA increase compared to the control group, however no other indicators of progression were assessed. Other dietary factors such as Vitamin D, genistein (soy) and selenium were not shown to affect PCa progression when assessed in an RCT design (20).

### 2.5.2. Prostate Cancer and physical activity

According to the World Health Organization (WHO) physical activity is the bodily movement generated by skeletal muscle which requires energy expenditure and includes
movement for leisure, transport and working purposes (116). Exercise is type of physical activity, which is repeated and planned with the aim to improve or maintain physical fitness (117). Physical activity has been suggested to play a protective role in PCa risk and progression, by increasing cardiorespiratory and muscular fitness, reducing BMI and improving fatigue associated with PCa treatment and health related quality of life (24,69,118-120).

Observationally, physical activity has also been found to be associated with reduced rates of PCa specific and all-cause mortality; however, RCTs where physical activity interventions were delivered to patients have not investigated the effects on PCa progression, but rather on proxy measures of survival and quality of life such as PSA levels, fatigue levels and fitness, most likely due to the fact that PCa progression can take many years to develop.

A systematic review of dietary and physical activity randomised interventions found that none of the physical activity interventions had consistent effects on PSA measures of progression in men who underwent hormone therapy, androgen deprivation therapy or radiotherapy and the same patterns were observed in RCTs which combined both nutritional and physical activity aspects (20). Different types of physical activities produce different changes in the body, for example cardiovascular exercise mainly benefits weight loss, cardiorespiratory fitness, and fat related measures, while resistance training stimulates and increases muscle mass. In the context of cancer, cardiovascular and resistance training can help with different aspects of the issues experienced by people living with cancer, such as decreased fitness, muscle loss and fatigue (121). Hackshaw et al. identified four RCTs which had a physical activity intervention and all of them included resistance training. There was no evidence of change in PSA levels in any of the trials in the intervention groups (resistance training alone, resistance training and cardiovascular training and cardiovascular training alone), however only one of the four trials was at low risk of bias, with the other three having an unclear or high risk of bias (20,120,122-124). Increased levels of physical activity, including brisk walking has been suggested to decrease PCa progression and PCa specific mortality, however only two randomised controlled trials have investigated the effects of brisk walking in men with PCa $(32,118,125-127)$. One pilot trial which randomised participants to group walks once a week for an hour and encouraged participants to reach 10,000 steps per day over 11 weeks, found increased high-density lipoprotein and decreased low-density lipoprotein and C-reactive protein levels in the intervention group compared to control (126). However, no indicators of progression such as PSA levels were investigated.

In another trial, which randomised men with localised PCa who underwent prostatectomy, to a brisk walking intervention for 30 minutes, 5 times a week, indicators of progression were not assessed $(34,127)$.

Many mechanisms have been proposed that link physical activity and PCa risk and progression. A recent review of exercise training in obesity and PCa suggests that exercise training can impact metabolic health (lower circulating glucose and insulin levels, reduce adipose tissue mass and whole-body inflammatory markers), as well as alter genes which are particular susceptible to hypermethylation following bouts of exercise training $(128,129)$. One proposed hypothesis is that physical activity triggers epigenetic alterations in the calcium release activated channel regulator 2A gene, a gene encoding a key regulator of calcium channels. A recent study found that men with localised PCa, who exercised vigorously once or more per week were found to have a reduced risk of developing metastases over the average follow-up time (11.3 years) (130). In the subset of participants with tissue samples the study found some evidence of differently methylated promoter region of calcium release activated channel regulator 2A between men who exercised vigorously once or more per week compared to men who did not (130). The evidence supports calcium release activated channel regulator 2 A as a mediator between physical activity and PCa progression.

While there is some evidence that increased physical activity, both cardiovascular and resistance training, may be beneficial to men with PCa, the effect on PCa progression remains uncertain.

### 2.6. Conclusions

PCa is one of the most commonly diagnosed cancer worldwide. Survival is generally high for localised cases, and there is a large proportion of cases diagnosed which do not pose clinical problems. However, some cancers advance fast and lead to worse outcomes. Due to the fact that some cancers are diagnosed while the disease is still indolent (i.e., not clinically significant), overtreatment is common, and this can lead to lifelong negative effects on urinary, erectile and bowl functions. There are no established or highly predictive indicators of PCa progression at present, with the best performing ones being the Gleason score and PSA readings over time. Risk factors for PCa progression are not yet established, however it is believed that diet and physical activity may play a role. As it takes a long period of time
for PCa to progress it is difficult to accurately establish risk and protective factors for prognosis and most of the research has been performed using proxies of progression such as PSA levels in randomised studies.

# Chapter 3. Introduction to 

## metabolomics

### 3.1. Introduction

The previous chapter focused on the clinical and epidemiological aspects of PCa, presenting measures of incidence, prevalence and survival, as well as dietary and lifestyle factors influencing PCa progression. This chapter introduces and reviews the background evidence relating to factors influencing PCa progression with a focus on circulating metabolomic measures, diet and lifestyle research. The chapter opens with a general introduction to metabolomics and its application in the field of cancer. It then presents current evidence of the application of metabolomics in nutritional and physical activity epidemiology and their potential role as intermediate markers of disease progression. The chapter ends by highlighting the complex mechanisms of PCa progression, with a focus on the role of various classes of metabolites in PCa metabolic pathways. Understanding the way metabolites affect PCa metabolic pathways is a crucial aspect in assessing the role that metabolites may have in PCa progression and their role as intermediate biomarkers of disease progression.

### 3.2. The field of metabolomics

Metabolomics is the large-scale study of metabolites, which are small molecules within cells, biofluids, tissues or organisms that collectively form the metabolome (131). The metabolome encompasses all metabolites in a biological sample and is dynamic in nature. Metabolomics studies small molecules that are end products of many functions in the body. Metabolites are generally defined as small, low-weight molecule intermediates and end products of biochemical reactions with a mass lower than 1,500 Daltons (Da) $(132,133)$. Metabolomic profiling includes the detection and quantification of metabolites and can reflect changes occurring from different conditions in lifestyle, diet and genetic factors (10). The assessment of metabolites, which are downstream of changes in genes and proteins, can indicate
functional alterations in pathways affected by multiple pathological states (133). Therefore, metabolomic profiling can provide a comprehensive and dynamic description of the biological state of an organism through the integration of genetic regulation, enzyme activity, and metabolic reactions (133).

Another advantage of metabolomics is the potential it holds to demonstrate altered cellular activity within various disease states, using a much smaller number of measures compared to other 'omics fields such as genomics or proteomics (133). Metabolomic analysis can be conducted both at cellular and systemic levels, by measuring metabolites in the peripheral system, for example in serum or plasma which can be particularly useful when assessing systemic metabolomic markers in metabolically complex diseases such as cancer (134). Advancements in the instruments used to detect and quantify metabolites, which are presented in more detail later in the chapter, along with developments in the field of bioinformatics, allowed improvements in the knowledge of the metabolome $(10,135,136)$. Better characterisation and understanding of the metabolome provides an opportunity to inform the use of circulating metabolites as markers of disease risk and progression since metabolites hold the advantage of being downstream of genetic and proteomic changes (133). Metabolites may be useful potential indicators of functional alterations in pathways affected by multiple factors, integrating enzymatic, genetic and metabolic activity in a dynamic profile $(133,137)$.

### 3.2.1. Untargeted metabolomics

Untargeted metabolomics refers to the analysis of all measurable analytes in a sample, including chemically unknowns(138). Untargeted metabolomic studies measure a large number of metabolites and are frequently used in hypothesis generating studies where no specific set of metabolites are planned to be investigated. Mass spectrometry (MS) is a technique that is widely used in the study of untargeted metabolomics and uses the ratio between mass and charge in complex molecules to identify and quantify the structure of the molecule $(11,139)$. The sample must first be ionised, and the ionised compounds, which produce peak patterns, allow for the identification and quantification of the original molecule. To reduce the complexity of the original molecule and perform MS, the sample undergoes a separation step which can be done through gas chromatography or liquid chromatography, the most used techniques for spectral data in MS (11,139-141). By
performing the separation step, this technique's specificity, sensitivity and reproducibility are enhanced (Table 2) (10).

Gas chromatography MS and liquid chromatography MS methods provide a relatively costeffective method to identify and quantify a large number of metabolites, achieve high sensitivity and specificity and identify novel metabolites and their physical and chemical properties, even at very low concentrations $(11,142,143)$. However, the separation steps are time consuming, costly and can lead to a significant loss in sample quantity, with some techniques impacting the classes of metabolites that can be investigated (9-11). From both a bioinformatic and statistical point of view, correctly analysing the vast number of metabolites represents a real challenge, with various techniques and methodologies being implemented to avoid misinterpretation of results $(11,139,142)$. This is a particularly difficult problem in studies with large populations where the value of metabolomics lies in the high through-put, the low sample volume required and the low costs of assays.

Figure 3.1: Definitions of measures of test performance in the context of Gas chromatography or liquid chromatography in metabolomics

- Sensitivity refers to the ability to detect a signal for the metabolite at very low concentrations
- Specificity represents measures of how well the technique performs at correctly distinguishing a metabolite from other substances present in the sample
- Reproducibility is the ability of the technique to display consistency in signals across replicated experiments


### 3.2.2. Targeted metabolomics

Targeted metabolomics refers to the measurement of a pre-defined group of metabolites, which are normally identified and quantified using an established platform (138). Similarly to untargeted metabolomics, targeted metabolomics can be quantified through MS and suffer from the same problems, such as separation step related issues in gas
chromatography and liquid chromatography. In addition, inter-laboratory reproducibility is an issue particularly important in biomarker discovery research, as in order for it to be widely used, it must be easily reproducible across different laboratories (144).

Another technique for targeted metabolomics is using proton $(1 \mathrm{H})$ nuclear magnetic resonance (NMR) spectroscopy, where the compound is placed in a magnetic field and the resulting resonant frequencies allow the identification of the different molecules it contains (137). NMR holds a few advantages over the MS technique: it is a non-destructive, easily quantifiable, and inherently quantitative technique that requires a simple sample preparation and thus has high throughput at low costs $(137,145-147)$. It also provides excellent reproducibility, with predictable spectra, and allows a precise structure determination of the molecules contained within the compound (145). However, the NMR method requires a larger sample quantity and has poor sensitivity, therefore is only able to detect molecules present at much higher concentrations than MS, leading to a lower number of metabolites that can be identified and quantified $(145,146)$.

### 3.2.2.1 The Nightingale Nuclear Magnetic Resonance platform

The Nightingale NMR platform (Nightingale Health@, Helsinki, Finland) is a low-cost highthroughput metabolomics method that uses two different approaches of identification and quantification of metabolites to provide metabolomic profiles of both lipoprotein measures and metabolites with a low molecular weight (148). The sample preparation is minimal and includes multiple automated steps which allows for 96 samples to be prepared in approximately 2 hours (148). Serum or plasma samples can be analysed, using 100 or $350 \mu \mathrm{~L}$ volume of blood, respectively, in both fasted and non-fasted samples. The platform (as released in 2014), yields 228 metabolic measures, 150 of which are primary concentrations, with the rest being derived ratios (149). There are 14 lipoprotein classes investigated, with each of these having 12 measures quantified (149). The processing of the spectrum for one serum sample is done automatically in approximately 9 minutes (148). The platform has been previously used in epidemiological studies in multiple disease areas and biobanks $(149,150)$.

### 3.3. Metabolomics in cancer

Cancer is complex and dynamic disease and remains challenging to prevent and treat due to the lack of understanding around its aetiology and progression (151). Hanahan and Weinger (2000) introduced the concept of the "six hallmarks of cancer", a set of underlying principles which were extended to eight in 2011, that aim to break down the complexity of the disease, by characterising the essential alterations in cell physiology which support malignant growth (Table 3.1) $(152,153)$.

It has been long established that one hallmark of cancer is altered metabolism that allows the increased nutrient requirements for proliferation of cancerous cells to be fulfilled $(154,155)$. Metabolic alterations in tumours have also been proposed; however, it was the recent developments in the biochemical and molecular biological tools in areas such as imagining, cell culture, genetics, proteomics and metabolomics, that advanced the understanding of the mechanisms and importance of metabolic alterations in cancer research (156). Fundamental to the six hallmarks of cancer are genome instability, inflammation and metabolic reprogramming (153).

Metabolomics is a relatively new field compared to other 'omics studies, such as genomics, proteomics or transcriptomics. However, it has captured great interest in cancer research, due to its ability to provide a characterisation of the tumours' cellular physiology and biochemistry, by using metabolites as intermediate and end-point indicators of altered molecular pathways (157). Changes in both intra- and extra- cellular metabolite levels, which are observed in cancer metabolic reprogramming, can greatly affect gene expression, cellular differentiation and the tumour environment, making metabolites an important area of focus in cancer research (156).

Table 3.1: The eight hallmarks of Cancer according to Hanahan and Weinberg $(152,153)$.

| The six hallmarks of cancer (152) : | The two emerging hallmarks of cancer (153): |
| :--- | :--- |
| Inducing angiogenesis | Evading immune destruction |
| Resisting cell death | Reprogramming of energy metabolism |
| Evading growth suppressors |  |
| Enabling replicative immortality |  |
| Activating invasion metastasis |  |

### 3.4. Metabolomics, lifestyle and diet

Lifestyle and dietary factors have long been believed to influence cancer risk and progression and many studies investigated the link between them (24). However, it is difficult and sometimes impossible to assess the relationship between cancer and diet and lifestyle factors, particularly in observational studies due to a wide range of limitations (158). Some of the most common limitations of nutrition research are, recall bias, measurement error, inability to split combined effects from closely related exposures and nutrient bioavailability which varies with food type and cooking methods (159).

By measuring downstream components, metabolomics may be more suited to measure dietary exposures and thus provide more accurate estimates of disease risk in nutritional epidemiological research (160). In addition, metabolites have also been shown to be correlated with intake levels for alcohol, vitamins and food groups (e.g. citrus fruit) and be a good predictor of these (160). Metabolomic signatures can also be used to reflect broader dietary patterns in addition to specific dietary exposures. For example, a metabolic signature which contains 67 metabolites was shown to robustly reflect the adherence to a Mediterranean diet and also be causally associated with risk of coronary heart disease and stroke (34).

The link between physical activity and metabolomics has been studied for decades, and a recent review of the effects of physical activity on the metabolome concluded that various degrees and types of physical activity are associated with quantifiable changes in the metabolome, measured in both MS and NMR spectroscopy (162). Measures of overall physical activity and exercise have been shown to be associated with changes in levels of
fatty acid, cholesterol, glycolysis related measures and amino acid. The type, intensity, frequency and timeframe of the physical activity exposures have been suggested to affect the metabolomic response, suggesting the metabolomic response to physical activity is multifactorial dose dependent (162-166).

To provide definitive evidence on the causal effect of any potential protective intervention (and its harms), appropriately powered and well conducted (i.e., unbiased) RCTs are the gold standard. RCTs are costly, time and resource intensive, particularly when a long follow-up time is required. In addition, it can be unethical and physically challenging for participants to complete an intervention over a long period of time, which would be needed for cancer risk and progression. Assessing the effects of the intervention on appropriately chosen intermediate biomarkers could provide preliminary evidence of the effect of the intervention on long term outcomes. For instance, a shorter dietary or lifestyle intervention could be implemented to investigate its effects on metabolomic markers which could then be investigated separately to assess their role in cancer risk and mortality. Metabolites are good candidates for intermediates in dietary and lifestyle interventions as they are markers of genetic and proteomic functions, and they can provide an overall picture of the effects of an intervention on multiple groups of compounds and pathways $(9,12,13)$.

There has been a great increase in the number of nutritional studies that use metabolomics since metabolomic analysis can provide a deeper insight into the effects of the dietary intervention on metabolomic perturbations (167). Metabolomics shows great potential in advancing the understanding between lifestyle and dietary factors and disease and a recent review of metabolomic methodologies in human nutritional studies emphasises the field's potential and the need of harmonization in the future (167).

### 3.5. Metabolomics and Prostate Cancer progression

### 3.5.1. Circulating metabolites

To evaluate the evidence linking PCa progression and circulating metabolites I conducted a literature search on PubMed.

I used the following search strategy:
(Prostat*[TITLE]) AND (Cancer*[TITLE]) AND (metabolite*[TITLE] OR metabol*[TITLE] OR AMINO*[TITLE] OR GLYCOLY*[TITLE] OR LIPID*[TITLE] OR LIPOPROTEIN*[TITLE] OR glycoprotein*[TITLE] OR CHOLESTER*[TITLE] OR PHOSPHOLI*[TITLE] OR GLYCER*[TITLE] OR lipoprotein*[TITLE] OR FATTY[TITLE] OR KETON*[TITLE] OR CREATIN*[TITLE] OR ALBUMIN*[TITLE]) AND ( [surviv*[TITLE] OR progress*[TITLE] OR metasta*[TITLE] OR death[TITLE] OR mortality[TITLE])

## Inclusion criteria

All publication abstracts were accessed and assessed for inclusion into the literature review. Publications were included in the literature review if they satisfied all the following criteria:

- Full text was available
- The study was metabolomic
- The study presented original findings, or it was a review of available evidence (methodological, case reports or ecological studies were excluded)
- Progression measures were investigated

Following abstract screening, full-text assessment followed, and studies were included if they fulfilled the following criteria:

- Performed in human subjects, in plasma or blood
- Presented original findings

Given the limited evidence linking circulating metabolites and PCa progression, the search term used was broad and identified 261 publications. 196 publications were excluded for the following reasons: not from metabolomic studies ( $\mathrm{n}=156$ ), were commentary or correction ( $n=14$ ), case reports ( $n=8$ ), methodology publications ( $n=7$ ) or ecological study $(n=1)$, text was not available ( $n=7$ ), had no progression indicator $(n=3)$. Of the remaining 65, 29 were in human subjects, 13 were evidence reviews, 10 were in animals, 18 were in cell lines and 1 was in both cell lines and animals. Since this thesis investigates the role of circulating metabolites in relation to PCa progression, this literature review focuses on circulating metabolomic evidence. There were 17 publications that investigated the link between circulating metabolites and PCa progression. Table 3.2 presents the sample sizes, findings, strength and weaknesses of each study.

Figure 3.2: Selection process for inclusion of publications into the literature search


## Study quality

Overall, studies only assessed a few circulating metabolites, albumin, glucose and cholesterol being the most investigated. Only 2 of the 17 studies used a metabolomic platform with high throughput and thus investigated a large number of metabolites $(16,168)$. Most of the studies included had a retrospective case-control study design, and did not allow for survival analyses, since time to event was not measured. Low power was a predominant limitation of the included studies, with low number of participants which were occasionally further split into PCa disease phenotypes $(169,170)$.

## Main findings

Both studies which quantified a large number of metabolites (109-763 measures) were underpowered, with low number of participants (under 300). Few individual metabolomic measures were found to be linked to increased risk of PCa mortality, such as N-oleoyl taurine and fatty acid measures (polyunsaturated, acyl glycyl, monohydroxy acids, dicarboxylic acids) (16).

Lipids
Lipids in general and total free cholesterol were shown to be inversely associated with overall survival and radiographic progression-free survival in men with metastatic PCa , who underwent different treatments (168). However, the evidence around cholesterol measures remains unclear. The largest study identified in this literature review (5,893 PCa cases) found weak evidence that cholesterol was associated with PCa mortality (171). However, the study did not have any information on tumour characteristics or socioeconomic factors. A study in the Finish cancer registry, which covers $99 \%$ of all cancers diagnosed in Finland, also found weak evidence of association of cholesterol and progression, although PCa specific mortality analyses were not undertaken due to small number of events.

## Fatty acids

The link between fatty acids and PCa progression were assessed in three studies (16,172174). One study that investigated the effects of eicosanoids, found that hydroxyeicosatetraenoic acid and arachidonic acid were positively and inversely correlated with advanced stages of PCa, however no survival analysis was performed (173). A study conducted in men with low-risk PCa, from a previously completed RCT, found that omega-3 fatty acids were not associated with PCa progression(172). The progression indicator was defined as moving from a Gleason score of 6 to a Gleason score of 7 , at the 6 -month repeat biopsies which were performed(172). Another study which assessed a large number of metabolites, found that fatty acids (polyunsaturated, acyl glycyl, monohydroxy acids, dicarboxylic acids) were associated with increased risk of PCa morality (16). The study used data from the Alpha-Tocopherol, Beta-Carotene Cancer Prevention study which recruited Finnish men who were aged 50-60 years old and smokers from 1985 to 1988. Follow up data was available through the Finish cancer registry, and 92 of the 197 cases included had a PCa death event(175).

## Amino acids, fluid balance measures and ketone bodies

Four studies assessed the effects of amino acids on PCa progression (16,169,170,176). All studies found evidence of a link between amino acids with PCa progression. N -oleoyl taurine was associated with increased PCa mortality, in the only study of the four that assessed PCa mortality (16). Medium-chain acyl-carnitine hexanoylcarnitine and trimethylamine were positively and long-chain acyl carnitine, acetylcarnitine, isovalerylcarnitine, l-carnitine and isovalerylcarnitine, inversely associated with PCa progression (169). The study had a case-control design and it included multiple progression indicators such as repeat PSA measures, local lymph node and bone metastases. Another study found that creatine, creatinine and glutamine were associated with increased levels in early PCa cases and reduced levels in more advanced cases (170). The study also found histidine, leucine, valine, lysine and acetate to be inversely associated with PCa progression (170). When assessing the link between the amino acids and progression, the study compared men with early, advanced, metastatic and castrate resistant PCa to men with benign prostatic hyperplasia but did not conduct survival analyses. Homocysteine, cystathionine and cysteine were higher in men with biochemical recurrence compared to recurrence-free men (176). The same study showed that sarcosine was higher in urine but not in serum for men with biochemical recurrence when compared to men without biochemical recurrence. Albumin and albumin-globulin ratios were investigated in 6 of the studies (177-182). The evidence supports a role for decreased albumin with increased risk of PCa progression, however the threshold has not yet been established. In addition, albuminglobulin ratio is showing to be a good predictor of PCa progression $(180,181)$. The studies that investigated albumin and the albumin-globulin ratio were generally in men with more advanced instances of PCa. Therefore, it is unclear from the available evidence what the link between albumin and PCa progression is, in men with localised PCa (16).

Glycolysis measures
Glucose was the most glycolysis measure investigated in the selected studies. In the largest study glucose was not associated with PCa progression (171). However, the study did not take into account socioeconomic and cancer staging factors and compared low versus high levels rather than continuous measures (171). Citrate was found to be lower in metastatic and castrate PCa cases compared to benign prostatic hyperplasia cases (170).

Table 3.2: An overview of studies which examined metabolomics in relation to Prostate Cancer progression

| Author <br> name, year <br> and PMID | Study design | Population <br> size | Metabolite | Findings |  | Limitations |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |


|  |  |  |  | Medium-chain acyl carnitine positively associated with PSA progression ( $\mathrm{p}=0.036$ )(169) |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Zoni(2019),$31842810$ | Case-control | 10 PCa free and 42 PCa cases |  | Long-chain acyl carnitine inversely associated with local and bone progression ( $\mathrm{p}=.034$ ) |  |  |
|  |  | 11: low risk PCa <br> 12 high risk with no | 17 carnitine metabolites | No strong evidence of association with lymph node progression and short, medium or long-chain acylcarnitines | Overall, the study had long follow up time for non-metastatic cases (minimum 5 years) | Low power, small numbers in each of the risk categories investigated |
|  |  | progression 10 high-risk with progression |  | Acetylcarnitine(0.016) and isovalerylcarnitide (0.01) inversely associated with PSA progression | Multiple progression indicators assessed (PSA increase, local, node and bone metastases) | Low follow-up time in the low risk PCa category (7 years) |
|  |  | 9 metastatic |  | Hexanoylcarnitine $(\mathrm{p}=0.035)$ positively associated with PSA progression |  |  |
|  |  |  |  | L-carnitine $(\mathrm{p}=0.006)$ and isvalerylcarnitine ( $\mathrm{p}=0.027$ ) inversely associated with lymph node progression |  |  |
| Zheng$\begin{aligned} & (2020), \\ & 31758937 \end{aligned}$ | Cross- <br> sectional | 76 | 20 | Metabolomic patterns more clearly | Allowed comparison between metabolites in plasma, urine and tissue | Case-control study design, not allowing for time to event analyses |
|  |  | participants: | metabolites | defined in tissue compared to |  |  |
|  |  | 18 benign | (lipid, | serum(170) |  |  |
|  |  | prostatic | glucose and |  |  |  |
|  |  | hyperplasia <br> 16 early PCa | amino acid metabolism) | Overall patterns identified 9 important differential metabolites in PCa | Compared progression across multiple PCa progression phenotypes | Low power- reduced number of participants across 5 categories |

11 advanced
PCa
23 metastatic
PCa
8 castrate-
resistant PCa
progression in serum samples: citrate
creatine, histidine, lysine, glutamine,
creatinine, valine, leucine, and acetate

Citrate was lower in metastatic and castrate resistant PCa cases compared to benign prostatic hyperplasia ( $\mathrm{p}=0.004$ )

Glutamine levels were inversely
associated with PCa progression ( $\mathrm{p}=0.001$ )

Creatine ( $\mathrm{p}=0.002$ ) and
creatinine $(\mathrm{p}=0.005$ ) were inversely
associated with progression from benign
hyperplasia and castrate-resistant PCa

Trimethylamine was higher in early PCa compared to metastatic and castrateresistant PCa ( $\mathrm{p}=0.008$ )

|  |  |  |  |  | Large number of participants | Only quantified albumin |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Cook $\begin{aligned} & (2006), \\ & 16740758 \end{aligned}$ | Prospective cohort | 643 PCa <br> metastatic <br> cases | Albumin | Albumin was not associated with overall survival(177) | Clearly defined inclusion criteria as data was generated via a multicentre randomised controlled trial | Survival analysis outcome was overall survival not PCa specific mortality |
| Stabler (2011), $21853037$ | Nested casecontrol | 58 PCa cases, <br> 28 with <br> progression | Sarcosine, dimethylglyc ine, | Homocysteine, cystathionine and cysteine were higher in the progressed men (p<0.001) (176) | Age-matched cases and controls <br> Model adjusted for serum PSA< Gleason | Case-control study design, not allowing for time to event analyses |



| Moreel (2014), 24824038 | Prospective cohort | 120 <br> participants | Omega-3 <br> fatty acids | Eicosapentaenoic acid was higher in the non-progressed cases compared to the progressed cases in tissue but no association was observed in serum (172) | Randomised controlled trial design with welldefined inclusion criteria | Low power to detect association between serum omega-3 fatty acids and PCa progression |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  |  | Only total omega-3 serum measures were assessed |
|  |  |  |  |  |  | PCa progression indicator is not widely use in literature (Gleason score change from 6 to 7 ) |
| Pan (2021), <br> 34414685 | Prospective <br> cohort | 2,205 <br> participants | Albumin | Patients with low albumin had worse overall survival pre and post chemotherapy ( $\mathrm{p}<0.001$ ). (182) <br> Albumin independently predicted overall survival in both trials $(\mathrm{HR}=1.54$, $95 \% \mathrm{CI}: 1.34-1.78$; and $\mathrm{HR}=1.40$, 95\%CI:1.21-1.64) | Participants originated from two clinical trials, thus the studies had clearly defined inclusion criteria <br> Patients in the two trials followed different treatments | Patients not representative of most PCa cases since these were men with advanced stages of disease <br> Generally, included men had better health than men with metastatic PCa |
| $\begin{aligned} & \text { Lin (2021), } \\ & 34638448 \end{aligned}$ | Prospective <br> cohort | 274 <br> participants | Lipids and cytokines (109-763) | Lipid levels were generally associated with poorer overall survival, and were enriched with sphingolipids, particularly ceramides(168) | Three different cohorts of participants <br> Prospective study | Small number of participants compared to number of measures being investigated |
|  |  |  |  | Higher level of free cholesterol was associated with shorter overall survival | Multiple treatments |  |

Radiological progression free followed
similar associations with lipids as did
overall survival


|  |  | mortality in a time dependent analysis, <br> however the effects were only short <br> lived | through a centralised system which covers <br> 95 of prescriptions |
| :--- | :--- | :--- | :--- | :--- |


haemoglobin, neutrophil count and
Eastern Cooperative Oncology Group
performance status.


### 3.5.2. Non-circulating metabolites

Given the limited literature identified on circulating metabolites and PCa progression, some background evidence generated from mechanistic (metabolomic, genetic and proteomic) studies using PCa tissue, animal or PCa cell lines is provided. These studies generally focused on cellular pathways, which are not investigated in the Thesis, but are used in interpreting findings in my results chapters, particularly where no circulating metabolomic evidence exists. To adequately present these studies, manual searches were conducted aiming for more mechanistic evidence in PCa and its progression.

### 3.5.2.1. Introduction to Prostate Cancer molecular mechanisms

PCa cells produce citrate, PSA and polyamines (e.g. spermine) and therefore have a specific metabolic profile $(132,133,186)$. PCa progression has been linked to activation of androgen receptors, lipid and amino acid pathway abnormalities, and more generally to a deregulated metabolism $(187,188)$. In addition to the unique metabolic profile at tumour level, there is evidence to suggest that serum levels of metabolites may differentiate low, high-grade and lethal disease (189-191). This and the lack of prognostic factors for PCa progression make serum metabolites great candidates for PCa aggressiveness and progression biomarkers. However, as seen earlier in this Chapter, due to the long latency period of PCa to progression, few studies have assessed the role of metabolomics in PCa progression. Thus, the next section presents the most well-known metabolic pathways and the metabolites involved in these which have been shown to play a role in PCa risk and progression, in cellular, animal and human studies.

### 3.5.2.2. The Krebs cycle

The Warburg effect refers to the ability that cancerous cells have to produce adenosine triphosphate (ATP), an energy-carrying molecule though a glycolytic pathway (anaerobic) even in the presence of oxygen, unlike noncancerous cells which use oxygen (aerobic) during the Kerbs cycle $(132,154,155,192)$ (Figure 3.2). However, in the case of PCa, the Kerbs cycle and more recently, oxidative phosphorylation, have been found to also play an important role in cellular metabolism (193). In contrast to noncancerous prostate cells, cancerous PCa cells have a very low amount of citrate. This is a consequence of the very low levels of zinc found in PCa tissue, in the prostate peripheral zone which is the most
voluminous part of the prostate gland and the most common area for malignancy ( $85 \%$ ) $(186,194)$. Low levels of zinc in PCa cells is an unique aspect of PCa which allows extra ATP to be produced aerobically by prolifically allowing oxidative phosphorylation to occur, which leads to accelerated neoplasm formation (186). Since PCa is not glycolytic (i.e does not rely on glucose for the higher energy demand) and there is a high demand of lipids for membrane formation and intercellular signalling in cell proliferation, an increased lipid biosynthesis is a required metabolic change associated with prostate malignant transformations (133). For the lipid biosynthesis to occur, citrate must be converted to acetylcoA, a process that recruits a number of enzymes involved in fatty acid and cholesterol synthesis, which are androgen-regulated and have been shown to be abundant in PCa cells $(133,195,196)$. Another observed source for the acetyl CoA production has been suggested to be acetate obtained from circulation, with cancer patients exhibiting much lower levels of acetate in arterial blood compared to cancer free individuals $(194,197)$.

Figure 3.3: Net citrate production metabolic pathway.


Adapted from Costello and Franklin (2006) (194)

### 3.5.2.3. Fatty acid metabolites

Overall, observational epidemiological studies found that PCa is associated with increased lipogenesis, high levels of triglycerides, low high-density cholesterol (HDL) and high levels of low-density lipoprotein (LDL) in blood (198). PCa is characterised by an increased fatty
acids synthesis and overexpression of lipogenic and lipid-modifying enzymes at tumour level $(188,199,200)$. For example, carnitine which is a metabolite derived from lysine and methionine amino acids, has been shown to be a regulator of adaptive metabolic reprogramming in PCa cells $(188,201)$. Stearoyl-CoA desaturase-1 enzyme which coverts saturated fatty acids into monosaturated fatty acids is overexpressed in PCa tissue, androgen sensitive and castration-resistant PCa cells and has been linked to neoplastic transformation and tumour cell proliferation and invasiveness $(188,202,203)$.

### 3.5.2.4 Cholesterol related metabolites

As described in the previous sections, biosynthesis plays an important role in PCa. Furthermore, cholesterol makes up a third of the plasma membrane lipids, and thus has a critical part in membrane synthesis and steroidogenesis, two essential aspects of PCa survival and tumour upkeep $(188,198,204)$. Cholesterol is also the largest compound contained in the "lipid rafts", which are glycoprotein lipid membrane microdomains with regulatory functions in signal transduction and extracellular stimuli channelling (205). The concentration of cholesterol in lipid rafts and the degree of saturation of fatty acid chains in the phospholipid layer regulate these signals (188). The composition of the lipid raft can be in turn directly regulated by pathways that have the ability to alter cholesterol and lipid metabolism (206).

Cholesterol's important role in PCa progression is supported by the fact that cholesterol is a precursor for androgen synthesis and an agonist of the steroidogenic genes, with high levels of circulating cholesterol (hypercholesterolemia) in PCa patients being associated with higher levels of intratumor androgens and tumour cell proliferation $(188,204)$. Another adaptation of PCa cells in cholesterol biogenesis is their ability to esterify cholesterol to avoid cellular toxicity, while maintaining high cellular cholesterol levels, independent of free cholesterol, through the permanent activation of sterol regulatory element binding proteins (188). Sterol regulatory element binding proteins, which are regulated by androgen receptors, are elevated in PCa tissue, with levels dropping following androgen therapy (207). In addition, sterol regulatory element binding proteins are also activated in the phosphoinositide 3-kinases, PI3K pathway, which is in turn is activated by loss of the phosphatase and tensin homolog gene expression, a common (70\%) occurrence in metastatic PCa (188,196,208,209).

Statins, a cholesterol lowering drug, has been recently shown to improve biochemical recurrence free 5 - and 10-year survival in men who underwent radical prostatectomy (210). A prospective observational study in men with metastatic castration resistant prostate cancer undergoing taxane treatment, found that high serum cholesterol levels were independently (adjusted for clinical factors) correlated with shorter biochemical progression/first progression free survival, supporting the hypothesis that cholesterol metabolic pathways may play a role in PCa progression (187).

### 3.5.2.5. Amino acid metabolites

Sarcosine is an amino acid and an intermediate product in the synthesis and degradation of glycine. As seen has been suggested to play a role in PCa progression by stimulating metastatic PCa cells $(157,191,211,212)$. Previous studies that looked at metabolomics and PCa progression found that urinary and particularly tissue sarcosine were high in PCa cases which were progressing or had progressed to metastases, suggesting tissue sarcosine levels may be useful in monitoring or predicting disease progression $(132,191)$. Urinary and serum sarcosine were also associated with PCa and serum sarcosine had a very good discriminating performance (area under the receiver operating curve (AUROC)=0.97) for newly diagnosed PCa $(213,214)$. The Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial found an association of serum sarcosine with low grade PCa which was stronger than that of high-grade PCa and suggesting that plasma sarcosine may be a good biomarker for PCa aggressiveness (215). However, the association was stronger if the blood was drawn within two years of diagnosis and since $40 \%$ of the small number of aggressive PCa cases investigated had their bloods drawn 5 years post diagnosis, it is uncertain whether sarcosine is truly only associated with low-grade PCa or the observation is an artefact of low power (215). Since the association of sarcosine and low-grade PCa was stronger for samples taken closer to diagnosis, sarcosine may have the potential to be an early diagnostic measure (215).

Furthermore, recent evidence shows that sarcosine has a very good discriminating capability at differentiating between prostatic intraepithelial neoplasia (AUROC $=77 \%$ ) and PCa (AUROC $=79 \%$ ) in relation to noncancerous prostate hyperplasia (214). Although there is evidence for both urinary and serum sarcosine to be a potential marker of PCa diagnosis and progression, more replication work is required to fully assess the predictive value of sarcosine in this case $(132,189,216)$. Some replication studies found weak associations for
sarcosine as a biomarker of PCa; however, the discrepancy in findings is most likely due to the differences in sample selection, processing and analysis, a particular issue in nonstandard metabolomic analysis, as well as the limited sample sizes of the studies ranging from 25 to 2244 participants ( $132,191,213-218$ ).

Glutamine is an amino acid that is used in the biomass cycle for energy production (187). Glutamine metabolism has been found to be regulated by both AR signalling and MYC proto-oncogene, a regulator oncogene associated with neuroendocrine differentiation, introducing another pathway which may be critically involved in PCa progression $(187,219,220)$. Glutamate, the catabolised compound resulting from glutaminolysis, has been shown to be correlated with Gleason scores $\geq 8$ and was suggested to be a promising biomarker candidate for PCa aggressiveness (220). A recent study in patients with metastatic castration resistant PCa receiving treatment with taxanes, found that high levels of plasma glutamine were associated with shorter biochemical/clinical progression-free survival and overall survival in men with metastatic castration resistant PCa, introducing the potential role of glutamine metabolic pathways in PCa progression (187).

### 3.5.2.6. Glucose and related metabolites

Despite the increased requirement for energy, which in most cancers relies on aerobic glycolysis (Warburg effect), PCa cells use active glycolysis and oxidative phosphorylation and only rely on the Warburg effect for energy production in metastatic PCa (221-226). The increased glycolytic activity in advanced PCa cells supports cell proliferation and migration by altered activity and expression of key regulator enzymes such as glucose transporters, hexokinasae and phosphofructokinase (221,227-230). Glucose transporter 1, a subtype of glucose transporters, is overexpressed in PCa and has been associated with glycolytic activity, high glucose uptake and enhanced cell proliferation (207,229-231). Pyruvate is the end product of the glycolytic process and is converted into lactate, which has been suggested to play an important role in PCa $(226,232,233)$. In cell models lactate promotes proliferation of PCa cells, decreases apoptosis and increases migration and invasion of castrate resistant PCa cells (221,233-236). The glucose metabolic alterations presented illustrate the many complex mechanisms that are involved in PCa progression.

### 3.6. Chapter summary

With its unique metabolic profile and the lack of reliable predictive biomarkers of progression, localised PCa remains a challenge to manage. This is because many localised PCa instances remain indolent throughout a man's lifetime, but some will progress to metastasize and to cause death, and it is very difficult to accurately distinguish at diagnosis between indolent versus more aggressive PCa cases. The circulating metabolome, through its ability to detect changes at very small levels, can provide a comprehensive picture of the biological activity of PCa at various stages of the disease, and thus provide mechanistic insights into disease progression. Understanding how the multiple pathways that are involved in PCa metabolism work on their own, and interact with each other, is a complex and vital aspect of elucidating the mechanisms of progression. In addition, metabolomics is a useful tool for assessing the effects of diet and lifestyle interventions on PCa progression, by providing a detailed snapshot of metabolomic processes footprints. In conclusion, metabolomics can help investigate the causal and non-causal predictors of PCa progression.

# Chapter 4. Methodology 

### 4.1. Chapter introduction

This chapter introduces the five major datasets - from the ProtecT, CAP and PrEvENT randomised controlled trials (RCTs), UK biobank (UKBB) and PRACTICAL consortium and the statistical methods used throughout this Thesis. The latter are: the general analytical tools used for metabolomic data (Chapters 5 and 6); survival analysis (Chapter 5); intention-to-treat analysis (Chapter 6); instrumental variable analyses (Chapters 5 and 6); and the statistical methods employed in validating a predictive biomarker (Chapter 7).

### 4.2. Study populations

### 4.2.1. The Cluster randomised triAl of PSA testing for Prostate cancer and The Prostate Testing for Cancer and Treatment trials

### 4.2.1.1 Trials overview

The Cluster randomised triAl of PSA testing for Prostate cancer (CAP) trial is an RCT that aims to test the effectiveness of a single PSA test in men aged 50-69 years, in reducing PCa mortality and to assess whether it is cost-effective (237). The trial compares the group who received an invitation to attend population-based PSA testing for prostate cancer with standard National Health Service (NHS) care.

The Prostate Testing for Cancer and Treatment (ProtecT) trial is embedded within the intervention arm of the CAP trial (237) (See Figs 4.1 and 4.2). Men who responded to the invitation, had a PSA test as part of the CAP trial and who were diagnosed with localised PCa, were considered for entry into the ProtecT trial (237). The ProtecT trial aimed to assess the comparative effectiveness of the three major treatments at the time for localised PCa: radical prostatectomy, external beam three-dimensional conformal radiotherapy and active
monitoring. Some participants from the ProtecT trial were also followed up through the CAP trial. This thesis used data collected from both the CAP and ProtecT trial. However, given that the ProtecT trial was imbedded within CAP, for simplicity, with the exception of this Chapter, I will only be referring to the data as originating from the ProtecT trial.

The Prostate Mechanisms of Progression and Treatment (ProMPT) study is embedded in the ProtecT study and was responsible for the collection of samples for basic research. The aim of the ProMPT study was to help explore the molecular pathologies and mechanisms for tumour progression and developing new treatment strategies and novel markers. Men who attended the recruitment appointment for ProtecT were asked to complete an additional consent form for inclusion in the ProMPT study.

### 4.2.1.2 Ethical approval and consenting

All men provided written informed consent to be included in the ProtecT (ISRCTN08435261) and CAP (ISRCTN92187251) studies as well as the linked Prostate Cancer: Mechanisms of Progression and Treatment (ProMPT) study (ISRCTN20141297), the latter being responsible for the collection of the serum samples, BMI, exposure and diet data used in the present study. ProtecT (MREC/01/4/025) and ProMPT (MREC/01/4/061) received Multicentre Research Ethics Committee (MREC) approval (01/04/025). The CAP trial additionally had ethics approval for the flagging and reviewing of medical records for men with PCa (MREC/03/4/093 and MREC/03/4/093). Individual informed consent was sought from men who were alive at the time when the trial team were notified of a PCa diagnosis (237). The CAP trial had PIAG/NIGBECC approval which allows medical records to be reviewed post-mortem if a man died of a cause potentially linked to PCa (PIAG 1-05(f)/2006) (237). This applied to men who were no longer alive at the time of the notification and had no previous objection to their medical records being used for research (237).

### 4.2.1.3 Recruitment

Recruitment into the feasibility pilot of the ProtecT trial was undertaken between June 1999 and September 2001, in three cities in England and in 24 primary care centres (238). In 2001, recruitment into both trials in nine cities across England, Scotland and Wales commenced and finished in $2009(239,240)$.

As previously mentioned, the CAP trial is an extension of the ProtecT trial, which randomised primary care centres in the UK to either undertake the ProtecT prostate specific antigen (PSA) testing intervention or standard UK NHS management. The CAP trial created clusters of general practices which were randomised, blocked and stratified by geographical area by an independent statistician. The general practices were invited to participate into the trial and written consent was obtained. The overlap between the CAP and ProtecT trial recruitment is presented in Figure 4.1. In the control arm of the CAP trial, all men aged 5069 years, registered at the recruited practices were eligible to be included (239). Men were not contacted about their participation in the study, however they had an opportunity to opt-out from being followed-up through the general practice (237).

Figure 4.1: The overlap between the CAP and embedded ProtecT trial.


PSA $=$ prostate specific antigen
NHS=National Health Service
ProectT $=$ Prostate Testing for Cancer and Treatment
Reproduced with no changes from Turner et al, 2014, " Design and preliminary recruitment results of the Cluster randomised triAl of PSA testing for Prostate cancer (CAP)", license CC BY3.0 (237).

In the intervention arm of the CAP study, all men identified followed the ProtecT study protocol and had one PSA test. Following randomisation of general practices, in total, 337 practices were assigned to participate in the Protect trial. In the intervention arm (Figure 4.2) all registered men aged 50-69 years old were sent a written invitation for a PSA test, except for the following:

- Men with a life expectancy less than 10 years
- Men with a previous malignancy (except skin cancer)
- Men who underwent renal transplant or who were on renal dialysis
- Men with major cardiovascular or respiratory comorbidities
- Men who underwent bilateral hip replacement

Men who accepted the invitation were provided with an information sheet and a nurse appointment. At the appointment, the process was explained, trial eligibility was assessed, written consent was sought and a blood sample was taken for a PSA test amongst men who consented. Men who had a PSA result of at least $3.0 \mu \mathrm{~g} / \mathrm{L}$ were invited to attend another appointment where a rectal examination and a standardised ten-core transrectal-ultrasound guided prostate biopsy were performed. Men with a PSA higher than $20 \mu \mathrm{~g} / \mathrm{L}$ were excluded from randomisation in the 3-arm ProtecT treatment trial, as these were likely advanced PCa cases so were referred for specific NHS management.

The flow of men through the recruitment process is presented in Figure 4.1. Cancer diagnosis and staging were done at each centre based on PSA measurements, biopsy results and isotope bone scanning (for participants who had a PSA $\geq 10 \mu \mathrm{~g} / \mathrm{L}$ or Gleason score higher than 7). MRI was used for staging in some centres in the later part of the trial. Men diagnosed with benign prostate disease, or locally advanced (T3, T4) or advanced PCa, were excluded from the ProtecT treatment trial and were managed within routine NHS urological services. Men with a benign first biopsy were offered further biopsies if they had a free-tototal PSA ratio below $11 \%$, atypical small acinar proliferation or high-grade prostatic intraepithelial neoplasia. No further ProtecT trial follow-up occurred after the first round of PSA testing for PCa free patients or after referral to NHS for cases. However, all men were followed up for cancer and morality outcomes via linkage to NHS Digital under the auspices of the CAP trial.

Figure 4.2: Flow diagram of the Prostate Testing for Cancer and Treatment and the Cluster randomised triAl of PSA testing for Prostate cancer trial from invitation to diagnosis


PSA=Prostate specific antigen; PIN= Prostatic intraepithelial neoplasia; ASAP=Atypical small acinar proliferation Adapted from "Active monitoring, radical prostatectomy, or radiotherapy for localised prostate cancer: study design and diagnostic and baseline results of the ProtecT randomised phase 3 trial" , Lane et al, 2014, license CC BT 4.0 (239).

### 4.2.1.4 Randomisation and follow-up processes

Men who consented to be included in the ProtecT 3-arm treatment trial were randomly assigned to one of the 3 treatments using a computer system, which stratified by site and used stochastic minimisation to balance the trial arms in terms of age, Gleason score groups $(<7,=7,>7)$ and mean PSA results (from baseline and first biopsy measurements). Men who declined randomisation were followed-up identically to trial participants, and form part of a cohort within the study design and are referred in this thesis as 'eligible to be randomised'. Progression of PCa in the ProtecT trial was assessed at annual nurse-led follow-up appointments following a detailed schedule for each trial arm. Participant questionnaires on anxiety, depression, urinary symptoms, sexual function, and treatment-related quality of life were collected at 6 months post the first information appointment and every 12 months thereafter.

### 4.2.1.5 Follow-up variables

Men who were not 'eligible to be randomised' in the ProtecT trial were followed-up through the CAP study for cancer diagnosis and mortality status, but no other formally measured progression indicator (237). This was done passively through the Healthcare and Social Care Information Centre system where cancer registration or death events are flagged. For men who died, medical records were searched, and clinical facts were recorded by a researcher blinded to the death certificate information (237). An international Cause of Death Evaluation committee, blinded to trial arm and each man's death certificate, reviewed the information extracted by the researcher on all possible PCa deaths and assigned a definite, probable, possible, unlikely or definitely not PCa status (237). Definite and probable PCa mortality after a median 10-year follow-up is the primary outcome of the trial and for the purpose of all analyses in this thesis they are defined as PCa-specific deaths.

Similarly, to the CAP trial, the primary outcomes of the ProtecT trial was definite or probable PCa specific death at a median of 10 years of follow-up ascertained and adjudicated as above for the CAP trial. Multiple secondary outcomes, such as all-cause mortality, progression indicators (incidence of metastases, clinical progression), treatment complications, as well as patient reported outcomes were measured.

### 4.2.1.6. Variables used in this thesis

### 4.2.1.6.1. Progression variables

In the ProtecT trial, clinical progression was defined as evidence of regional metastases, diagnosis of clinical T3 or T4 disease, long-term androgen-deprivation therapy, ureteric obstruction, rectal fistula, or the need for a urinary catheter owing to local tumour growth. Metastatic disease was defined as bony, visceral, or lymph-node metastases on imaging or PSA levels above 100 ng per millilitre. PCa specific death was available for both the ProtecT and CAP trial. Where ProtecT mortality data was unavailable (e.g., in the case of 'not eligible to be randomised' participants), mortality status was assigned by using information gathered through the CAP trial.

### 4.2.1.6.2. Metabolomic measures

Baseline serum samples were collected as part of the ProMPT study, which is the linked study to the ProtecT trial and responsible for the collection of the serum samples. These were collected at the time of the participant's first visit (at the prostate clinic check) and were available for both cases and controls. The samples were aimed to be processed at the prostate clinic check (inverted and centrifuged) and stored in a cool box until they arrived in the laboratory the same day. In the laboratory all samples were checked and frozen at -80 Celsius. The samples were stored in the laboratory until they were shipped to the biorepositories. Samples were shipped again for metabolomic analysis. All shipping was done using dry ice. Proton nuclear magnetic resonance (NMR) was used to identify and quantify two hundred and twenty-seven metabolites using the Nightingale NMR platform (Nightingale Health®, Helsinki, Finland). Details of the platform has been described in Chapter 3 section 3.2.2.1.

### 4.2.1.6.3. Other variables

This thesis used demographic (sex, age, socioeconomic class), anthropometric (height, weight, BMI), lifestyle (physical activity, alcohol, smoking) and morbidity measures (heart attack, stroke, angina, heart failure, hypertension, high cholesterol, nervous trouble or depression, asthma, bronchitis). Overweight and obesity was defined as having a BMI higher than $25 \mathrm{~kg} / \mathrm{m}^{2}$. Harmful drinking was defined as having two or more alcoholic drinks daily, while the upper limit of alcohol consumption for moderate drinking was one
drink almost daily. PCa specific variables (T stage, Gleason grade and PSA level) were all obtained using the standardised definitions above. Men were split into two categories of Grade based on their Gleason scores: low (lower or equal to 7) and high (Gleason of 8 and higher).

### 4.2.2. The Prostate cancer: Evidence of Exercise and Nutrition Trial

### 4.2.2.1. Trial overview

The Prostate cancer: Evidence of Exercise and Nutrition Trial (PrEvENT) trial is a feasibility RCT which aimed to investigate if men who have undergone radical prostatectomy adhered to nutritional and physical activity interventions and assessed the effects of the intervention on intermediate outcomes. The full details of the trial protocol have been previously published (34).

### 4.2.2.2. Ethical approval

The PrEvENT trial (ISRCTN9904894) obtained full ethical approval from Cornwall and Plymouth Research Ethics Committee (REC) (ref: 14/SW/0056) and all men provided informed consent.

### 4.2.2.3. Participants and randomisation

Men who were diagnosed with localised PCa and whose planned management was radical prostatectomy were invited into the PrEvENT trial in one NHS Trust in the South-West of England. All men due to have radical prostatectomy who consented to participate in PrEvENT were first enrolled in a baseline cohort, completed preoperative questionnaires, and had bloods sampled (Figure 4.3). Six weeks after undergoing radical prostatectomy, men in the baseline cohort were then re-approached to be randomised within the PrEvENT trial. Exclusion criteria included inability to give informed consent, follow-up unavailability, a decision by the treating clinician which deemed the patient unsuitable to participate, major comorbidities that would make the participation difficult or impossible (e.g., uncontrolled congestive heart failure, angina, myocardial infarction, respiratory difficulties requiring oxygen or hospitalisation), regular lycopene supplementation or uptake of high levels of
physical activity. Randomisation was performed online using a system that ensured concealment of allocation.

Figure 4.3: The Consolidated Standards of Reporting Trials diagram of the The Prostate cancer: Evidence of Exercise and Nutrition Trial.


RCT=Randomised controlled trial; $P A=$ Physical activity
Adapted from "Phase II randomised control feasibility trial of a nutrition and physical activity intervention after radical prostatectomy for prostate cancer", Hackshaw-McGeagh LE, et al. (2019), license CC BY 4.0 (127).

### 4.2.2.4. Interventions and follow-up

The PrEvENT trial had a $2 \times 3$ factorial design and randomised men to both a dietary and a physical activity intervention. The dietary intervention consisted of either lycopene oral supplementation, a plant-based diet advice or no change in diet (control group). Men in the lycopene group were asked to ingest one 10 mg lycopene capsule (Holland and Barrett) daily which was supplied and paid for by the trial. The advice for the plant-based diet arm was to eat as many portions of fruit and vegetables as possible, with a minimum of 5 portions, to reduce dairy milk intake as much as possible, by swapping dairy milk with a non-dairy milk alternative. Men in the control diet group were asked to continue as normal with their diet.

Since the trial was set-up, the term plant-based diet has taken a different meaning and it now generally refers to an exclusive plant-based diet, which contains no animal products. Although the intervention is referred to as plant-based diet in the trial, in this thesis, with the exception of this Chapter, where PrEvENT trial terminology is used, I will be referring to the intervention as 'dietary advice'. This is to avoid confusion on the trial's intervention and the current literature around plant-based diets vs increased fruit and vegetable consumption.

Men in the physical activity intervention group were asked to walk at a brisk pace for 30 minutes on at least 5 days of the week in addition to their current levels of physical activity. For the purpose of this trial, physical activity and exercise terms were used interchangeably, however the definitions and differences of each were discussed in Chapter 2. Men in the physical activity control group were advised to continue with their current level of physical activity.

Men were asked to adhere to the intervention for 6 months and were prompted to attend nurse-led research appointments at trial baseline (randomisation), 3 - and 6 -months post randomisation. Blood samples were collected at baseline and 6 months, and questionnaires (physical activity, diet, mood, prostate symptoms) were completed at baseline, 3- and 6months post randomisation. Men received regular motivational, inspirational (e.g. recipe ideas) and gratitude communications in the form of text, emails and relating to their participation in the trial (34). Messages personalised for each interventional group were sent out at $1,2,5,8,13,15$ and 18 weeks post-randomisation (34).

### 4.2.2.5. Main variables

The primary outcomes of the trial were two feasibility related measures: randomisation rates and intervention adherence at 6 months. Secondary outcomes were blood biomarkers (antioxidants, metabolomic data, PSA) and self-reported patient outcomes (urinary symptoms, psychological factors, quality of life, fatigue, dietary and physical activity). Other data that were collected were demographic (sex, age, socioeconomic class) and anthropometric (height, weight, BMI) measures.

### 4.2.3. United Kingdom Biobank (UKBB)

The UKBB is a large-scale study that was established with the aim of allowing detailed investigations of genetic and nongenetic factors in middle and old age diseases. The UKBB is a prospective population-based study which recruited over 500,000 participants aged 40-69 years between 2006 and 2010. There were 22 assessment centres across the UK, and participants from a varied socioeconomic, ethnic and urban-rural mix took part. Data was collected on a wide range of phenotypes, through questionnaires, interviews, tissue collection and physical measurements (Appendix A, Table A2). Additional assessments have been undertaken in UKBB to enhance the phenotyping data (Appendix A, Table A3). This thesis used genetic data from 115,078 participants (males and females) in UKBB which identified 2,542 unique genetic variants associated with 245 metabolites (Appendix A, Table A4) (35).

### 4.2.4. The Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL) consortium

The PRACTICAL consortium includes 133 different study groups from across the world and aims to collaboratively combine data from studies to investigate the inherited risk of PCa and its progressions (23). The consortium was established in 2008 and at present it contains samples from over 120,000 PCa cases with mortality status and genetic data available for

75,672 participants of which 7,914 died of PCa. This thesis used the meta-analysed genetic data generated through a Cox regression analysis with left truncation and right censoring of the 75,672 participants from 43 studies (Appendix A, Table A4).

### 4.3. Observational statistical methods

### 4.3.1. Metabolomic methods

### 4.3.1.1. Identification and quantification of metabolites

As described in Chapter 3 there are multiple methods for identifying and quantifying metabolites, such as mass spectrometry and NMR. This chapter discusses NMR metabolites measured and identified using the Nightingale NMR platform (see Chapter 3, section 3.2.2.1), which assays over 200 metabolites, mostly lipid parameters, and generates ratios of metabolites, which are used to proxy enzyme activity (241).

### 4.3.1.2. Metabolomic data methods

### 4.3.1.2.1. Data cleaning

The results obtained from the Nightingale NMR platform have undergone internal quality control and therefore minimal additional data cleaning was required. Basic data checking and cleaning was performed to ensure the metabolites were correctly assigned to men in the PrEvENT trial. All metabolites with undetected levels of metabolites were set to the minimum value observed in the dataset.

### 4.3.1.2.2. Metabolite processing

All individual level results that were more than 4 times the interquartile range (IQR) away from the group median were considered outliers, given that they were more likely to be analytical rather than biological outliers and they were dropped(242). To compare the magnitudes of the associations of metabolic traits with different scales and units, all metabolic trait concentrations were standardised using $z$-scores in all studies in this thesis $(243,244)$. Chapter 8 uses additionally $\log$ (natural) transformed the metabolites to reduce skewness and follow data processing procedures consistent with those of my Nightingale Health collaborators, an important step for between-group comparability.

### 4.3.1.2.3. High correlation adjustment

A large number of the metabolites were analysed many of which were also highly correlated with each other, and the samples sizes of the studies assessed were small in comparison to the number of participants. To overcome the high correlation of metabolic traits and the multiple testing, a Principal Component Analysis (PCA) method was employed (245-247). PCA is a statistical method that allows multiple variables to be assessed simultaneously, to reduce dimensionality in multivariable analysis $(245,246,248)$. This data-driven approach investigates the dataset as a whole and generates a number of components that explain the variance observed, by assessing the metabolite correlation matrix. Each principal component can be linked to an amount of variance explained, using eigenvalues. The investigator should therefore set a suitable level of variance that would be appropriate in the study, and based on the variance explained by the eigenvalues, determine the number of components. The number of components will then be used to adjust the significance threshold using the equation below:

$$
\boldsymbol{\alpha}_{\text {adjusted }}=\frac{\boldsymbol{\alpha}_{\text {set }}}{N_{\text {Principle Components }}}
$$

PCA method has been used in genomics and metabolomic epidemiological studies as a tool to visualise structure and reduce data dimensionality $(245,248,249)$. I set the variance parameter at 95\% based on previous metabolomics epidemiological studies that took this statistical approach and I calculated the adjusted significance threshold using the number of principal components based on the explained variance of $95 \%(243,244,250)$. The adjusted pvalue obtained was used as indication of strength of association, however no formal statistical threshold was set.

### 4.3.2. Survival analysis

Survival analysis represents the collection of statistical methods that investigate the time to an event. For example, if the outcome was death, survival analysis would aim to investigate the time between origin (entry into the study) and time of death, by creating a survival function. The two main probability terms used to describe and model survival analysis are
survival and hazard (251). The survival probability represents the probability that an individual survives from origin until the specified time, while hazard is the probability that the individual under observation has an event at the specified time (251).

### 4.3.2.1. Censoring

Time zero is the point when the analysis time starts and is different depending on the study. For example, in an RCT time zero starts the day of randomisation while in a retrospective study of disease risk, time zero is the day the subject became considered "at risk" of that disease. The time ends upon the occurrence of an event with a well-defined criterion (e.g., death). Censoring occurs when the real survival time (i.e., the period from time zero to time at which the event occurs) of the subject is unknown. This can happen for various reasons and depending on the situation it carries different names (Table 4.1). Right-censoring is the most commonly occurring type of censoring (252).

Table 4.1: Censoring in survival analysis.

| Censoring type | Description | Examples |
| :---: | :---: | :---: |
| Right censoring | The event has not happened before the end of the follow-up period | - Subject did not experience the outcome before the end of study follow-up <br> - A patient is lost to follow-up before they experienced the event <br> - Patient experiences a different event that does not allow follow-up to be continued |
| Left censoring | The event occurred before the observation period started (time zero) | - The subject experienced the event of interest before the observation period started |
| Interval censoring | When patients come in and out of observation | A subject is follow-up at discrete time points and the time of the event happening can only be estimated between the two points between which the event occurred |

### 4.3.2.2. Kaplan Meier survival curves

Kaplan-Meier (KM) survival curves are generated by splitting the given length of time in multiple small time intervals and plotting the probability of surviving at each of the intervals $(253,254)$. There are three assumptions that the KM require:

1. Participants who are censored have the same survival as those who are still being followed up, at any given time.
2. The survival probability is identical regardless of the time of entry into the study.
3. The event occurs at the specified time.

Given the assumptions, the probability formula at any given time is:
$S_{t}=\frac{N_{\text {start }} N_{\text {outcome }}}{N_{\text {start }}}$
$N_{\text {start }}=$ Number of subjects outcome-free at the start
$N_{\text {outcome }}=$ Number of subjects with the outcome

Therefore, at each time interval, the survival probability is calculated by diving the number of subjects who did not experience the outcome during the time period by the total number of patients who were outcome-free at the beginning of the time period ("at risk"). Since the events are independent (i.e., one event occurring will not change the probability of another event occurring), the total probability of survival at any time is given by the multiplication of all probabilities preceding that time. KM curves are useful tools to assesses summary survival patterns (e.g., median survival) particularly when comparing different groups. A vertical gap between curves suggests that at a specific time, one group had a larger proportion of subjects not experiencing the outcome, while a horizontal gap suggests one group was slower to experience the same proportion of outcomes (253).

### 4.3.2.3. The log-rank test

While KM survival curves are useful tools to investigate overall survival patterns, they are unable to ascertain the statistical significance strength of the difference between multiple curves. The log-rank test evaluates whether differences observed between groups achieve statistical significance by testing the null hypothesis (i.e., that there is no difference between the two groups in terms of survival) $(253,255)$. The log-rank test uses the number of estimated events and compares it to the number of observed events and the value obtained is used to generate a p-value, using a $\chi 2$ distribution with $n-1$ degrees of freedom (251). The formulas for the calculations are presented below:

Log-rank test: $\chi_{2}=\sum_{i=1}^{n} \frac{\left(o_{i}-E_{i}\right)^{2}}{E_{i}}$, where
$\mathrm{n}=$ number of groups
$O i=$ number of observed events in group i
$E i=$ number of expected events in group i

### 4.3.2.4. Cox regression

KM survival curves and the log-rank test allow comparisons between survival time for two groups or more. However, they do not consider the effect that other variables could have on the survival function. The distribution of survival time accounting for covariates can be described by a fully parametric survival model. This distribution can also be described semiparametrically by relaxing parametric assumptions on time dependence. The Cox proportional hazards model is a semiparametric survival model in which the hazard ratio is constant over time. The mathematical formula for the Cox regression distribution is given by:
$\mathrm{h}(\mathrm{t})=h_{0}(t) \times\left(\beta_{1} x_{1}+\beta_{2} x_{2} \ldots \ldots \beta_{p} x_{p}\right) \quad$,where
$\mathrm{t}=$ survival time
$\mathrm{h}(\mathrm{t})=$ hazard function determined by covariates $x 1, x 2 \ldots x p$
$\beta=$ effect size of covariates
$h 0=$ baseline hazard

The proportional hazards assumption can be investigated via three methods.

## 1. Graphical approach $\log$ minus $\log$ plot

Since the survival function involves two exponentiations (one of the hazard ratio and one of the explanatory variable) a log minus log transformation will produce two curves. If the two curves are parallel or nearly parallel, then the proportional hazard assumption is not violated, since the distance between the two lines which is the difference in predictor values remains constant over time. As the assessment is subjective in nature, a conservative approach should be taken, and strong evidence of violation should only include instances where the curves cross or meet.

## 2. Goodness of fit test

The goodness of fit compares the observed and estimated values of the survival values. Schoenfeld residual is a type of goodness of fit test that is used to test the proportional hazards assumption and uses the difference between the observed and estimated explanatory variables for subjects who experienced an event (256). Schoenfeld residuals are independent of time because the hazard ratio is constant, thus if the graphical representations of the Schoenfeld residuals show a relationship with time, this would indicate a violation of the proportional hazards assumption. Since this method relies on statistical significance determined by p-value, artificially strong evidence can be detected in large sample sizes and potential violation can be dismissed as weak evidence $(256,257)$.

## 3. Applying a time-dependent covariate

Another method to test the violation of the proportional hazard assumption is to include an interaction term of time and the variable being investigated into the model. Using Wald or likelihood ratio statistics, both of which are methods of goodness of fit for statistical models. If the interaction terms show strong evidence at improving the model, then violation of the proportional hazards should be considered.

No one method described is better and as each comes with limitations; it is therefore advised that more than one test is performed to comprehensively investigate the violations of proportional hazards assumptions (257-259).

### 4.3.3. Predictive biomarker validation analysis

### 4.3.3.1 Confusion matrix

Confusion matrix is a table used in statistics when investigating classification performance(260). The confusion matrix counts the frequency of observed outcomes against predicted outcomes (Table 4.3) (260). Multiple measures can be derived from the confusion matrix and the measures used in this thesis are presented in Table 4.4.

Table 4.2: Confusion matrix table model for an outcome A

| Observed outcome A | Predicted outcome A |  |
| :--- | :--- | :--- |
|  | No | Yes |
| No | True negative (TN) | False positive (FP) |
| Yes | False negative (FN) | True positive (TP) |

Table 4.3: Confusion matrix measures

| Measure | Formula |  | Definition |
| :--- | :---: | :---: | :--- |
| Classification <br> accuracy |  | $\frac{\mathrm{TN}+\mathrm{TP}}{\mathrm{TN}+\mathrm{FP}+\mathrm{FN}+\mathrm{TP}}$ | The ratio of correctly classified <br> predictions (negatives and <br> positives) to all predictions made |
| Positive <br> predictive value <br> (PPV) | $\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}$ | The probability that a positive <br> prediction is truly positive |  |


| Negative <br> predictive value <br> (NPV) | $\frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FN}}$ | The probability that a negative <br> prediction is truly negative |
| :--- | :---: | :--- |
| Sensitivity | $\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}$ | The ratio of correct classified <br> positive predictions to total <br> observed positives |
| Specificity | $\frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}}$ | The ratio of correctly classified <br> negative predictions to total <br> observed negatives |

TN=True negative; $F P=$ False positive; $F N=$ False negative; $T P=$ True positive
Classification accuracy is a measure of how good the classifier is at making true predictions. PPV and NPV are useful measures that assess how likely it is that given a predicted positive or negative outcome this is truly positive or negative.

Sensitivity, or true positive fraction, refers to the ability of a test to detect an individual with the positive outcome as positive, while specificity, or true negative fraction, is the ability of the test to designate the individual with the negative outcome as negative. The false positive fraction (FPF) is 1 -specificity, and it represents the proportion of false positives (individuals detected through the test as positive but who do not have the outcome) from all individuals who do not have the outcome.

## Cohen's Kappa

Cohen's Kappa is a measure of agreement between two classifiers (261). In the context of the confusion matrix, the Kappa coefficient compares the agreement between the predicted and observed classifiers, by taking into account imbalances in the classifiers' distributions(262). Thus, the Kappa coefficient can provide additional information to overall accuracy when dealing with unbalanced data. The Kappa coefficient is calculated using the following formula:
$\kappa=\frac{\rho_{0}-\rho_{\mathrm{e}}}{1-\rho_{\mathrm{e}}}$, where
$\rho_{0}$ is the overall accuracy
$\rho_{\mathrm{e}}$ is a measure of how well that predicted values agree with the observed values by chance. In the confusion matrix $\rho_{\mathrm{e}}=\rho_{\mathrm{e} 1}+\rho_{\mathrm{e} 2}$, where $\rho_{\mathrm{e} 1}$ represents the probability that the predictors agree with observed values in the "No" category, while $\rho_{\mathrm{e} 2}$ represents the same but in the "Yes" category. The formula for calculating $\rho_{\mathrm{e} 1}$ is presented below:
$\rho_{\mathrm{e} 1}=\left(\frac{\mathrm{FP}+\mathrm{TP}}{\mathrm{TN}+\mathrm{FP}+\mathrm{FN}+\mathrm{TP}} \times \frac{\mathrm{TN}+\mathrm{FP}}{\mathrm{TN}+\mathrm{FP}+\mathrm{FN}+\mathrm{TP}}\right)+\left(\frac{\mathrm{FP}+\mathrm{TP}}{\mathrm{TN}+\mathrm{FP}+\mathrm{FN}+\mathrm{TP}} \times \frac{\mathrm{FN}+\mathrm{TP}}{\mathrm{TN}+\mathrm{FP}+\mathrm{FN}+\mathrm{TP}}\right)$

Kappa can take values from -1 to 1 , where -1 represents perfect disagreement, 0 represents no difference and 1 represents perfect agreement between the predicted and observed classifiers.

### 4.3.3.2 Receiver operating curves (ROC)

To evaluate the performance of predictive models, ROCs and area under the curve statistics were used. The ROC plots the sensitivity vs 1 -specificity and serves as test performance tool to assess how accurately a test can discriminate between those with and those without the outcome of interest (263). By overlapping the two distributions (sensitivity and 1specificity), the ROC allows the assessment of each of the two measures at a specific cut-off point. For example, if we set a cut-off at a specific sensitivity value, we can then calculate the FPF value at that point. If we changed the sensitivity cut-off value, then the FPF value would also change. The ROC curve is built by plotting the two distributions at multiple cut-offs values. In the hypothetical ROC curve from Figure 4.3, I present the discriminating power of test 1 and 2 to predict outcome $X$. The closer the curve is to the top left of the graph, the better the test is at discriminating the outcome. For example, in Figure 4.4, Test 1 has a better discriminating ability than Test 2 , since at any chosen sensitivity value, it will have a higher FPF. The orange line represents the chance level (i.e., test that generates a positive or negative result unrelated to true outcome status). The area under the receiver operating curve (AUROC) (or c-statistic) is a summary measure of the whole ROC curve and is computed by calculating the area under the ROC curve. The AUROC value represents the ability of the test to distinguish between the positive and negative individuals. The closer the AUROC is to 1 , the better the model is at distinguishing between the individuals with and without the outcome. An AUROC of 0.5 indicates that the model has no discriminating capability and will follow the chance line (orange line in Figure 4.3).

Figure 4.4: A receiver operating curve for Test 1 and 2 for outcome $X$.


### 4.3.3.3 Calibration curves

In addition to discrimination, which was discussed above, another key element of performance is calibration. Calibration represents the reliability of the predicted risk estimates in matching the observed proportions of the outcome (264). In a well calibrated model, for example in a study where PCa death is the outcome, this would mean that participants who had a predicted higher risk of PCa death will truly experience the outcome. In this thesis I chose calibration curves to evaluate the calibration for the models investigated, since a visual representation provides more information than numerical indicator (265). Calibration slopes describe the linear relationship between predicted risk of the outcome and the observed proportion outcome as the predicted risk increases(266). Figures 4.5 a and b show different calibration curves and their interpretation.

Figure 4.5: Theoretical calibration curves examples


Reproduced from Van Calster et al, 2019, with permission (266). The diagonal for each curve represents perfect calibration.
a. The dashed line represents underestimation of predicted risk, while the dotted line represents overestimation.
b. The dashed line represents overly extreme risk estimates. The line both dashed and dotted shows an underestimation and overly extreme risk estimates. The dotted line represents estimate risks that are not extreme enough

Calibration is very important in clinical practice, particularly if used in decision-making (266). Even if an algorithm has a good discrimination ability, it can still be poorly calibrated. In clinical practice, unreliable risk estimates (i.e., poorly calibrated) could give patients and healthcare professional inaccurate expectations, which could lead to misguided personal or treatment choice (266). Since calibration and discrimination are independent of each other and serve different purposes, when assessing the performance of an algorithm, both measures should be evaluated.

Many factors can influence the calibration of risk predictions. Disease incidence can affect calibration, if patterns are different in the population in which the algorithm was developed compared to those in which it is used. Other factors such as healthcare policy, treatment and screening guidelines may change the population over time which in turn can lead to changes in the calibration of algorithms linked to those change (266-270). There are also methodological causes for poor calibration such as statistical under and overfitting, measurement error and inadequate modelling strategy (266).

### 4.3.3.4 Predictive metabolomic risk models

### 4.3.3.4.1 The all-cause mortality metabolomic risk score model (all-cause mRS)

An all-cause mRS was developed in a study of 12 cohorts, in 44,168 participants, with age at baseline between 18 and 109. Over the cohorts' follow-up (mean follow-up range: 2.76-16.70 years) 5,512 participants died (271). The metabolomics profiling was done using the Nightingale NMR platform and the model used 63 of the 226 measures to avoid overfitting (271). Briefly, the score was generated by using Cox proportional hazards regression to identify the metabolites associated with all-cause mortality, and a forwards-backwards stepwise process based on successive rounds of meta-analyses (271). Of the 63 metabolites included, 14 reached the Bonferroni $p$-value threshold and were therefore included in the all-cause mRS.

### 4.3.3.4.2 The PCa mortality metabolomic risk score model (PCa mRS)

Collaborators at Nightingale Health developed a PCa mRS in UKBB using the platform. The score was derived in the general population, on metabolite plasma samples from 55,000
male participants from UKBB, with ages between 37 and 70 , using 37 clinically validated biomarkers from the Nightingale NMR panel (Nightingale Health@, Helsinki, Finland). The score aimed to predict PCa specific mortality in pre-diagnostic blood samples of participants in UKBB. The model was trained in half of the population ( $\mathrm{N}=24,895$ participants and 1,234 PCa deaths), using penalised logistic regression with L1/LASSO regularisation and five-fold cross validation to optimize the regularization parameter $\lambda$. The score performance was evaluated in the remaining half of the dataset using ROC curves and AUROC measures. The score performed well at predicting all-cause mortality in men in UKBB (AUROC=0.71). In addition, the score's performance was evaluated in FINRISK 1992 and 1997, which are population-based cohorts, followed-up for cancer diagnoses and mortality (all-cause mortality only), through a national registry system.

### 4.3.3.4.3 Applying the models' weights

In order to evaluate the performance of the models in my dataset, I applied the model weights for each of the metabolites included in the score and generated the overall score for each participant. This was done by following the same pre-processing steps used in the model development. Firstly, I dropped measurements that were outside four interquartile range in each dataset and scaled to standard deviation (SD) and log transformed all concentration values. I then multiplied the weights to my observed values for each metabolite and summed these together to obtain the overall risk score value for each participant.

### 4.4. Causality statistical methods

In this section I introduce intention to treat analysis, IV and Mendelian Randomisation (MR) analyses, which I used in Chapter 5 and 6. Their aim is to assess the causal link between metabolite levels and PCa mortality and causal effects of randomised interventions on intermediate biomarkers. In addition, this chapter provides a succinct introduction to MR, concepts and methodologies employed in this thesis.

### 4.4.1. Intention to treat analysis

Intention to treat analysis (ITT) is the most commonly used method of analysing data in a RCT. It includes the analysis of all participants who were randomised, in their respective
groups, regardless of adherence, actual received treatment or reasons for withdrawal $(272,273)$. The aim of the ITT is to assess if the different interventions prescribed to the participants produced an effect in the pre-determined outcomes. ITT can be performed by using statistical methods (e.g., linear regression for a continuous outcome) to compare differences in outcomes between the intervention groups (arms) and the control groups. The strengths and limitations of ITT are presented in Table 4.2.

Table 4.4: Strength and consideration for inference of ITT analyses

| Strengths | Considerations for inference |
| :--- | :--- |
| A comprehensive trial strategy included at |  |
| various stages of the trial (design, conduct |  |
| and analysis) | Leads to an attenuated treatment effect <br> amongst all compliers, as it analyses those <br> participants who did not receive treatment as <br> having received treatment |
| Maintains the prognostic balance resulting <br> from random allocation | Dilution due to noncompliance |
| Unbiased treatment effect | Introduces heterogeneity if noncompliant and |
| compliant subjects are analysed together |  |, | Difficult to interpret if significant crossover |
| :--- |
| randomised participants to the analysis stage | between treatment arms occurs | Limits bias due to ad hoc subgroups |
| :--- |
| Greatest generalisability |

Despite the considerations listed above, that impact on inference (e.g., the magnitude of the causal effect amongst compliers), ITT remains the recommended strategy for analysing RCT results by the Consolidated Standards of Reporting Trial (CONSORT).

### 4.4.2. Introduction to instrumental variable analysis

IV analysis is a statistical method used to make and estimate causal inferences, by controlling for unmeasured confounders. IV analysis uses a variable (or instrument) that is robustly correlated with the exposure of interest (e.g., treatment), but which has no direct effect on the outcome, to estimate the effect of the exposure on the outcome (274). There are three important assumptions in IV analysis which if not met could invalidate the analysis
and bias the results, namely that the instrument: i) is robustly associated with the exposure of interest (relevance assumption); ii) does not share common causes with the outcome (independence assumption) that influence the exposure-outcome relationship; and iii) only affects the outcome through the exposure of interest (exclusion restriction criteria) (Figure 4.6). IV analysis is often conducted in two stages. The first stage assesses how well the "instrument" predicts the exposure via the F-statistic. As the F-statistic increases, the risk of weak instrument bias decreases and causal effect estimation is more reliable (274-277). The second stage uses this information to determine the causal effect of the exposure on the outcome.

Figure 4.6: Diagram of the three IV assumptions

$1=$ relevance assumption; $2=$ independence assumption; $3=$ exclusion restriction criteria

### 4.4.3. Instrumental variable analysis in RCTs

ITT analysis estimates the causal effect of prescribing an intervention or treatment, rather than the causal effect of the exposures modified by the intervention. To establish the causal effect of the treatment dosage on the measured outcomes, instrumental variable (IV) analysis can be conducted (278). IV analysis, which is explained in more detailed in section 4.4 of the thesis, can be conducted in an RCT and along with the ITT results, can provide a more detailed interpretation of the results, particularly if there is interest in assessing the
causal role of the treatment (in those who adhere to it) and its dosage on the outcome. This is performed using 2-stage least squares regression. An F-statistic and $r 2$ for the first stage regression was computed to assess the IV assumption, that the instrument is sufficiently associated with the exposure. The F-statistic assesses the strength of the instrument, the higher the F-statistic, the lower the risk of weak instrument bias. Where the F-statistic suggested a high risk of weak instrument, the second stage of the regression was not performed. If the F-statistic was suitable, I considered the instrument to be at low risk of weak instrument bias and ran the second stage of the regression, which estimated the causal effect of the instrumented exposures (274-277).

### 4.4.4. Introduction to Mendelian Randomisation (MR)

### 4.4.4.1 Introduction

MR is a method that uses germline genetic variation to assess causality between modifiable risk factors and health outcomes (279). Details of MR have been published extensively over the last decade (280). Briefly, germline genetic variants, which are fixed at conception, are used to proxy modifiable risk factors, in an instrumental variable framework and used to establish causality on health outcomes of interest (280). At conception, germline variants that are associated with a particular trait are randomly distributed with respect to other unrelated traits and environmental factors, akin to random assignment of exposure to treatment or intervention in a conventional RCT (279-281). An experiment can be constructed in which people are naturally randomly allocated to a particular exposure based on the presence or absence of a genetic variant associated with the exposure of interest. Similar to a traditional RCT, and in its simplest form, MR based on one single nucleotide polymorphism (SNP) yields an "exposed group" of individuals who have the variant which can be compared with an "unexposed group" of individuals without that variant to determine if the outcome occurs more or less frequently in the exposed group compared to the unexposed one. By comparing the outcomes by groups, it can be established whether the relationship between exposure and outcome of interest is likely to be causal (282). DNA, although itself unmodifiable, operates through modifiable pathways. MR exploits this to identify modifiable exposures that can be used for disease prevention and therapeutic strategies, although a fundamental assumption is that of "gene-environment equivalence":
that the downstream physiological effects of modifying an exposure are the same whether they are genetically or non-genetically triggered.

### 4.4.4.2 Conducting Mendelian Randomisation

There are multiple types of MR studies. In one sample MR the genotypes, risk factors and outcomes all come from the same set of people. In a two-sample MR, two different sources of data are used to conduct the study, with data on genotyping and risk factors in one study and genotyping and health outcome data in another. There are important challenges in understanding MR studies, based on the type of study and the assumptions made $(279,283)$. The three key assumptions that are made about the instrumental variables (see section 4.4.1.) allow an MR study to be valid. In brief, instrumental variables should be associated with the risk factor of interest, they should be independent of the outcome, and they should only affect the outcome via the risk factor (Figure 4.4) (283-286).

Testing the assumptions:

1. Instrument strength (relevance strength)

Instrument strength is vital for a well conducted MR. A weak instrument will have a low statistical power for hypothesis testing, will amplify bias from the core instrumental variable assumptions and will bias the results towards the outcome in one-sample MR and towards the null in two-sample MR even with very large sample sizes (287). Weak instruments can be detected using the F-statistic, and a value of under 10 is indicative of a weak instrument (276).
2. Independence assumption

Rather than proving independence, the aim of this assumption testing is to dismiss dependence. By evaluating the effects of the genetic instrument on a wide range of characteristics, inferences on the plausibility of the independence (that the instrument is not influenced by confounders) can be made. This assumption can be investigated using negative control outcomes or populations, or measuring covariates that could confound the variant-outcome association $(288,289)$.

## 3. Exclusion restriction

The exclusion restriction (instrument only affects the outcome through the exposure) can be assessed through multiple methods. The most common sensitivity methods are the standard
inverse-variance weighted (IVW) methods, the median estimator, MR Egger, weightedmode based estimation and MR-PRESSO $(287,290,291)$. The Wald ratio is a statistical test which assess whether a set of parameters are equal to some value and is used in most of these methods. The Wald ratio estimate is the causal estimate obtained for each genetic variant and is equal to the gene-outcome association divided by its gene-exposure association. The IVW method can be used to estimate the causal effect of the exposure on the outcome, by combining the ratio estimates for each genetic variant. All genetic variants must be valid IVs and not be correlated (i.e., not in linkage disequilibrium) for the IVW method to produce an unbiased estimate. Even if only one genetic variant is invalid, the IVW will yield bias estimates. However, the median ratio estimator allows for up to $50 \%$ of the genetic instruments to be invalid and will produce a consistent estimate of the causal effect, although inefficiently. To improve this, a weighted median estimator was developed, where the weight assigned to each ratio is generated using empirical density functions of the ratio estimates. This method will yield consistent causal estimates as long as $50 \%$ of the weight will come from valid IVs (292). MR Egger is a statistical method that allows a causal estimate to be calculated when the IV assumptions do not hold, by setting an estimated intercept rather than zero as is the case in IVW.

Since many biological processes are yet to be fully understood due to their complexity, it is unlikely that a single genetic variant could satisfy the three assumptions, which could lead to using a weak instrument, unless the link between the risk factor and outcome is well understood (293). It is therefore important and ideal to use more than one genetic variant when conducting MR to robustly infer causal relationships. Genetic variants can affect the outcome of interest not just through the risk factor pathway, but other pathways which would invalidate them as instruments. This is referred to as horizontal genetic pleiotropy and is a particularly common issue in 'omics analysis. Vertical genetic pleiotropy occurs if the genetic variant affects the risk factor though a pathway which in turn affects the outcome. Unlike horizontal pleiotropy, vertical pleiotropy does not invalidate the instrumental variable and will not bias the findings. Comprehensive analysis methods have been developed which address the issue of pleiotropy in MR studies and have been well documented and are presented above $(290,292,294,295)$. Multiple statistical methods should be employed for causal inferences, particularly methods that make different assumptions, to provide a comprehensive assessment on whether the IV assumptions are satisfied (296).

### 4.4.4.3 Strengths and Weaknesses of Mendelian Randomisation studies

Unlike traditional observational studies, MR studies can produce unbiased, unconfounded estimates of the effect of modifiable risk factors on the health outcomes, by proxying exposures through genetic variants which are randomly distributed at birth. This is particularly important when dealing with confounding by common causes and reverse causation. Reverse causation can occur when pre-cancerous or early stages of cancer lead to metabolic changes that result in a positive association that is not causal, but potentially predictive. Such an effect could also result in inverse associations of metabolic changes with later or more advanced cancer stages, when undertaking an analysis that compares late vs early stages.

Provided the IV assumptions are met, MR studies are less prone to confounding by environmental factors or to reverse cause, and there are various analyses that can be undertaken to assess these issues. Furthermore, MR studies are less likely to be invalidated by selection bias than observational studies, since it has been shown that selection bias only introduce significant bias and Type 1 error when the selection effects are large (297). Unlike observational studies, MR studies are also not prone to diluted strength of associations due to "attenuation by measurement error" since the differences in exposure levels are across lifetime and thus unlikely to suffer from measurement imprecision (298).

MR studies can assess a wide range of exposures, some of which may not feasibly or ethically be tested in an RCT design which would ultimately provide the best evidence on causality. For example, assessing the link between smoking during pregnancy and offspring autism related disorders could not be ethically investigated in an RCT due to the potential harm to the participants and their offspring, however an MR study was conducted and did not find enough evidence to support a causal relationship between maternal smoking and offspring autism related disorders (299).

With the development of large genome wide consortia, and the two-sample MR, the causal impact of many risk factors on a large number of health outcomes can be investigated. MR studies can be quickly and easily run using freely available online platforms, such as MR Base (https://www.mrbase.org/), where both data and user-friendly interfaces allow this (300). MR can also provide evidence on causality between intermediate phenotypes and
disease, which could generate potential targets for interventions. Lifetime, hard to measure or dose dependant exposures, which are an issue in traditional epidemiological studies, can be addressed by the MR framework as well.

Population stratification, participant overlap, weak instrument bias, linkage disequilibrium and pleiotropy are common MR limitations; however, they can be investigated by measuring other factors as described above (279,283,293,301-303) (Appendix A, Table A5). In some cases of quantitative approaches in MR a very large sample size may be required for hypothesis testing, and this may not be achievable, thus leading to an underpowered study and biased results. Canalization and development stability refer to compensatory mechanisms that could develop during intra-uterine periods when exposure to certain factors could trigger life-long immunity to that factor via tissue structural and functional alteration (304). This concept of developmental compensation continues to affect MR studies as there are no current methods to address it.

### 4.4.5. Mendelian Randomisation techniques used in this thesis

### 4.4.5.1 Rationale for using Mendelian Randomisation

The evidence around risk and protective factors for PCa survival is still unclear. Although cholesterol has been suggested to play a crucial role in PCa and its progression, observational epidemiological studies have produced conflicting evidence and have been unable to establish any causality (305-308). An MR study which assessed the causal effect of lipid fractions (LDL, HDL, TG) on PCa, found some evidence that higher circulating LDL and TG levels increase and that a variant in the HMGCR gene, which mimics the lowering effect of statins on LDL, decreases the risk of developing aggressive PCa (308). Vitamin D is a compound that has been considered to affect PCa risk and progression for decades, with evidence from animal, cell and epidemiological studies, as well as clinical trials (309-312). The hypothesis generated from epidemiological studies, that Vitamin D, through Vitamin D receptors, decrease PCa progression, was investigated in a MR study. The authors found that a causal link between Vitamin D and PCa progression was unlikely, however the study only had one instrument for Vitamin D, putting it at high risk of weak instrument bias (313). Alcohol, a carcinogenic in many cancers, has an uncertain role when it comes to PCa risk
and progression, with meta-analyses highlighting the inconsistent evidence (314-319). An MR study performed in 23,868 PCa cases and 23,091 controls in the PRACTICAL dataset has found that genetically instrumented alcohol intake was associated with increased mortality in men with localised PCa, but not with PCa risk (320). Observational epidemiological studies have found that BMI and measures of adult stature may influence PCa risk, however few studies assessed their effects on PCa mortality $(321,322)$. Another MR study conducted in the PRACTICAL dataset found that genetically increased height and BMI levels were associated with increased mortality of PCa specific and all-cause mortality, respectively (323).

### 4.4.5.2 Investigating the Instrumental Variable assumptions

To assess the strength of the instrument (assumption 1) I conducted two-stage least square regression as described in section 4.3.4 to assess the strength of each metabolite instrument used in the MR analyses. The F-statistic and $r_{2}$ obtained from the first stage of the regression indicated the strength of the association between the genetic instrument and metabolite. A large F-statistic indicates that the risk of weak instrument bias is minimised. For the independence assumption, I investigated whether the genetic instruments used were reported to be associated with potential confounders. Lastly, I investigated the exclusion criteria through sensitivity analyses (MR Egger, simple mode, weighted median, weighted median). In addition, I checked if the genetic variants used to instrument the metabolites showed signs of horizontal pleiotropy, that is that they were associated with other metabolites which were likely on different pathways.

### 4.4.5.3 Collider bias

Collider bias occurs when conditioning on a variable that is a common effect of two variables. Collider bias is a particular problem in studies of disease progression since the factors that are shown to influence progression could be an artefact of the associations of the same factors with disease risk. For example, in a study of PCa progression, when only cases are selected, associations may artificially arise between the PCa risk factors (if at least one is associated with both risk and progression) leading to spurious associations with progression $(324,325)$. For unmeasured confounding between incidence and progression, any factor associated with incidence will also affect prognosis, with direction and size dependent on the incidence mechanism (326).Therefore, in an MR study of PCa progression within the cases, collider bias can occur since the independence assumption, that states the genetic
instrument has to be independent of the factors that confound the associations between exposure and PCa progression, could be violated $(283,297)$.

An example of how collider bias can invalidate results, is the conclusion reached in two studies that glucose-6phosphate dehydrogenase deficiency, which is associated with severe malaria, is protective over cerebral malaria $(327,328)$. A study that assesses collider bias in these studies found that the risk reduction in cerebral malaria could have been entirely attributed to collider bias (325). Only few methodologies that investigate the effects of collider bias in genetic studies have been proposed.

One proposed method uses the residuals from the regression of genetic effects on prognosis and the genetic effects on incidence (329). This method assumes that the genetic effects on incidence and outcome are independent, which is generally not the case in metabolomic research, since metabolites share many pathways (326). Another method that addresses collider bias subtracts the causal effect a trait on the outcome calculated using MR from the total genetic variant outcome estimate, however this method cannot be applied in case-only studies.

Thus, in this thesis I selected the slope-Hunter correction, a method that uses model-based clustering. The method allows the genetic effects on incidence and outcome to be correlated and estimates a correction factor using genetic variants that are only associated with incidence (326). This method was shown to eliminate or substantially minimise collider bias and performed better than the other two methods, particularly in case-only studies and in instances where there are correlated genetic effects on incidence and outcome (326). The method produces a corrected estimate of the causal association between the genetic variant and outcome by subtracting from the biased estimate the effect of the genetic variant that only affects the incidence adjusted using the correction factor. The formula for the corrected estimate using Hunter-slope method, adapted from Mahmoud et al, 2022 states:

$$
\hat{\beta}_{G P}=\hat{\beta}_{G P}^{\prime}-\hat{b}_{1} \hat{\beta}_{G I} .
$$

where,
$\mathrm{G}=$ genetic variants
$\mathrm{P}=$ outcome
I = incidence trait
$\hat{\beta}_{G P}=$ unbiased estimate of the association of the genetic variants and outcome
$\hat{\beta}_{G P}^{\prime}=$ biased estimate of the association of the genetic variants and outcome
$\hat{b}_{1}=$ estimated correction factor
$\hat{\beta}_{G I}=$ estimate of the association of the genetic variants and incidence trait I

### 4.4.6. Mendelian Randomisation in feasibility trials

### 4.4.6.1 Introduction

The following section contains excerpts from a methodological publication I wrote in conjunction with Rhona Beynon and is currently available in preprint format (330). All text included in this thesis originating from this methodological publication was solely written by myself. The idea originated from a previous publication of Rhona Beynon where the method described in this section was developed, while the manuscript I wrote in conjunction with Rhona Beynon focused on the methodological aspects that should be applied or taken into account when using the method.

Feasibility studies are small-scale studies, often with a randomised controlled trial (RCT) design, which aim to assess whether an intervention can be delivered, and a future trial
performed, in a particular setting (281). Feasibility trials do not set out, and are not typically powered, to evaluate the effects of the intervention on clinical outcomes, such as disease risk or progression. Furthermore, participants are not usually followed-up for enough time to assess long-term outcomes, and thus ITT investigating clinical outcomes is rarely used in feasibility trials (281). In addition, even though RCTs are the gold standard for assessing the efficacy or effectiveness of treatments or interventions, there are potential limitations and threats to causal inference (Table 4.5). For example, non-adherence in randomised trials (i.e., people failing to adhere to their assigned intervention) could result in attenuated estimates of the causal effect of perfect adherence with the intervention on the outcome in the ITT analysis (introduced in section 4.3.3). On the other hand, people who do not adhere cannot get any potential positive effect of the intervention. Thus, if there was a clear adverse effect in patients who do not benefit from treatment, then an ITT analysis would not capture the full extent of that adverse effect amongst the intervention group, resulting in an inflated estimate of the benefit: harm trade off caused by the intervention. As such, ITT may not reflect the potential causal effect of the intervention on the outcome, as it does not consider participants' behaviours and their effect on adherence and outcomes.

Table 4.5: Potential issues in feasibility RCTs

| Issue | Discussion | Further reading |
| :---: | :---: | :---: |
| INTERNAL VALIDITY |  |  |
|  | Inadequate blinding in an RCT, both for participants and trial staff, can introduce bias in |  |
| Inadequate blinding | the way the intervention is delivered, participants are followed-up and change participants' behaviours. | $(331,332)$ |
| Non-compliance in | Non-compliance refers to the failure of patients |  |
| the intervention | to follow the treatment or intervention assigned | $(331,333)$ |
| group. | by randomisation. |  |



> participants and the trial findings may not be generalisable. Low recruitment is particularly important in feasibility trials as the sample size is generally small.


#### Abstract

Although feasibility trials cannot test for any effects of the interventions on outcomes, they may give some indication of the effects of a novel intervention on short-term intermediate endpoints, such as minimally invasive biomarkers of a biological effect of the intervention which may lie on the causal pathway to clinical endpoints. These biomarkers may be established measures of clinical significance, for example Gleason score measurements in a study of PCa progression, or more novel biomarkers of cancer progression, such as measures of the epigenome, metabolome or proteome $(34,281)$. Such biomarkers can serve as supportive endpoints, which may then predict the clinical and long-term impact of a novel intervention. If such biomarkers are directly influenced by the intervention and in turn modify clinical outcomes, they may be used to both determine the potential magnitude of the effect of the intervention on long-term outcomes and to advance understanding of biological mechanisms which underlie intervention effects.

While MR can be applied in feasibility studies directly to estimate the causal effect of the exposure (intervention) on the clinical outcome of interest, this relies on the availability of genetic data and power, which is often inadequate in feasibility studies. Thus, an alternative approach would be an MR implemented using a two-sample approach based on summary genetic association data from a genome-wide association study (GWAS) of the exposure


 (sample 1) and a GWAS of the outcome (sample 2).This section presents a two-stage randomisation analysis (Figure 4.7) process for estimating the effects of novel interventions on long-term clinical outcomes using data from feasibility trials on supportive endpoints. The two-stage process involves estimating causal effects of an intervention on supportive endpoints from a feasibility RCT (stage 1), which is then coupled with estimates of the causal effect of the supportive endpoints on outcomes obtained from large-scale GWASs (stage 2). The process relies on knowledge and application of three statistical techniques: intention-to-treat analysis (339); instrumental variable analysis (278); and MR analysis $(280,284)$.

Figure 4.7: Two-stage randomisation in feasibility trials diagram


Stage 1: ITT, stage 2: MR

### 4.4.6.2 Method overview

In the first stage, ITT and IV analysis are conducted to establish the effects of the novel intervention on supportive endpoints. For a biomarker to be a valid surrogate endpoint i) it must be highly correlated with the long term clinical outcome that it is proxying; and ii) the effects of the intervention on it fully capture it's net effect on long-term clinical or public health outcomes $(340,341)$. However, in proof-of-concept trials, as is the case of feasibility RCTs, the supportive endpoint does not necessarily need to be a surrogate endpoint as it can still provide evidence around the biological mechanisms of action of the interventions being assessed. Those supportive endpoints altered by the intervention are then investigated further in the second stage of the process, an MR analysis, to predict the potential long-term effect of the intervention on clinically or public health relevant outcomes. This second stage requires the establishment of germline genetic variants (SNPs) that are robustly associated with the supportive endpoint of interest, and which can act as proxy instrumental variables for these endpoints. Such genetic instruments are usually identified from previously
conducted, large-scale and replicated GWAS. After establishing the SNPs that will be used as instruments, those SNPs are integrated with data from large GWAS consortia investigating the outcome of interest using the two-sample MR approach. This second stage assesses the causal effect of the intermediates on outcomes, such as disease risk or progression.

By considering the results from the two stages, it is possible to predict the potential impact of the novel intervention on long-term clinically or public health relevant outcomes via the estimated effect of the intervention on supportive endpoints. In addition, physiologically related adverse effects and potential off-target effects (including unsuspected adverse effects) of the intervention, that it was not specifically designed to produce, can be assessed via phenome-wide association studies (PheWAS) that test associations between the intervention as proxied by the instrumental variable and multiple phenotypes $(342,343)$. Using this two-step approach, the predicted effects of novel interventions on the outcomes of interest and on off-targets effects provides information that could potentially be used to justify the development and conduct of a large-scale, definitive phase III RCT that directly measures the clinical or public health outcome of interest.

### 4.5. Chapter summary

This Chapter presented the methods of the four studies of which data I used in the Thesis and their measured variables on which the analyses are based. In addition, I provided an overview of analytical and statistical methods for analysing metabolomics data in a cohort study (Chapter 5) and RCT setting (Chapter 6). I introduced MR as a tool to assess causality between metabolites and long-term clinical outcomes and its application in feasibility trials. Finally, I introduced the general concepts of prognostic replication analysis which have been used in Chapter 7.

## Chapter 5. Assessment of systemic metabolic biomarkers to detect

PCa progression in the Prostate Testing for Cancer and Treatment (ProtecT) trial

### 5.1. Chapter overview

In this chapter, I present: i) associations of baseline circulating metabolites in relation to Prostate Cancer (PCa) progression using survival analysis in the Prostate Testing for Cancer and Treatment (ProtecT) randomised controlled trial (RCT); and ii) estimates of causal associations between circulating metabolites and PCa mortality in the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL) consortium.

The Chapter opens with a succinct summary of the methodology and statistical analyses employed, with the main description of the methodology discussed in Chapter 4 . The results section is split into two parts. The first part presents the results of the survival analysis for the association of baseline circulating metabolites and PCa progression in the ProtecT trial. A broader measure of PCa disease progression (clinical progression, metastases or PCa death) and a more specific definition of progression which includes PCa metastases or PCaspecific death only are used. The second part presents the causal analysis, using Mendelian randomisation (MR), assessing the link between circulating metabolite levels and PCa progression (PCa-specific mortality only). The UK Biobank (UKBB) genome-wide association study (GWAS) was used to identify genetic instruments for metabolites and the PRACTICAL consortium to investigate how these genetically instrumented levels relate to PCa mortality. The chapter ends with a triangulation of evidence by bringing together the observational and causal evidence in the discussion and summary sections.

### 5.2. Research questions

1. Are circulating metabolite levels at baseline associated with subsequent risk of clinical progression in men with localised PCa? Clinical progression has been defined as evidence of metastases, diagnosis of clinical T3 or T4 disease, long-term androgen-deprivation therapy, ureteric obstruction, rectal fistula, the need for a urinary catheter owing to local tumour growth or PCa death.
2. Are circulating metabolite levels at baseline associated with subsequent risk of PCa metastases or PCa-specific mortality in men with localised disease?
3. Do circulating metabolite levels at baseline have a causal role in subsequent PCaspecific mortality?

### 5.3. Methods

In this section I briefly present the methods I used to analyse the ProtecT trial. A detailed description of the ProtecT trial, UKBB and PRACTICAL consortium and the methodology used in this Chapter can be found in Chapter 4 (23,35,237-240,344).

### 5.3.1. Study population

The main survival analysis used data from the ProtecT trial and included all men with localised disease for whom I had metabolomics data ( $\mathrm{n}=1,827$ ). Survival analyses using the complete case sample were also conducted; that is, men on whom there was complete anthropometric, lifestyle and morbidity questionnaires data ( $\mathrm{n}=1,232$ ). In addition, a sensitivity analysis was run on a subset of the ProtecT data, which was generated by censoring the dataset at 10 years of follow-up. This was done since there were few events after this timepoint and there was some evidence of the proportional hazards' assumption being violated.

For the two-sample MR analysis, which assessed the causal link between metabolites and PCa death, data from the UKBB GWAS $(\mathrm{n}=115,078)$ was used as the exposure dataset (sample 1) which identified genetic instruments for metabolites. The PRACTICAL consortium GWAS study of PCa mortality ( $\mathrm{n}=75,672$ ) was used as the outcome dataset to identify genetic instruments associated to PCa mortality (sample 2).

### 5.3.2. Measures

In this chapter, for the survival analyses I used circulating metabolites as my exposure variables and clinical progression and PCa metastases or PCa-specific death as outcome variables for the main (minimally adjusted) and complete case analysis. Due to small numbers of PCa-specific deaths, I grouped these with metastases and throughout this Chapter they are analysed as a combined outcome (metastases or PCa death). These progression categories follow the ProtecT trial definitions which are discussed in detail in Chapter 4. The definition of clinical progression in the trial includes metastases or PCa death, thus the clinical progression category in the current analyses includes the metastases or PCa death category.

In the main analysis I included the following potential confounders: Gleason score, tumour, nodes and metastasis (TNM) stage, baseline PSA, study centre and age. Gleason score, TNM stage and baseline PSA are the three well established marker of PCa aggressiveness and thus could be confounders of the metabolite PCa progression relationship. Given that the ProtecT trial is a multi-centre study, differences between centres may result from different geographies and social economic aspects of the areas and internal study specific procedures. This would introduce a heterogeneity in the data, which may confound the observed results. Age is another potential confounder given that it is associated with higher PCa mortality.

For the fully adjusted analysis (in the complete case analysis) in addition to the variables used in the minimally adjusted analysis I also included the following variables which could be related to the progression outcomes, given that they are known risk factors for all-cause mortality: BMI, self-reported alcohol consumption, smoking and morbidity measures (angina, coronary or myocardial heart attack, stroke, hypertension, high cholesterol, asthma, bronchitis, emphysema, diabetes, nervous trouble, or depression). Details on generating nuclear magnetic resonance (NMR) metabolomic data, as well as the processing of it are described in detail in Chapter 4. Briefly, 229 metabolites were generated using the Nightingale NMR platform (Nightingale Health@, Helsinki, Finland), which were then processed to address correlation, distribution skewness and analytical outliers $(149,244)$. PCa specific death was the outcome for the MR analysis, as defined in each of the PRACTICAL studies, generally obtained through mortality registries from various countries (23).

### 5.3.3. Statistical analyses

### 5.3.3.1. Descriptive analysis

Baseline characteristics were stratified by progression status at the end of follow up, (men who did not progress, those who clinically progressed, and those who experienced PCa metastases or PCa death). For continuous variables, means and standard deviations were computed, and for categorical variables percentages were used.

### 5.3.3.2. Missing data

There were no missing data for the variables included in the main analysis (minimally adjusted). For the variables used in the fully adjusted analysis, $33 \%$ of men had missing data. Whilst weight was collected by nurses at the ProtecT screening and enrolment visit, height, along with all other variables included in the fully adjusted model, were collected via questionnaires if men entered the additional ProMPT epidemiological study at the same visit. The pattern of the variables missing is not assumed to be at random given that these were participants who did not consent to enter the ProMPT study and return their first questionnaire. However, participants in the complete case analysis ( $67 \%$ of the ProtecT cohort) had similar age, PSA, weight and proportion of clinical progression and metastases or death compared to participants who had missing questionnaire data. Whilst not employed in this thesis, multiple imputation could have been used to attempt to fill values for the variables with missing data. Multiple imputation, in this context, would use the metabolite measures to compute the missing values for BMI. However, since the measures are highly correlated with each other and with BMI, using metabolite measures to compute BMI could lead to issues of collinearity and inflated variances, which would make the multiple imputation method unreliable and difficult and potentially impossible to implement (345). Therefore, since multiple imputation would not produce reliable estimates, and the two groups are believed to have similar demographic and clinical characteristics thus unlikely to introduce substantial bias, I excluded the observations with missing BMI values in the fully adjusted analysis.

### 5.3.3.3. Survival analysis

I used Kaplan-Meier (KM) survival curves to visualise the differences in survival patterns for the categorical values included as covariates in the multivariable survival analysis:

Gleason groups, BMI (normal (BMI $\leq 25 \mathrm{~kg} / \mathrm{m}^{2}$ ) vs overweight or obese (BMI $>25 \mathrm{~kg} / \mathrm{m}^{2}$ ), TNM stage (T1 vs T2, as T3 were ineligible to be included in the ProtecT trial) and age groups (under 60, 60 to 65 and over 65). I undertook Cox proportional hazards regression to examine the risk of developing PCa progression (clinical progression and PCa metastases or PCa death) per standard deviation increase of metabolite level, in the minimally adjusted model.

Sensitivity analyses were conducted by running the following Cox regression analyses: i) a minimally adjusted model in the complete case subset; ii) a fully adjusted model in the complete case subset. Evidence of violation of the proportional hazard assumption was investigated using the "estat phtest" command in STATA (346). The command tests whether Schoenfeld residuals are time independent. There were only few events after 10 years of follow-up, mainly because most participants did not have follow-up after 10 years. There was some evidence of nonproportional hazards after this timepoint (age groups and weight category survival curves overlap). To address this, an additional sensitivity analysis was conducted with the minimally adjusted model in the main dataset censored at 10 years of follow-up.

Given the large number of correlated metabolites which were investigated and the limited number of observations in the dataset, I employed principal component analysis to adjust the threshold for statistical significance. Description of the reasoning and methodology of this method has been provided in Chapter 4 section 4.3.1.2.4. There were 15 principal components that explained $95 \%$ of the variance and calculated an adjusted p-value of 0.003 $(0.05 / 15)$, with associations under the threshold treated as providing strong evidence. I also investigated potential suggestive evidence i.e. where associations did not meet the $p$-value threshold but were $\mathrm{p}<0.05$ and considered other measures, such as effect size and confidence intervals, to assess the strength and importance of the evidence (347).

### 5.3.3.4 Mendelian randomisation

Detailed descriptions of IV and MR analyses as well as the techniques employed in this Chapter are presented in Chapter 4, section 4.4. In summary, two-sample MR analyses were used to assess the causal link between metabolites and PCa-specific mortality, building genetic instruments for each metabolite (Sample 1) and generating causal estimates for the instruments on PCa-specific mortality (Sample 2). First, an adjusted effect size for the PCa mortality GWAS was calculated, using the Slope Hunter correction for collider bias (326). F-
statistic and $r_{2}$ measures were computed to assess the strength of association between the individual genetic instruments and each metabolite. The inverse variance weighted (IVW) method was used to estimate the causal effect if more than one genetic variant was available and the Wald estimate for instances where only one genetic variant was present. As a sensitivity analysis, where possible (i.e., if three or more instruments were available), the causal effect estimates were also computed using MR Egger, weighted median, weighted mode and simple mode methods. Where genetic variants were found to be associated with the majority of metabolite measures, the MR analysis was re-run after exclusion of such genetic variants. One of the strengths of this MR study is its ability to assess a wide range of metabolites, with genetic instruments available for 246 metabolites. Rather than using a multiple testing adjusted p-value, I assessed the strength of evidence by looking at the $95 \% \mathrm{CI}$, effect size and p-value together to minimise the risk of dismissing important findings which could be penalised by adopting a strict statistical threshold of significance (347-349).

### 5.3.3.5 Correlation analysis

To assess the consistency between the results from the observational and causal analyses, the logged odds ratio of developing metastases or PCa death for each metabolite in the analysis of the ProtecT trial was plotted against the logged odds ratio for PCa mortality in the MR study. To assess the consistency between the observational and causal estimates, a forest plot was constructed which presents the odds ratios in the ProtecT trial alongside the odds ratios in from the MR analysis for each metabolite. Logistic regression with metastases or PCa death as a binary outcome and individual metabolites as explanatory variables, with the covariates from the minimally adjusted Cox regression model (Gleason, age, centre, stage, baseline PSA) was used to calculate the odds ratio for metastases or PCa death in the ProtecT trial.

### 5.4. Results

### 5.4.1. Baseline descriptive analysis

The baseline characteristics for the ProtecT trial participants are presented in Table 5.1. Overall, 201 (11\%) men with PCa experienced clinical progression over the median 9.5 (8.0-
10.9 interquartile range) years of post-randomisation follow-up. At baseline, men who experienced clinical progression had a mean age of 63 years, $27.4 \%$ were overweight or obese and $4.0 \%$ had a Gleason score higher than 7 . Sixty-four men experienced metastases or PCa-specific death over the follow-up period. They had a mean baseline age of 63 years, $26.6 \%$ were overweight or obese and $4.7 \%$ had a Gleason score higher than 7 . There were no meaningful differences between men who progressed compared to men who did not progress (clinical progression) by age, BMI, alcohol intake or smoking status. Baseline PSA was lower in men who did not progress compared to men who progressed clinically (mean concentrations $5.56 \mathrm{ng} / \mathrm{ml}$ and $6.62 \mathrm{ng} / \mathrm{ml}$, respectively, mean difference $=1.06 \mathrm{ng} / \mathrm{ml}$, $95 \% \mathrm{CI}: 0.63-1.48, \mathrm{p}<0.001$ ). The proportion of men with a Gleason score higher than 7 was greater in men whose disease clinically progressed compared to men whose disease did not $(4.0 \%$ and $1.7 \%$, respectively, difference in percentage $=2.3 \%, 95 \% \mathrm{CI}: 0.4 \%-0.4 \%, \mathrm{p}=0.03)$.

Table 5.1: Baseline characteristics, by progression category in the The Prostate Testing for Cancer and Treatment trial 6

| Variable | No disease progression | Clinical progression | Metastases or PCa death |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | N | Mean (SD) | N | Mean (SD) | N | Mean (SD) |
| Age (years) | 1626 | $62.04(4.94)$ | 201 | $63.0(4.9)$ | 64 | $63.30(5.13)$ |
| Body mass index (kg/m2) | 1098 | $27.05(3.57)$ | 147 | $27.4(3.6)$ | 48 | $26.55(2.90)$ |
| Overweight and obese (\% BMI>25) | 790 | 72 | 77 | 77.8 | 33 | 68.75 |
| PSA baseline (ng/ml) | 1626 | $5.56(2.89)$ | 201 | $6.6(3.1)$ | 64 | 6.53 (2.95) |
| Gleason score <=7 (\%) | 1598 | 98.3 | 193 | 96.0 | 61 | 95.31 |
| Gleason score >=8 (\%) | 28 | 1.7 | 8 | 4.0 | 3 | 4.69 |
| Alcohol Intake | 1189 |  | 157 |  | 49 |  |
| Hazardous drinker (\%) | 50 | 4.21 | 8 | 5.1 | 3 | 6.12 |
| Moderate drinker (\%) | 1094 | 92.0 | 144 | 91.7 | 46 | 93.9 |
| Non-drinker (\%) | 45 | 3.8 | 5 | 3.2 | 0 | 0 |
| Smoking Status | 1193 |  | 158 |  | 50 |  |
| Current smoker (\%) | 135 | 11.3 | 23 | 14.6 | 8 | 16 |
| Former smoker (\%) | 592 | 49.6 | 79 | 50.0 | 28 | 56 |
| Never smoker (\%) | 466 | 39.1 | 56 | 35.4 | 14 | 28 |

PSA $=$ Prostate specific antigen

### 5.4.2. Survival analysis

### 5.4.2.1. Associations of potential confounders with Prostate Cancer progression

The Kaplan Meier (KM) curves for clinical progression and metastases or death by Gleason, BMI, stage, and age groups are presented in Figures 5.1-5.8.

### 5.4.2.1.1. Clinical Progression

Of the $201(11 \%)$ instances of clinical progression, 193 occurred in men with a Gleason lower or equal to 7 , and 8 in men with a Gleason score higher than 7 (Table 5.1). The KM curve shows that the probability of developing clinical progression was consistently higher for men with a Gleason score higher than 7 throughout the follow-up period, compared to men with a Gleason score lower or equal to 7 (log rank test $\mathrm{p}=0.01$ ) (Figure 5.1a). There were 147 clinical progression events for which BMI was measured, 77 ( $52.4 \%$ ) of which occurred in the overweight group (BMI>25kg/m²). Overall, the probability of developing clinical progression was similar between the normal weight and overweight groups (log rank test $\mathrm{p}=0.4$ ) (Figure 5.1b). Of the total 201 clinical progression events, 125 were to men with a T1 stage and 76 to men with a T2 stage (Figure 5.1c). There was evidence that men with stage T2 had a higher probability of developing clinical progression overall compared to men with stage T1 (log rank test p <0.001) (Figure 5.1c). There were 60, 61 and 80 clinical progression events in men under the age of 60,60 to 65 and 65+ years, respectively. The KM curves showed similar probabilities of developing clinical progression by age groups overall (log rank test $\mathrm{p}=0.09$ ) (Figure 5.1d).

Figure 5.1: Kaplan Meier plot of Prostate Cancer clinical progression by categorical confounding variables in The Prostate Testing for Cancer and Treatment trial (Gleason score grouping, weight category, Tumour, Node andMetastasis stage, age group)


Each of the lines in the graphs represent the probability of not experiencing clinical progression in specific groups of participants:
(A) Gleason $\leq 7$ and Gleason $>7$
(B) Normal weight $(\mathrm{BMI} \leq 25)$ and overweight $>25$
(C) TNM stage T1 and T2
(D) Age $<60,60-65$ and $\geq 65$

### 5.4.2.1.2. Metastases or PCa death

Of the $64(3.5 \%)$ metastases or PCa death events, only 3 were in men with a Gleason score higher than 7. The KM curve shows some, but not strong evidence of a gradual increase in the probability of developing metastases or PCa death in participants with a Gleason score lower or equal to 7 (log rank test $\mathrm{p}=0.14$ ) (Figure 5.2a). However, only three events were observed in the group of men with a Gleason score higher than 7, therefore there was little statistical power to detect a difference in survival between the two groups. There were 15 normal weight and 33 overweight ( $\mathrm{BMI}>25 \mathrm{~kg} / \mathrm{m}^{2}$ ) participants who developed metastases
or PCa death. The KM curves presented in Figure 5.2b show a similar gradual increasing probability of developing metastases or PCa death in both groups throughout the study $(p=0.59)$. Of the 64 men who developed metastases or PCa death, 36 and 18 had stage T1 and T2, respectively. The KM curve shows both groups had an increasing probability of developing metastases or PCa death ( $\mathrm{p}<0.001$ ) (Figure 5.2c). There were 20 men who experienced metastases and PCa death in the under 60-year-old category, 14 in the 60-to-65year group and 30 in the over 65 s . The probability of developing metastases or PCa death increased in all three groups throughout the follow-up period ( $\mathrm{p}=0.04$ ) (Figure 5.2.d).

Figure 5.2: Kaplan Meier plots of Prostate Cancer metastases and death by categorical confounding variables in the The Prostate Testing for Cancer and Treatment trial (Gleason score grouping, weight category, Tumour, Node and Metastasis stage, age group)


The lines in the graphs represent the probability of not experiencing metastases or PCa death in specific groups of participants:
(A) Gleason $\leq 7$ and Gleason $>7$
(B) Normal weight $(\mathrm{BMI} \leq 25)$ and overweight $>25$
(C) TNM stage T1 and T2
(D) Age $<60,60-65$ and $\geq 65$

### 5.4.2.2 Associations of metabolic traits with Prostate Cancer progression and metastases or Prostate Cancer death in Cox regression

The full list of results for each metabolite in relation to PCa clinical progression and metastases or PCa death are available in Tables B1 and B2, Appendix B. For simplicity, Figure 5.3 presents all metabolites with the exception of lipoprotein subclass measures for which only the total lipid value is shown. Overall, there was limited evidence of associations between individual metabolites and clinical progression, metastases or death, with none of the associations surpassing my pre-defined statistical threshold ( $\mathrm{p}<0.003$ ). There was suggestive evidence of associations of lipoproteins, cholesterol, fatty acid ratios, amino acids and measures of glycolysis and fluid balance and disease progression (HR range from protective, OR as low as 0.78 , to positive ORs up to 1.31). Generally, patterns observed for clinical progression followed those for metastases or PCa mortality (Figure 5.3). The effect sizes were larger in magnitude but estimated with wider confidence intervals in the metastases or PCa death analysis compared to the clinical progression category, probably because the numbers of these outcomes were fewer (Figure 5.3 and Table B1 Appendix B). In sensitivity analyses, the associations observed in the fully adjusted model generally followed the same patterns as those in the minimally adjusted model (Table B1, B2 and Figures B1, B2, B3, Appendix B). The direction and magnitude of associations were largely maintained after censoring the data at 10 years of follow-up (Table B1, B2 and Figure B4, B5 and B6, Appendix B).

Figure 5.3: Forest plot of hazard ratios for clinical progression and metastases or death, in the minimally adjusted Cox models.


VLDL=Very low-density lipoprotein; $L D L=$ low-density lipoprotein; $H D L=$ high-density lipoprotein; $C=$ Cholesterol; MUFA=monounsaturated fatty acids; PUFA=polyunsaturated fatty acids; n3=omega; n6=omega 6

### 5.4.2.2.1 Associations of progression with lipid and lipid related measures

Clinical progression
There was some evidence suggesting that clinical progression was positively associated with circulating measures of very low-density lipoprotein (VLDL) and triglycerides and inversely associated with non-triglyceride measures of high-density lipoprotein (HDL) (Figure 5.3 and Table B1, Appendix B). The largest effect-sizes were for triglycerides in medium HDL (not shown on the graph $)(\mathrm{HR}$ per standard deviation increase $=1.14,95 \% \mathrm{CI}: 1.00-1.30 ; \mathrm{p}=0.05)$ and HDL3 cholesterol ( $\mathrm{HR}=0.86,95 \% \mathrm{CI}: 0.75-0.99, \mathrm{p}=0.04$ ). The associations observed in the minimally adjusted model were generally maintained in the fully adjusted model, with comparable effect sizes and precision (Table B1 and Figure B2, Appendix B).

## Metastases or PCa-specific death (cumulative measure)

There was some evidence that metastases or PCa-specific mortality measures were associated with medium, very large and extremely large VLDL, some triglyceride measures and particle diameter for low-density lipoprotein (LDL) (Figure 5.3 and Table B2, Appendix B). The largest effect-size was for concentration of chylomicrons and extremely large VLDL particles (not shown on graph) (HR=1.26; 95\% CI: 1.04-1.53; p =0.02) (Table B1, Appendix B). There was some evidence for inverse associations of metastases and PCa specific mortality with small HDL measures, small, medium, and large LDL, and multiple intermediate density lipoprotein measures (Figure 5.3; Table B2, Appendix B). The largest effect-size was for cholesterol esters in small HDL (HR=0.78, 95\% CI: 0.61-0.99; p =0.04) (Figure 5.3; Table B2, Appendix 5). The associations observed for metastases and death were maintained in the fully adjusted model, with comparable effect sizes and precision to those in the minimally adjusted model, with the exception of LDL particle size which showed a stronger effect size in the adjusted model (minimally adjusted model: $\mathrm{HR}=1.26,95 \% \mathrm{CI}: 0.98$ $1.55, \mathrm{p}=$ vs. fully adjusted model: $\mathrm{HR}=1.39,95 \% \mathrm{CI}: 1.07-1.82, \mathrm{p}=0.01$ ) (Table B2 and Figure B3, Appendix B).

### 5.4.2.2.2 Associations of progression with fatty acid measures

There was limited evidence for an association of clinical progression, metastases or death with fatty acid measures in the minimally adjusted model, with small effect sizes and confidence intervals crossing the null (Figure 5.3, Tables B1 and B2, Appendix B). The largest effect sizes for clinical progression were observed for the ratio of omega-6 fatty acids (HR= $0.90 ; 95 \%$ CI: $0.79-1.04, \mathrm{p}=0.13$ ) and the ratio of monounsaturated fatty acids to total fatty acids ( $\mathrm{HR}=1.11,95 \% \mathrm{CI}: 0.97,1.27, \mathrm{p}=0.12$ ). For metastases and PCa death, the largest effect sizes were for the ratio of saturated fatty acid to total fatty acid (HR=1.20; 95\% CI: 0.93-1.55, $\mathrm{p}=0.16$ ) and degree of unsaturation ( $\mathrm{HR}=0.82,95 \% \mathrm{CI}: 0.64-1.06, \mathrm{p}=0.13$ ). Similar patterns were observed in the fully adjusted model for clinical progression (Table B1 and Figure B1, Appendix B). The associations for metastases or death were stronger in the fully adjusted model, with larger magnitude effect sizes but wider confidence intervals (Table B2 and Figure B3, Appendix B). The association for the degree of unsaturation had the largest difference in effect size in the fully adjusted compared to the minimally adjusted model and was stronger but estimated with similar precision (HR=0.77; 95\%CI: 0.57-1.05, p=0.10) (Table B2 and Figure B3, Appendix B).

### 5.4.2.2.3 Associations of progression with glycolysis related measures

There was limited evidence that clinical progression was associated with measures of glycolysis. In the minimally adjusted model, there was some evidence that citrate was inversely associated with clinical progression ( $\mathrm{HR}=0.87,95 \% \mathrm{CI}: 0.75-1.00, \mathrm{p}=0.05$ ) (Figure 5.3 and Table B1, Appendix B). The association was maintained in the fully adjusted model, with similar effect sizes and precision (Table B1 and Figure B2 Appendix B). There was weak evidence of association for metastases or PCa death with glycolysis related measures, with direction and effect sizes of associations similar to those in the clinical progression category but with wider confidence intervals (Figure 5.3 and Table B2, Appendix 5).

### 5.4.2.2.4 Associations of progression with amino acids and ketones measures

There was limited evidence of associations between clinical progression, metastases or PCa death and amino acids and ketones in the minimally adjusted model, with most hazard ratios centred around the null (Figure 5.3 and Tables B1 and B2, Appendix 5). The largest
effect size for clinical progression was observed for glutamine which was inversely associated with clinical progression (HR=0.90; 95\% CI: 0.78-1.04, $\mathrm{p}=0.15$ (Table B1, Appendix B). The effect size and precision were similar in the fully adjusted model compared to the minimally adjusted one (Table B1 and Figure B2, Appendix 5). Leucine had the largest effect size for metastases and PCa death in the minimally adjusted model ( $\mathrm{HR}=1.19 ; 95 \% \mathrm{CI}: 0.94-$ $1.50 ; \mathrm{p}=0.14$ ), with the association becoming stronger but estimated with less precision in the fully adjusted model (HR=1.29; 95\%CI:0.98-1.70; p=0.07) (Table B2 and Figure B3, Appendix B).

### 5.4.2.2.5 Associations of progression with inflammation and fluid balance measures

There was some evidence that clinical progression and metastases or PCa death were associated with some measures of fluid balance and inflammation (Figure 5.3, Tables B1 and B2, Appendix B). Although not strong, there was some evidence that glycoprotein acetyls were associated with clinical progression ( $\mathrm{HR}=1.18 ; 95 \% \mathrm{CI}: 1.03-1.35 ; \mathrm{p}=0.01$ ), and metastases or PCa death ( $\mathrm{HR}=1.31, \mathrm{CI}: 1.04-1.64, \mathrm{p}=0.02$ ) in the minimally adjusted model, with the effect sizes being maintained in the fully adjusted model (Tables B1 and B2 and Figures B1, B2 and B3, Appendix B).

### 5.4.2.3. Censored sensitivity analysis

To address the lack of follow-up for most participants and the few events observed after 10 years of follow-up, a sensitivity analysis was run which censored the data at 10 years of follow-up for clinical progression and metastases or PCa death. The results are presented in Tables B1 and B2 and Figures B5 and B6 in Appendix B. Generally, the associations between metabolite levels and clinical progression and metastases or PCa death were consistent between the non-censored and censored analysis, with similar effect sizes and precision. The exception was LDL particle size which showed a larger effect size but similar precision in the metastases or PCa death category (uncensored minimally adjusted $\mathrm{HR}=1.23 ; 95 \% \mathrm{CI}: 0.98$ 1.55; $\mathrm{p}=0.08$ and censored minimally adjusted $\mathrm{HR}=1.31$; 95\%CI:1.03-1.67; $\mathrm{p}=0.03$ ) (Figure B6, Appendix B).

### 5.4.2.4 Cox Proportional hazard assumption

The results for the test of the Cox proportional hazard assumption are presented in Table B3, Appendix B. There was no strong evidence to suggest that the proportional hazard assumption did not hold (global $\mathrm{p}>0.003$ ). There was weak evidence to suggest that the proportional hazard assumption was violated for some HDL, VLDL and fatty acid measures in the clinical progression category (global p<0.05).

### 5.4.3. Causal analysis using Mendelian randomisation

The F-statistic and $\mathrm{r}^{2}$ for the genetic instruments of the metabolites in UKBB GWAS studies are presented in Table B4 in Appendix B. Generally, all instruments showed a strong association with the metabolites that were taken into the second step, with the lowest Fstatistic of 49 , and $\mathrm{r}^{2}$ between 0.002 and 0.15 . The results of the two-sample MR are presented in Figure 5.4 and Table B5 in Appendix B.

Overall, I found evidence to support causal roles of cholesterol related molecules, lipoproteins, fatty acids and amino acids on PCa mortality. Total, free and esterified cholesterol, some measures of IDL, LDL and VLDL, sphingomyelins, apolipoprotein B, total omega-3 fatty acids, docosahexaenoic acid, and valine were found to causally increase PCa mortality risk. Weaker evidence found that beta-hydroxybutyrate alanine, the concentration of branched-chain amino acids (leucine, isoleucine and valine) and lactate may also be causally linked to increased PCa mortality. I also found evidence of causal association of histidine, phospholipids to total lipids ratio in small and large HDL cholesterol, ratio of triglyceride to total lipids in very small VLDL and ratio of omega- 6 to omega- 3 with decreased PCa mortality.

The consistency between the PCa mortality associations in the ProtecT study and the causal estimates from the MR analysis are shown in Figures 5.5 and 5.6 and Supplementary Table B6. Figure 5.5 and presents a scatter plot of the logged odds ratios in the ProtecT trial and MR study, showing an $\mathrm{r}^{2}$ value $<0.001$, suggesting inconsistent estimates between the observational and causal odds ratios. However, when comparing the odds ratios between each individual metabolite in the observational (ProtecT) and MR (PRACTICAL) only LDL, HDL and total measures of cholesterol look inconsistent (Figure 5.6). All other metabolites have small effect sizes (OR close to 1) and are estimated with large confidence intervals. In Figure 5.6 there are a number of outliers, with larger effect sizes which will contribute to
shifting the overall trend and could explain the slope nearing the zero value. Sensitivity analyses for the MR analysis conducted using MR Egger, simple mode, weighted median, weighted mode methods are presented in Table B5 in Appendix B and the results are consistent with those using the inverse variance weighted method.

### 5.4.3.1. Lipid and lipid related measures

There was evidence that total sphingomyelins, apolipoprotein B, cholesterol, total esterified cholesterol, and measures of IDL, LDL and VLDL increased subsequent PCa mortality, except the ratio of triglycerides to total lipids in very small VLDL which decreased subsequent PCa mortality. I found that HDL related cholesterol measures decreased subsequent PCa mortality, except free cholesterol measures in HDL which increased PCa mortality (Figure 5.4 and Table B5, Appendix 5). The largest effect size was observed for the ratio of cholesterol esters to total lipids in large LDL which were associated with higher PCa mortality risk (OR=1.32; 95\%CI:1.07-1.62; p=0.008) (Figure 5.4 and Table B5, Appendix 5). The sensitivity analyses found that the associations were generally maintained, with similar direction effect sizes (Appendix B, Table B5).

### 5.4.3.2. Fatty Acid measures

There was some evidence that total fatty acids and docosahexaenoic were causally associated with increased PCa mortality and the ratio of omega-6 to omega-3 fatty acids with decreased PCa mortality. The largest effect size was seen for the ratio of linoleic acid to total fatty acid which was associated with decreased PCa mortality ( $\mathrm{OR}=0.85 ; 95 \% \mathrm{CI}: 0.68$ 1.05; $p=0.13$ ) (Figure 5.4 and Table B5, Appendix B). The estimates had similar direction and effect sizes in the sensitivity analyses (Table B5, Appendix B).

### 5.4.3.3. Glycolysis related measures

There was weak evidence to suggest that glycolysis measures were causally associated with PCa mortality (Figure 5.4 and Table B5, Appendix B). Lactate showed one the largest effect sizes in my MR analysis (OR=1.49; 95\%CI:0.95-2.34; $\mathrm{p}=0.09$ ), however estimated with
imprecision. Sensitivity analyses found that the effect size fluctuated from 1.38 ( $95 \% \mathrm{CI}: 0.15-$ 12.9; $\mathrm{p}=0.8$ ) to 1.63 ( $95 \% \mathrm{CI}: 0.80-1.51, \mathrm{p}=0.24$ ), with wide confidence intervals.

### 5.4.3.4. Amino acids and ketones measures

There was evidence that alanine, valine and the concentration of total branched-chain amino acids (leucine, isoleucine and valine) were causally associated with increased PCa mortality, (Figure 5.4 and Table B5, Appendix B). Histidine showed some causal link with decreased PCa mortality. Sensitivity analyses found similar direction and size of the effects for these metabolites; however, the effects were estimated with less precision and wider confidence intervals (Table B5, Appendix B). Beta-hydroxy butyrate showed one of the largest effect sizes in this MR analysis, however the estimate was estimated with imprecision (OR:1.43; $95 \% \mathrm{CI}: 0.89-2.30 ; p=0.15)$. Sensitivity analyses found similar effect sizes but estimated with larger confidence intervals (Table B5, Appendix B).

### 5.4.3.5. Inflammation and fluid balance measures

There was weak evidence that inflammation and fluid balance measures were linked to PCa mortality, with small effect sizes estimated with imprecision (Figure 5.4 and Table B5, Appendix B). Albumin showed the largest effect size of association with decreased PCa mortality (OR= $0.89 ; 95 \% \mathrm{CI}: 0.73-1.08 ; \mathrm{p}=0.25$ ).

Figure 5.4: Forest plot of MR estimates (odds ratios) using IVW or Wald ratio and their $95 \%$ CI for the effects of metabolites on PCa mortality among 67,758 PCa cases in the PRACTICAL consortium, using instruments developed in the UKBB GWAS.

| Lactate |  | 늘 |
| :---: | :---: | :---: |
| 3-Hydroxybutyrate |  |  |
| Acetone |  | + |
| Isoleucine |  | - |
| Total concentration of branched-chain amino acids (leucine + isoleu |  | 불 |
| Valine |  | 블 |
| Cholesteryl esters to total lipids ratio in large LDL |  | 뭄 |
| Cholesterol to total lipids ratio in large LDL |  | - |
| Leucine |  | - |
| Cholesterol in large LDL |  | - |
| Alanine |  | 들 |
| Cholesteryl esters in large LDL |  | - |
| Total lipids in large LDL |  | -물 |
| Total lipids in LDL |  | - |
| Cholesteryl esters in LDL |  | -물 |
| LDL cholesterol |  | - |
| Cholesterol to total lipids ratio in medium LDL |  | - |
| Free cholesterol to total lipids ratio in very small VLDL |  | -들 |
| Clinical LDL cholesterol |  | - |
| Free cholesterol in LDL |  | - |
| Total cholesterol minus HDL-C |  | - |
| Phospholipids in large LDL |  | - |
| Free cholesterol in large LDL |  | -1- |
| Free cholesterol in medium LDL |  | - |
| Free cholesterol in medium VLDL |  | - |
| Free cholesterol in small LDL |  | - |
| Phospholipids in LDL |  | -1 |
| Concentration of medium LDL particles |  | - |
| Phospholipids in medium LDL |  | - |
| Phospholipids in small LDL |  | - |
| Concentration of large LDL particles |  | - |
|  |  | 2 |

Odds ratios per SD (95\%CI)






$V L D L=$ Very low-density lipoprotein; $L D L=$ low-density lipoprotein; $H D L=$ high-density lipoprotein; $I D L=$ intermediate-density lipoprotein

Figure 5.5: Scatter plot of the logged odds ratios of developing metastases or Prostate Cancer death in the observational (Prostate Testing for Cancer and Treatment) and logged odds ratios for Prostate Cancer mortality in the MR (Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome) studies.

## Scatterplot of the logged odds ratios in the observational (ProtecT) and MR(PRACTICAL) studies



MR=Mendelian randomisation; ProtecT =Prostate Testing for Cancer and Treatment; PRACTICAL=Prostate Cancer Association
Group to Investigate Cancer Associated Alterations in the Genome

The x -axis represents the odds ratios for PCa specific death for each metabolite in the MR study using instrument genetic data from UKBB GWAS studies and outcome genetic data form the PRACTICAL consortium. The y-axis represents the odds ratios of developing metastases or PCa specific death in relation to each metabolite in the ProtecT trial. Only metabolites that have odds ratios for both the ProtecT and UKBB MR studies are presented. The red line is the slope.

Figure 5.6: Forest plot of odds ratio for metastases or death in the Prostate Testing for Cancer and Treatment trial and odds ratio of Prostate Cancer mortality in the Mendelian Randomisation analysis in the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome consortium.


[^2]
### 5.5. Discussion

### 5.5.1. Main findings

In this chapter I assessed the potential link between metabolites and PCa progression using observational and causal evidence. I found suggestive observational evidence (ProtecT results) that individual metabolites were associated with PCa clinical progression, metastases or PCa specific death, with lipoproteins, cholesterol and glycolysis related metabolites, fluid balance and inflammation measures showing the strongest evidence. I found causal evidence (PRACTICAL MR results) that cholesterol related metabolites, lipoproteins, amino acid measures and fatty acids may be linked to PCa mortality. Total, free and esterified cholesterol, some measures of IDL, LDL and VLDL, sphingomyelins, apolipoprotein B, total fatty acids, docosahexaenoic acid and valine increased PCa mortality. Some HDL-related measures, ratio of triglyceride to total lipids in very small VLDL, histidine and the ratio of omega-6 to omega-3 fatty acids were causally associated with decreased PCa mortality. To a lesser extent since the evidence was not strong, lactate, alanine, the total concentration of branched-chain amino acids and beta-hydroxybutyrate were causally linked to increased PCa mortality. Although the metabolites identified as potentially affecting PCa mortality in the causal analysis and clinical progression or metastases or PCa mortality in the observational analysis differ or are discordant, both studies support a role for cholesterol and lipoproteins in PCa progression to metastasis and death. In the causal analysis, the instruments used to proxy metabolite levels estimated a lifetime effect whilst in the observational analysis the effect of metabolites on PCa progression was assessed over a period of 10 years after a PCa diagnosis. This could explain the lack of correlation
observed between the effect estimates from the two analyses. In addition, the less precise estimates in the ProtecT trial could be a consequence of the much lower number of progression events observed in the ProtecT trial than in the PRACTICAL consortium, which again explains the lack of correlation observed between the observational and causal estimates. Also, the disease pathologies were different between the two studies, with the ProtecT trial only including participants with localised cases (T1 and T2) while the PRCATICAL consortium included all stages of disease. Lastly, it has been previously found in the ProtecT trial that metabolites, particularly lipoprotein and cholesterol measures were associated with PCa risk (350). This suggests that the metabolomic analysis in the ProtecT trial could suffer from collider bias, since if metabolites were associated with risk of developing PCa, they could lead to spurious associations with PCa progression. Outcome ascertainment for PCa death was performed by a committee who evaluated clinical notes along the mortality records in the ProtecT trial, while for studies included in the PRACTICAL consortium these were generally done via mortality records alone $(323,351)$. In addition, the data in the PRACTICAL consortium comes from many countries, and thus the progression indicator could be influenced by other factors not present in the ProtecT trial, as a result of each country's healthcare and surveillance systems.

My findings highlight the important role that metabolites may play in PCa progression, particularly through cholesterol, fatty acid and amino acid pathways. They also present opportunities for identifying novel causal biomarkers which could become intervention targets for delaying PCa progression.

### 5.5.1.1. Lipids and lipid related measures

While the observational analysis did not find any strong associations (p>0.003) of PCa progression with measures of cholesterol in the main analysis, the measures with the highest effect sizes (e.g., phospholipids in LDL, free cholesterol in medium LDL) appeared stronger after adjusting for BMI, morbidity, and lifestyle factors. This was not
surprising given that increased adiposity has been previously shown to alter metabolic profiles (243). I found that multiple measures of cholesterol, such as VLDL (small, medium, extremely large), LDL (overall measures, small, medium, and large), IDL (overall measures), total, free and esterified cholesterol, sphingomyelins and apolipoprotein B showed evidence of being causally associated ( $\mathrm{p}<0.05$ ) with increased PCa mortality. Generally, the metabolites with the largest effect size in the causal analysis showed inconsistency in direction and effect size in the observational analysis for metastases and PCa death. For example, some lipoprotein measures such as LDL cholesterol, total cholesterol and esterified cholesterol which showed large effect sizes on increasing PCa mortality in the causal analysis, showed inverse associations in the observational analysis. However, the estimates for these metabolites were measured with imprecision in the observational analysis, with large confidence intervals, thus making the inconsistencies hard to interpret.

A previous study of metabolites and PCa risk in the ProtecT study found some evidence that these metabolites (HDL and LDL measures, total free and esterified cholesterol) are associated with PCa risk. This suggests that collider may be an issue in my observational analysis, since the metabolites were found to decrease the risk of developing PCa progression. This could explain the opposite direction of association observed in the ProtecT trial compared to the MR analysis. On the other hand, the genetic variants used for instrumenting the effects of the strongest associated cholesterol-related metabolites (total lipids, cholesterol, cholesterol esters, phospholipids, free and total cholesterol in large LDL measures) were associated with 112 lipoprotein measures (IDL and VLDL), apolipoprotein B and fatty acid measures. VLDL is a subclass of lipoprotein which contains apolipoprotein B and is responsible for the transport of synthesized lipids, thus making it likely for the two to be on the same pathway which links cholesterol to PCa mortality. The degree of unsaturation, which is a fatty acid measure, is also likely to be on a shared pathway of cholesterol and PCa mortality since cholesterol esters have fatty acids, which can be either saturated or unsaturated, attached to it. It is therefore likely that rather than horizontal pleiotropy, where these genetic variants affect PCa mortality through other pathways other than cholesterol ones, this is a case of vertical pleiotropy
in which multiple cholesterol, apolipoprotein measures and fatty acid measures share the same pathways and thus do not invalidate my MR analysis.

Deregulated lipid metabolism have been suggested to play an important role in PCa progression, particularly as unlike most cancerous cells, PCa cells are not glycolytic and rely on lipid biosynthesis for the extra energy demands (133). Furthermore, cholesterol is of particular interest in PCa progression, as it is a precursor of androgen synthesis and agonist of steroidogenic genes and believed to contribute to the cancerous cell proliferation mechanism $(188,204)$. PCa cells have also been shown to have the ability to esterify cholesterol to avoid cellular toxicity, whilst maintaining high levels of cholesterol required for proliferation, through the androgen regulated sterol regulatory element binding proteins $(188,196,208,209)$. In vivo, statins, a cholesterol lowering drug was shown to improve the 5- and 10-year biochemical recurrence free survival in patients who underwent radical prostatectomy (210). My findings support a causal role for cholesterol in the mechanisms of PCa progression, with higher circulating levels of cholesterol related measures, such as total cholesterol and total esterified cholesterols, cholesterol measures in large LDL particles being causally linked to increased PCa mortality.

### 5.5.1.2. Amino acids and ketone measures

I found some evidence to suggest a role for amino acids and ketones in relation to PCa progression in both the observational study and causal analyses. Leucine and isoleucine which were found to have the largest effect sizes in relation to increased metastases or PCa mortality in the observational study also showed some evidence of association with increased PCa mortality in the causal analysis. Alanine, valine and betahydroxybutyrate were found to be causally linked to increased and histidine to decreased PCa mortality. The strongest evidence was for valine and histidine. However, the MR findings were not supported by the observational analysis, with inconsistent
direction and effects of the associations with both PCa clinical progression and metastases and PCa death which were estimated with poor precision.

One genetic variant (rs1260326) was found to be associated with multiple metabolic measures from different subclasses, including amino acids. The variant is located in the regulatory area of the glucokinase regulatory gene which encodes a protein involved in the inhibition of glucokinase in liver and pancreatic cells. The gene has been previously suggested to play a role in the development of many conditions such as type 2 diabetes mellitus, non-alcoholic liver disease and metabolic syndrome (352-356). This suggests that using the genetic variant, which is associated with a range of other metabolites, to instrument for the acetate and acetoacetate could lead to invalid MR findings and would be an instance of horizontal pleiotropy. However, after the exclusion of this genetic variant in the sensitivity analyses, the estimates obtained were similar in effect size and precision to the main results.

Previous cellular evidence supports the link between leucine uptake, proliferation and malignant transformation of PCa cells in castrate resistant PCa, suggesting increased levels of leucine may play a role in PCa progression $(191,357)$. Histidine has been previously found to affect inflammation, and in a study of PCa metabolic phenotypes, lower levels of histidine was observed in patients with the most aggressive PCa (358360). Another study found serum histidine to be reduced in castrate resistant PCa, however the mechanisms of action are still not well understood (170). However, one major limitation of the current evidence of the link between circulating amino acids levels and PCa progression is that the evidence relies on studies with small sample sizes. In addition, progression indicators were not measured using follow-up data as would be expected in studies of disease progression. Rather, progression was assessed simply by comparing the levels of amino acids between participants with various PCa stages and phenotypes $(170,358)$. This may not accurately reflect the link between the metabolite levels and disease progression, given that more advanced stages will have been exposed to other factors, such as androgen deprivation therapy, and androgen exposure has been previously shown to alter amino acid metabolism $(170,361)$.

One reason for the opposite direction observed in my MR analysis for histidine and valine compared to the observational analysis, could be due to low power in the observational analysis, with effect estimates near one and large confidence intervals crossing the null. Another reason for this could be collider bias in both the MR and observational analyses. In an MR study of PCa risk in relation to metabolites, histidine and valine were not found to be associated with PCa risk, suggesting that collider bias may not affect my MR estimates of PCa progression. However, collider bias cannot be fully excluded, despite the inclusion of a collider bias correction method in my MR analysis, since the GWAS study of metabolites used to perform the MR study on PCa risk had less genetic variants ( $\mathrm{n}=881$ ) than the present GWAS $(\mathrm{n}=1600)$ to instrument the metabolites, and was performed in a smaller number of participants $(24,925$ vs 115,078$)$ $(350,362)$. An observational study of PCa risk in the ProtecT trial found some evidence that both valine and histidine were associated with reduced risk of developing PCa. These findings suggest that the association observed in the ProtecT trial for histidine with reduced metastases or PCa death may be partly due to collider bias, since histidine was shown to increase the risk of developing PCa.

### 5.5.1.3. Inflammation and fluid balance measures

In the ProtecT trial, I found some, but not strong, observational evidence of glycoprotein acetyls, which are markers of chronic inflammation, with increased PCa progression (clinical progression and metastases or PCa death), with large effect sizes and narrow confidence intervals. In my MR analysis, there was weak evidence of a causal link between glycoprotein acetyls and decreased PCa mortality, with causal estimates having small effect sizes and wide confidence intervals. The heterogenous glycoproteome of PCa tumours has been previously investigated to better understand and predict PCa detection, grading and differentiation between aggressive and nonaggressive disease, however it has yet to be fully explored (363-365). A recent study into the complexity and dynamism of the glycoproteome in PCa tissue found glycoprotein markers associated with PCa progression (366). My findings suggest that, in addition to the role observed at
tissue level, glycoproteins may also play an important role in PCa progression at circulating level however, the link is unlikely to be causal.

### 5.5.2. Strengths and limitations

This study has multiple strengths and some limitations. Firstly, it provides both observational and causal epidemiological evidence on the link between metabolites and PCa progression. The UK-wide ProtecT RCT is a population-based PSA screening trial in which participants diagnosed with PCa have similar characteristics to those in unscreened populations, making it a highly generalisable trial $(239,367)$. In addition, the ProtecT trial had a high follow-up rate at 10 years for primary outcomes (93\%) suggesting good internal validity $(33,239,368)$. Nightingale NMR platform measures metabolite levels from multiple subclasses, and has been previously shown to detect metabolites with high reproducibility and accuracy $(362,369)$. The genetic data used in the MR analysis originated from large consortia and biobank datasets, with average sample sizes of 117,313 and 67,758 for the exposure and outcome datasets, respectively. This allowed the investigation of the causal effects of metabolites on PCa mortality using an adequately powered two-sample MR study.

A limitation of this study is the inclusion of localised only PCa cases in the ProtecT trial which tend to progress slowly, and I was unable to investigate advanced disease which may have a different aetiology. Although the ProtecT trial found similar PCa mortality across treatment groups, there was evidence of increased clinical progression and metastases in the active surveillance groups compared to the other treatment options (prostatectomy and radiotherapy) (19). Since treatment was not a covariate in the main Cox regression model, if treatment affected metabolites levels in addition to clinical progression and metastases, then treatment could be a confounder and thus bias the estimates in the observational analysis. The limited follow-up time and low number of progression events in the ProtecT trial, both of which are to be expected in localised PCa cases, impacted on the power to detect associations of metabolites and PCa progression events, yielding estimates with hazard ratios near 1 and large confidence intervals. In addition, the power was further reduced due to the large number of metabolites being
assessed. All these aspects made it difficult to triangulate the observational and causal evidence, given the lack of consistency between the two, and the different limitations of each methodology.

Although principal component analysis was employed to address multiple testing, this approach can be too conservative leading to low power in detecting true associations, especially given the low number of observed events $(244,248)$. Whilst genetically instrumenting the metabolites, there were 42 SNPs per instrumented metabolite. Since multiple classes of metabolites were identified as causal, and metabolites can be highly correlated, it is difficult to establish without further mechanistic work, how these metabolites interact with one another and at organism level to cause PCa progression. This is an essential aspect that warrants further investigation, as progression regulatory mechanisms and their downstream consequences are complex. More in-depth investigations of the independent role of metabolites on progression mechanisms can be conducted, for example using multivariable MR, in which a collection of genetic variants is used to predict a set of exposure variables (370). Such further analyses can explore the effects of metabolites in particular areas of regulatory mechanisms which characterise progression. This could potentially identify mediating mechanisms, and metabolomic biomarkers that can be used as therapeutic targets in clinical and public health practice (371-373). Both the ProtecT trial and the studies within the PRACTICAL consortium are largely based on men of white ethnicity so may not be representative of the disease aetiology in other ethnicities e.g., in black men where there is a higher prevalence of disease.

### 5.6. Chapter Summary

In this Chapter, I investigated the link between baseline circulating metabolites measured using NMR and measures of PCa progression using observational (ProtecT trial) and causal analysis (based on the PRACTICAL GWAS consortium) evidence. I found suggestive observational evidence of a link between metabolite subclasses (lipoproteins, amino acids, cholesterol and glycolysis related measures, fluid and inflammation measures) and PCa progression in men with localised disease. I found
causal evidence that total cholesterol, total esterified cholesterol, sphingomyelins, apolipoprotein B, measures of IDL, LDL and VLDL cholesterol and total concentration of branched-chain amino acids (leucine, isoleucine and valine) and valine increased PCa mortality. The causal analysis also found that some HDL-related measures, histidine and the ratio of omega-6 to omega-3 fatty acids were causally associated with decreased PCa mortality. However, the observational evidence did not support the causal findings, with effect estimates centred around the null or in the opposite direction to the causal estimates and large confidence intervals. Despite inconsistencies in effect magnitude and direction, I found causal evidence and some observational evidence to support previously proposed theories on the altered amino acid, lipid and inflammation mechanisms as hallmarks of PCa progression $(76,77,187,188)$. My findings support the use of metabolomics in PCa progression research, in disease aetiology and progression biomarker development. Whilst this Chapter focused on the observational and causal link between metabolites and PCa progression, the next Chapter aims to explore the potential of metabolites in predicting PCa progression.

## Chapter 6. Effects of exercise and

 nutrition on the metabolome of men with localised Prostate Cancer: the Prostate Cancer Evidence of Exercise and Nutrition trial (PrEvENT) randomised controlled trial and a Mendelian randomisation analysis
### 6.1. Chapter overview

In the previous chapter I investigated the link between circulating metabolites and Prostate Cancer (PCa) progression. The observational associations were tested in the Prostate Testing for Cancer and Treatment (ProtecT) trial and causal estimates generated using data from the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL) consortium. As seen in Chapter 2, fruit and vegetable consumption, dairy and lycopene intake and physical activity have been hypothesised to influence PCa progression. However, the evidence has been
contradictory and no causal link has yet been established. Metabolites are the byproducts of many cellular processes and thus can contribute to the understanding of some of the mechanisms underlying cancer metabolism. Mendelian randomisation (MR) can advance knowledge of disease aetiology, through assessment of causality between exposures and outcomes. The aim of this Chapter is to investigate: i) the effects of randomised dietary (lycopene supplementation, dietary advice to increase fruit and vegetable and decrease dairy milk consumption) and physical activity (brisk walking) interventions on the circulating metabolome of men with localised PCa in the PrEvENT trial; and ii) investigate the causal effect of the metabolites altered by the interventions on PCa mortality by applying MR in the PRACTICAL consortium.

The chapter starts with a brief overview of the study populations and methodologies used, with the full details presented in Chapter 4. The results section is split into three parts. The first two parts present the results from the intention to treat (ITT) analysis, followed by the instrumental variable (IV) analysis in the PrEvENT feasibility randomised controlled trial (RCT), investigating the effects of lycopene supplementation, dietary advice to increase fruit and vegetable and decrease dairy milk consumption, and brisk walking interventions on the metabolites of men with localised PCa. The ITT analysis presents the changes in metabolite levels as a result of prescribing the intervention. The IV analysis shows the adherence-adjusted effects that the intervention had on the metabolite levels while retaining the randomised allocation. The third part presents the results of the MR analyses, estimating the causal effects of the metabolites which were found to be altered in the PrEvENT trial, on PCa mortality. The chapter ends with a discussion of the triangulated evidence from the PrEvENT trial and MR analyses, the implications for future research and the limitations of this study.

### 6.2. Research questions

1) Is the metabolome of men with PCa altered following lifestyle interventions comprising either: i) lycopene supplementation; or ii) dietary advice to increase fruit and vegetable consumption and reduce dairy milk consumption; or iii) a brisk walking physical activity intervention?
2) What is the causal effect of the metabolites altered by these lifestyle interventions on PCa mortality?

### 6.3. Methods

This section briefly describes the methodology employed in this Chapter, with full details of the methodology, the PrEvENT trial, UK Biobank (UKBB) and the PRACTICAL consortium presented in Chapter 4 (5,23,237,239,344,374).

### 6.3.1. Study population

From the 81 men recruited into the PrEvENT trial, bloods and follow-up information were available for 74 . The same dataset of study men $(n=74)$ was used in both the unadjusted and adjusted ITT analyses as well as the IV analysis. The UKBB genetic-wide association study (GWAS) dataset ( $\mathrm{n}=115,078$ ) was used to identify genetic instruments for the metabolites altered by the intervention, and the PRACTICAL mortality GWAS study ( $\mathrm{n}=75,672$ ) was used to generate the causal estimates for the instrumented metabolites on PCa death.

### 6.3.2. Measures

All measures used in this analysis were collected as part of the PrEvENT trial with the exception of the 229 serum metabolomic measures, which were quantified after the trial ended, from banked samples, using the Nightingale NMR (nuclear magnetic resonance) platform (Nightingale Health®, Helsinki, Finland). The blood samples were collected from the participant at the trial enrolment visit and aimed to be processed and frozen as soon as possible on the same day. The centrifuging of the samples was aimed to be started within 3 hours of collection. However, no formal information was collected on the length of time that the samples were stored before being frozen. The samples were frozen at -80 degrees Celsius and remained in the freezer until shipping to the NMR laboratory. All samples were transported on dry ice and were defrosted in the freezer overnight. In total 156 metabolites were quantified at both trial baseline (before
intervention) and 6-month follow-up time (6-months post randomisation). This included measures for 14 lipoprotein subclasses (small, medium, large and very large highdensity lipoprotein (HDL), low-density lipoprotein (LDL) and very low-density lipoprotein (VLDL)), apolipoproteins (A and B), fatty acids and their ratios, glycolysis related metabolites, amino acids, ketone bodies, and fluid and inflammation measures. For the unadjusted ITT and IV analyses I used the following variables: metabolite values, intervention or control arm, serum lycopene, and self-reported daily portions of fruits and vegetables, millilitres of dairy milk intake and step count per day. For the adjusted analyses, I also used age and smoking status.

### 6.3.3. Statistical analyses

### 6.3.3.1. Descriptive Analyses

The baseline characteristics for the participants were stratified by intervention arm. The distribution of continuous variables was presented using mean and standard deviation (SD) and percentage for the categorical variables.

### 6.3.3.2. Intention-to-treat analyses

Effects of the interventions on serum lycopene, fruit and vegetable and dairy milk intake, and step count were investigated using linear regression in an ITT analysis. The effects of the dietary and physical activity interventions on serum metabolic traits were also evaluated by an ITT analysis, comparing the metabolic traits at follow-up for each of the intervention arms in relation to the control group. As described in Chapter 4 section 4.3.1.2, metabolic trait concentrations were transformed to standard deviation (Zscored) and principal component analysis was used to establish a multiple testing significance threshold. Linear regression was used to examine the associations between follow-up metabolic traits and the intervention arms separately for the dietary and physical activity interventions, using the control group as reference. Robust standard errors were used as the metabolic traits were skewed and the residuals were not normally distributed (375). The principal component analysis identified 12 principal components that explained $95 \%$ of the variance in the metabolites and thus the adjusted
p-value was 0.004 ( $0.05 / 12$ ). When discussing precision, I use the pre-defined value ( $p=0.004$ ) as indicative of strong evidence and the traditional $p$-value $(p=0.05)$ for suggestive evidence. However, I assessed the strength of the associations by looking at the effect size and precision rather than simply just relying on a p-value threshold for statistical significance (347).

To assess the similarities in metabolic perturbations caused by the multiple interventions, I assessed correlations of the effect estimates obtained in the ITT analysis, between all intervention groups: dietary advice vs lycopene supplementation groups, brisk walking vs lycopene supplementation groups and brisk walking vs the dietary advice groups. This was done by fitting a linear regression model between the ITT results for each of the pair, which generated a slope, intercept and $\mathrm{r}^{2}$ statistic.

### 6.3.3.3. Instrumental variable analyses

I employed an IV analysis approach to estimate the complier-adjusted causal effects of the exposures (274). This approach maintains the randomisation status while accounting for non-adherence (men who received the intervention but did not adhere) and contamination in the controls (men who did not receive the intervention but adhered to the intervention). I used the randomisation status as the IV, to assess the full magnitude of the causal effects of changes in the exposures (serum lycopene, fruit and vegetable intake, dairy milk intake and physical activity) on follow-up metabolite measures ( $274,278,375$ ). This was performed using 2-stage least squares (2SLS) regression implemented in the STATA statistical software suite (346). Applying Ordinary Least Squares regression I computed the F-statistic and $\mathrm{r}^{2}$ for the first stage regression to assess the IV assumption that the instrument (diet or physical activity intervention arm) is sufficiently associated with the exposures: serum lycopene, self-reported fruit and vegetable intake and dairy intake, and self-reported number of steps.

For the physical activity intervention, the number of steps per day (in thousands) and percentage of days with at least 10,000 steps were investigated as instruments. I chose the measure of step count per day since the intervention would be expected to increase the step count of men. I also chose a frequency measure (number of days achieving 10,000 steps) because men in the intervention arm were asked to brisk walk for 30
minutes, on at least 5 days of the week, in addition to their current exercise. If men adhered, the extra exercise would lead to at least an extra 5,000 steps on the brisk walking days, thus driving the total number of steps per day closer to 10,000 .

All IV regression coefficients were calculated in units of 1-SD metabolite concentration per one unit increase in exposure ( $\mu \mathrm{mol} / \mathrm{L}$ for lycopene, number of servings for fruit and vegetable intake, 100 mL for dairy milk intake and mean number of steps per day and percentage of days with at least 10,000 steps for physical activity). IV analyses were conducted in the dietary advice vs control groups, using self-reported dairy milk intake and number of fruit and vegetable servings per day as the exposure. In the physical activity intervention, I chose mean steps per days as exposure, since it had similar suitability to be an instrument (F-statistic and $\mathrm{r}^{2}$ ) compared to percentage of days with more 10,000 steps. In addition, it had a better interpretability and public health message (i.e., metabolite levels change per one thousand steps per day).

### 6.3.3.4. Sensitivity analyses for intention-to-treat and instrumental variable analyses

No covariates were included in the main analysis (unadjusted) because the distribution of confounders was expected to be balanced across the arms with successful randomisation. Before undertaking any analyses, I set out to also produce an ITT analysis, adjusted for baseline metabolic measures, and any other baseline characteristics that looked unbalanced (244). When conducting the analyses, I investigated baseline characteristics between the intervention groups to check which variables showed some evidence of being unbalanced. Although randomisation should lead to random allocation of participants to the intervention arms, in the present feasibility RCT I detected some evidence ( $\mathrm{p}<0.05$ ) that certain metabolites, age and lifestyle factors were unbalanced across intervention arms (Table 6.1 and Tables C1, C2 and C3, Appendix C). Based on the differences at baseline in these potential confounders, post-hoc sensitivity analyses were performed for both ITT and IV analyses in the diet arm, adjusting for baseline metabolic traits and smoking in the lycopene arm, and metabolic traits in the dietary advice arm. In the physical activity arm, post-hoc groups sensitivity analyses were performed by adjusting for age and baseline metabolic traits. The aim of the
adjusted ITT analysis was to investigate the effects that the potentially unbalanced variable had on the effect size and precision of the main results. All participants with follow-up metabolite data also had baseline data and thus they were all included in the sensitivity analysis.

### 6.3.3.5. Causal analysis

A detailed description of the use of MR in feasibility studies is presented in Chapter 4, Section 4.4.4. Briefly, a two-stage randomisation process was employed. Figure 6.1 presents the two-stage randomisation processed applied in the PrEvENT trial, and PRACTICAL consortium. The first stage assessed the effects of the intermediate endpoints (metabolites) which were altered by the intervention, using ITT and IV analyses. In stage 2, causal estimates of the genetically instrumented metabolites on PCa mortality are generated. First, the metabolites were instrumented using genetic variants from the UKBB GWAS (i.e genetic variants associated with the metabolites are identified). The instruments (genetic variants associated with the metabolites) were then used to run the MR analysis which estimates the causal effect that these instruments (genetic variants) have on PCa mortality using genetic mortality data in the PRACTICAL consortium. By combining the two stages, I was able to identify metabolites the that the interventions in the PrEvENT altered and estimate the causal effect of these altered metabolites on PCa mortality.

The two-sample MR analysis used in this Chapter was conducted as part of Chapter 5's causal analysis (Section 5.6.3). In this Chapter, I only present the causal effects on PCa mortality of the metabolites for which I found some evidence of being altered by the interventions. MR analysis is not meant to provide definitive evidence of the causal effect of the interventions on PCa mortality, but rather indicate potential mechanisms of action and provide evidence for future work. In addition, the measurement of metabolites in the PrEvENT trial was not a primary outcome, the sample size was small, and the statistical threshold obtained through principal component analysis was conservative. For these reasons, I took through to the MR analysis the metabolites with the strongest evidence of association ( $p$-value $<0.05$ ) observed in the trial and treated the investigated metabolites with caution. In addition, due to the low power to detect
changes in metabolites in the PrEvENT trial, I conducted an ad hoc sensitivity analysis and selected the metabolites to be taken forward to the MR analysis based on their effect sizes in each of the trial's unadjusted ITT analysis. I compared the metabolites which were taken forward using the $\mathrm{p}<0.05$ cut-off to the top $10 \%$ of metabolites within each arm. For each of these two sensitivity analyses, when estimating the causal effect, I used the Wald ratio or inverse variance weighted as the main estimating method. As sensitivity analyses to the causal estimation process, I used alternative estimating methods (MR Egger, simple mode, weighted median and mean), where possible (if more than 3 instruments available).

Figure 6.1: The two-stage randomisation process in the Prostate Cancer Evidence of Exercise and Nutrition trial and Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome consortium.

Stage 1: The effects of the interventions on metabolites (PrEvENT)


Stage 2: The causal estimates of altered metabolites on PCa mortality

$I T T=$ Intention-to-treat; $I V=$ Instrumental variable; $M R=$ Mendelian randomisation; $\operatorname{PrEvENT}=$ Prostate Cancer Evidence of Exercise and Nutrition trial; PRACTICAL = Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome consortium; $U K B B=$ UK biobank; $P C a=$ Prostate Cancer

### 6.4. Results

### 6.4.1. Descriptive analysis

Eighty-seven men were invited to take part in the trial and 81 men were randomised. Three participants withdrew, three were lost to follow-up and one participant did not have blood samples taken at his 6-month follow-up appointment (Chapter 4, Figure 4.2). In total, metabolomics data were available for 81 and 74 participants at trial baseline and 6-month follow-up, respectively. The participants had a mean age of 64.0 years (SD=6.54), a BMI of $26.6 \mathrm{~kg} / \mathrm{m}^{2}(\mathrm{SD}=3.39)$ and a mean baseline PSA of $0.17 \mathrm{ng} / \mathrm{ml}$ ( $\mathrm{SD}=1.25$ ) following prostatectomy.

Table 6.1 presents the baseline characteristics for men, by intervention arm, for each of the dietary and physical activity groups. For most of the clinical, sociodemographic and lifestyle factors investigated, there were no marked differences across the intervention arms within the dietary or physical activity intervention groups, as would be anticipated given the randomised design of the study. However, in the dietary intervention, the participants in the control arm, compared to those in the lycopene and dietary advice arms, appeared to have higher PSA, were more likely to be smokers or ex-smokers and drink hazardous levels of alcohol. The prevalence of PCa family history was higher in the lycopene arm compared to the dietary advice and control arms. Overall, there were few people with diabetes in the trial, with noninformative numbers across the trial arms. In the physical activity group, age was higher in the brisk walking arm than in the control arm.

Table 6.1: Baseline characteristics of the men by intervention group

|  | Dietary arm ( $\mathrm{N}=81$ ) |  |  |  |  |  | Physical activity arm ( $\mathrm{N}=81$ ) |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Lycopene |  | Dietary advice |  | Control |  | Brisk Walking |  | Control |  |
|  | N | Mean (SD) | N | Mean (SD) | N | Mean (SD) | N | Mean (SD) | N | Mean (SD) |
| Age (years) | 28 | 62.54 (7.98) | 27 | 64.63 (5.44) | 26 | 65.06 (5.78) | 42 | 65.77 (5.49) | 39 | 62.19 (7.12) |
| Body mass index (kg/m2) | 24 | 26.16 (3.61) | 26 | 26.71 (3.57) | 26 | 26.82 (3.07) | 39 | 26.58 (3.34) | 37 | 26.56 (3.49) |
| PSA, ( $\mathrm{ng} / \mathrm{ml}$ ) | 28 | 0.04 (0.13) | 27 | 0.00 (0.01) | 26 | 0.43 (2.10) | 39 | 0.03 (0.10) | 34 | 0.33 (1.83) |
| Lycopene (umol/l) | 27 | 1.36 (0.46) | 26 | 1.26 (0.37) | 26 | 1.20 (0.40) | 40 | 1.34 (0.42) | 38 | 1.21 (0.40) |
| Fruit \& Vegetable intake (servings/day) | 25 | 8.43 (4.10) | 26 | 8.82 (5.51) | 25 | 7.31 (2.39) | 36 | 8.85 (4.83) | 41 | 7.60 (3.50) |
| Amount of dairy milk intake(100mL/day) | 25 | 3.08 (1.62) | 26 | 2.80 (1.86) | 26 | 2.37 (1.57) | 40 | 2.97 (1.80) | 37 | 2.50 (1.55) |
|  | N | Percent (\%) | N | Percent (\%) | N | Percent (\%) | N | Percent (\%) | N | Percent (\%) |
| Smoking status |  |  |  |  |  |  |  |  |  |  |
| Current smoker | 0 | 0 | 0 | 0 | 2 | 7.69 | 1 | 2.7 | 1 | 2.5 |
| Ex-smoker | 9 | 37.5 | 13 | 48.15 | 16 | 61.54 | 21 | 56.76 | 17 | 42.5 |
| Never Smoker | 15 | 62.5 | 14 | 51.85 | 8 | 30.77 | 15 | 40.54 | 22 | 55.0 |
| Family history PCa |  |  |  |  |  |  |  |  |  |  |
| Yes | 11 | 44.0 | 3 | 11.11 | 2 | 7.69 | 9 | 24.32 | 7 | 17.1 |
| No | 10 | 40.0 | 14 | 51.85 | 16 | 61.54 | 16 | 43.24 | 24 | 58.5 |
| Don't know | 4 | 16.0 | 10 | 37.04 | 7 | 26.92 | 11 | 29.73 | 10 | 24.4 |
| Alcohol |  |  |  |  |  |  |  |  |  |  |
| Non-Drinker | 14 | 50.0 | 9 | 33.33 | 8 | 30.77 | 17 | 43.59 | 14 | 33.3 |
| Moderate | 11 | 39.29 | 16 | 59.26 | 11 | 42.31 | 17 | 43.59 | 21 | 50.0 |
| Hazardous | 3 | 10.71 | 2 | 7.41 | 7 | 26.92 | 5 | 12.82 | 7 | 16.7 |
| Diabetes |  |  |  |  |  |  |  |  |  |  |
| Yes | 1 | 4.0 | 2 | 7.41 | 1 | 3.85 | 3 | 8.11 | 1 | 2.4 |
| No | 22 | 88.0 | 23 | 85.19 | 21 | 80.77 | 29 | 78.38 | 37 | 90.2 |
| Unknown | 2 | 8.0 | 2 | 7.41 | 4 | 15.38 | 5 | 13.51 | 3 | 7.3 |

[^3]To investigate whether the interventions achieved what they set out to do (that is, the lycopene arm to increase serum lycopene, the dietary advice to increase fruit and vegetable and decrease dairy milk intake, and the brisk walking to increase step count), I compared levels of serum lycopene, self-reported fruit and vegetable and dairy milk intake, and selfreported step count for each arm, at the 6-month follow-up (Table 6.2).

Serum lycopene was highest in the dietary advice arm, with a mean concentration of 1.51 $\mu \mathrm{mol} / \mathrm{L}(\mathrm{SD}=0.50)$, followed by the lycopene arm, with a mean concentration of $1.30 \mu \mathrm{~mol} / \mathrm{L}$ $(S D=0.45)$. There was evidence that serum lycopene levels were altered in the dietary advice $\operatorname{arm}(p=0.03)$. There was weak evidence to suggest that serum lycopene was altered in the lycopene supplement arm ( $\mathrm{p}=0.47$ ).

Mean fruit and vegetable mean intake was higher in dietary advice (mean=9.89, SD=0.50) and lycopene (mean $=8.54, \mathrm{SD}=4.63$ ) arms compared to controls. However, only the dietary arm showed strong evidence of having higher intake of fruit and vegetable compared to the control arm ( $\mathrm{p}=0.04$ ). Dairy milk intake was lowest in the dietary advice arm (mean=0.39, $\mathrm{SD}=1.03$ ). I found strong evidence to suggest that dairy milk intake was lower in the dietary advice compared to the control arm ( $\mathrm{p}<0.001$ ). There was weak evidence to suggest that serum lycopene, fruit and vegetable, and dairy milk intake were different in the brisk walking arm compared to controls.

The mean number of steps per day was comparable across the dietary arms, with weak evidence to suggest that it was different in the lycopene and dietary arms compared to controls. The mean number of steps per day was higher in the brisk walking group (mean $=8,844, \mathrm{SD}=3,171$ ) compared to controls (mean $=7,494, \mathrm{SD}=2,364$ ). There was some evidence to suggest that the physical activity intervention altered the mean number of steps per day (mean difference $=1,350,95 \% \mathrm{CI}: 41.6-2658, \mathrm{p}=0.04$ ).

Table 6.2: Effects of the interventions on lycopene serum levels, self-reported fruit and vegetable, and dairy milk intake and step count.

| Diet arm ( $\mathrm{N}=74$ ) |  |  |  |  |  |  |  |  | Physical activity arm (N=74) |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Lycopene |  | Dietary advice |  | Control |  | Lycopene vs. control |  | Dietary advice vs. control |  | Brisk Walking |  | Control |  | Brisk walking vs. control |  |
|  | N | Mean (SD) | N | Mean (SD) | N | Mean (SD) | Mean <br> difference <br> ( $95 \% \mathrm{CI}$ ) | p-value* | Mean <br> difference <br> (95\%CI) | p-value* | N | Mean <br> (SD) | N | Mean <br> (SD) | Mean <br> difference <br> ( $95 \% \mathrm{CI}$ ) | p-value* |
| Lycopene (umol/l) | 28 | 1.30 (0.45) | 25 | 1.51 (0.50) | 21 | $\begin{gathered} 1.20 \\ (0.49) \end{gathered}$ | $\begin{gathered} 0.10 \\ (-0.17 ; 0.38) \end{gathered}$ | 0.47 | $\begin{gathered} 0.31 \\ (0.031 ; 0.60) \end{gathered}$ | 0.03 | 38 | 1.34 (0.48) | 36 | $\begin{gathered} 1.34 \\ (0.51) \end{gathered}$ | $\begin{gathered} 0.008 \\ (-0.22 ; 0.24) \end{gathered}$ | 0.95 |
| Fruit \& Vegetable intake(servings/day) | 27 | 8.55 (4.63) | 25 | 9.89 (4.72) | 19 | $\begin{gathered} 7.12 \\ (3.09) \end{gathered}$ | $\begin{gathered} 1.42 \\ (-1.16 ; 4.0) \end{gathered}$ | 0.28 | $\begin{gathered} 2.76 \\ (0.15 ; 5.38) \end{gathered}$ | 0.04 | 37 | 8.52 (4.20) | 34 | $\begin{gathered} 8.77 \\ (4.64) \end{gathered}$ | $\begin{gathered} -0.24 \\ (-2.3 ; 1.85) \end{gathered}$ | 0.82 |
| Amount of dairy milk intake $(100 \mathrm{~mL} /$ day $)$ | 27 | 2.90 (2.16) | 25 | 0.39 (1.03) | 19 | $\begin{gathered} 2.65 \\ (1.61) \end{gathered}$ | $\begin{gathered} 0.25 \\ (-.76 ; 1.26) \end{gathered}$ | 0.62 | $\begin{gathered} -2.26 \\ (-3.29-1.24) \end{gathered}$ | <0.001 | 37 | 2.01 (1.97) | 34 | $\begin{gathered} 1.89 \\ (2.12) \end{gathered}$ | $\begin{gathered} 0.11 \\ (-.86 ; 1.08) \end{gathered}$ | 0.82 |
| Mean step count (steps per day) | 28 | $\begin{gathered} 8,315 \\ (3,045) \end{gathered}$ | 24 | $\begin{gathered} 8,133 \\ (2,633) \end{gathered}$ | 21 | $\begin{gathered} 8,047 \\ (3,001) \end{gathered}$ | $\begin{gathered} 268 \\ (-1,404 ; \\ 1,939) \end{gathered}$ | 0.75 | $\begin{gathered} 85 \\ (-1,645 ; \\ 1,815) \end{gathered}$ | 0.92 | 37 | $\begin{gathered} 8,844 \\ (3,171) \end{gathered}$ | 36 | $\begin{gathered} 7,494 \\ (2,364) \end{gathered}$ | $\begin{gathered} 1,350 \\ (42 ; 2,658) \end{gathered}$ | 0.04 |

$S D=$ standard deviation


### 6.4.2. Intention-to-treat analysis

ITT analysis was used to investigate the effects of each of the interventions in relation to individual metabolites. The results of the ITT analysis are presented in the next sections, separately for each of the dietary and physical activity interventions.

### 6.4.2.1. Dietary Group

I found only weak evidence to suggest the dietary interventions altered the metabolome of men with PCa. Overall patterns showed a reduction in circulating serum levels of lipoprotein particles, cholesterol, glycerides, phospholipids, apolipoproteins, most fatty acids, amino acids and glycolysis related metabolites, ratio of saturated fatty acid to total fatty acids, albumin and glycoprotein acetyls (dietary advice only). Circulating large and very large HDL related measures, LDL and HDL particle size, degree of unsaturation, fatty acid ratios, citrate (dietary advice only), glutamine (dietary advice only), glycine and acetate there showed an increase associated with both interventions. Generally larger effects were observed in the dietary advice group compared to the lycopene group (Figure 6.2a). Weak evidence showed that alanine decreased in the lycopene arm. In the dietary advice arm small HDL concentration, lactate, pyruvate, saturated fatty acid to fatty acid ratio decreased and some fatty acid ratios (linoleic and omega-6 fatty acids to total fatty acids) increased. However, there was no strong statistical evidence of change in individual metabolites in either the lycopene supplementation or the dietary advice arms compared to the controls, with no associations reaching my pre-specified multiple testing significance threshold of $\mathrm{p}<0.004$ (Tables $C 4$ and C5, Appendix C). In the lycopene arm, the largest effect size was observed for alanine ( $\beta=-0.61 ; 95 \% \mathrm{CI}=-$ $1.18,-0.046 ; p=0.035$, where $\beta$ represents the change in SD for the metabolic trait). In the dietary arm, the largest effect sizes were observed for pyruvate ( $\beta=-0.68 ; 95 \% \mathrm{CI}=-1.28$, $-0.082 ; p=0.026$ ) and acetate ( $\beta=0.64 ; 95 \% C I=0.054 ; 1.22 ; p=0.03$ ).

After adjustment for baseline metabolic trait levels in the dietary and additionally for smoking in the lycopene arm, most metabolites maintained their effect size and direction. The exceptions were some VLDL, HDL and triglyceride, as well as fatty acid measures (degree of unsaturation, percentage of omega- 6 to totally fatty acids) which reversed their direction of association following adjustment (Figure 6.2b). Alanine showed strong evidence of association in the lycopene intervention in the adjusted
model ( $\beta=-0.86$; 95\% CI $=-1.35,-0.37 ; p=0.001$ ) (Table C6, Appendix C). After adjustment for metabolic traits in the dietary advice arm, the observed effect in the unadjusted analysis on acetate ( $\beta=0.63 ; 95 \% \mathrm{CI}=-0.02,1.27 ; \mathrm{p}=0.06$ ) and pyruvate $(\beta=$ $-0.60 ; 95 \% \mathrm{CI}=-1.14,-0.056 . ; \mathrm{p}=0.03$ ) were maintained (Table C7, Appendix C).

### 6.4.2.2. Physical activity Group

Overall, there was some, but not strong evidence that the brisk walking intervention altered the metabolome of men with PCa. There was a reduction in VLDL, LDL and their associated particle sizes, and an increase in HDL and its associated particle sizes in the physical activity arm compared to controls (Figure 6.2c). Measures of cholesterol levels generally decreased, except HDL cholesterol measures which increased. Triglycerides and fatty acids experienced a reduction overall, with an increase in the degree of unsaturation. Glycolysis related metabolites decreased except glycerol which showed an increase. Branched-chain amino acids and glycoprotein acetyls decreased, and ketone bodies increased. There was some, but not strong evidence that multiple measures of VLDL and saturated fatty acids decreased and some fatty acid ratios increased. The ratio of linoleic acid to total fatty acids ( $\beta=0.56 ; 95 \% \mathrm{CI}=0.12,1.01 ; \mathrm{p}=$ 0.01 ) and cholesterol esters in extremely large VLDL ( $\beta=-0.53 ; 95 \% \mathrm{CI}=-0.98,-0.07 ; p=$ 0.024 ) showed the strongest statistical evidence of associations (Table C8, Appendix C). However, there was no strong evidence that the brisk walking intervention altered individual metabolic traits at my pre-specified multiple testing significance threshold ( $\mathrm{p}<0.004$ ).

After adjustment for baseline metabolic trait levels and age, most metabolic measures maintained the direction and magnitude of effect observed in the unadjusted analysis, except for HDL measures which showed a change in direction. There was strong evidence that brisk walking decreased measures of triglycerides, saturated fatty acids, total phosphoglycerides and phosphatidylcholine (Figure 6.2d). The largest effect size was observed for triglycerides in large HDL ( $\beta=-0.57 ; 95 \% \mathrm{CI}=-0.9,-0.23 ; p=0.001$ ) (Table C9, Appendix C).

Figure 6.2: Forest plots of the overall effects (Intention-to-treat) of the intervention vs control on serum metabolites, by intervention arm in the unadjusted and adjusted analyses
(A)


Closed symblols: $\mathrm{P}<0.004 ;$ Open symbols: $\mathrm{P}>=0.004$
(B)


## (C)

- Brisk Walking

Closed symblols: $\mathrm{P}<0.004$; Open symbols: $\mathrm{P}>=0.004$

(D)


- Brisk Walking

Closed symblols: $\mathrm{P}<0.004$; Open symbols: $\mathrm{P}>=0.004$

VLDL=Very low-density lipoprotein; LDL=low-density lipoprotein; HDL=high-density lipoprotein; C=Cholesterol; MUFA=monounsaturated fatty acids; PUFA=polyunsaturated fatty acids; $n 3=$ omega; $n 6=$ omega 6; SFA $=$ saturated fatty acids

The circles and squares represent the $\beta$ coefficient of the ITT analysis (linear regression), between the metabolites and intervention arms at follow-up. The $\beta$ coefficients are in units of 1 SD metabolite concentration. The closed symbols indicate values that reached the pre-defined statistical significance level ( $\mathrm{p}<0.004$ ). The horizonal bars are the $95 \%$ confidence intervals. (A) Overall effects of the dietary advice and lycopene arms vs control on individual metabolites in the unadjusted ITT analysis (B) Overall effects of the dietary advice and lycopene arms vs control on individual metabolites in the adjusted (baseline metabolites for the dietary advice arm and baseline metabolites and smoking for the lycopene arm) ITT analysis (C) Overall effects of the brisk walking arm vs control on individual metabolites in the unadjusted ITT analysis (D) Overall effects of the brisk walking arm vs control on individual metabolites in the adjusted (baseline metabolic traits and age) ITT analysis.

### 6.4.2.3 Correlation of ITT results across the dietary and physical activity interventions

To compare how consistent the metabolic changes from each of the dietary and brisk walking interventions were with each other, I regressed the results from the individual metabolite ITT and graphed these on scatterplots (lycopene vs dietary advice, lycopene vs brisk walking, and brisk walking vs dietary advice). The correlation slopes ( $\beta$ ) generated using linear regression, intercept and $\mathrm{R}^{2}=$ are presented in Figure 6.1. There was some correlation between the effects of the lycopene supplementation and the dietary advice on metabolic measures at 6-months follow-up ( $\beta=0.60 \pm 0.04, R^{2}=0.64$ ) (Figure 6.3A). After adjusting for baseline metabolic measures, the correlation was maintained ( $\beta=0.62 \pm 0.03, \mathrm{R}^{2}=0.67$ ) (Figure 6.3B). There was a weak correlation between the effects of brisk walking advice with lycopene supplementation at the 6-month follow-up ( $\beta=0.19 \pm 0.12, \mathrm{R}^{2}=0.02$ ) (Figure 6.3C). There was some correlation between the effects of brisk walking and dietary advice ( $\beta=0.51 \pm 0.08, \mathrm{R}^{2}=0.21$ ) (Figure 6.3 D ) on metabolic measures at 6-months follow-up.

Figure 6.3: Comparison of the overall effects of the interventions on metabolites from the Intention-to-treat analysis intervention arm vs control.
(A)

(C)

(B)

(D)


The plots present the SD difference in metabolite concentration between the intervention arm and control. Each dot on the scatterplot represents a different metabolite. The dotted light grey line has a slope of 1 and an intercept of zero. The dark grey dotted line represents the linear fit of the two models. The closer the two lines are to overlap, the more similar magnitudes and direction of the associations.
(A) Comparison of the overall effects of metabolites between the lycopene arm vs controls ( x -axis) and dietary advice vs controls (y-axis), from the unadjusted ITT analysis.
(B) Comparison of the overall effects of metabolites between the lycopene arm vs controls ( x -axis) and dietary advice vs controls (y-axis), from the ITT analysis, adjusted for baseline metabolic traits and smoking.
(C) Comparison of the overall effects of metabolites between the lycopene arm vs controls (x-axis) and brisk walking vs controls (y-axis), from the unadjusted ITT analysis.
(D) Comparison of the overall effects of metabolites between the dietary advice arm vs controls ( $x$-axis) and brisk walking vs controls (y-axis), from the unadjusted ITT analysis.

### 6.4.3. Instrumental variable analysis

IV analysis was used, with the randomised intervention assignment as the instrument, to investigate the causal effect of changes in serum lycopene, fruit and vegetable and dairy milk intake on metabolites at 6 -months of follow-up. The F-statistic and $\mathrm{R}^{2}$ from the unadjusted linear regression results (first step) of the exposures (serum lycopene, fruit and vegetable, dairy milk intake, step count measures) and intervention arms are presented in Table 6.3 and the p-value is presented in Table 6.2. There was weak evidence to suggest that the lycopene intervention modified levels of serum lycopene ( $\beta$ $=0.1 ; 95 \% \mathrm{CI}:-0.17,0.37 ; \mathrm{p}=0.46 ; \mathrm{F}$-statistic $=0.57$ ). In the dietary advice arm, there was some evidence that the intervention modified levels of self-reported fruit and vegetable intake ( $\beta=2.76 ; 95 \%$ CI: $0.25,5.3 ; p=0.03 ; F$-statistic $=4.9$ ) and strong evidence that it modified self-reported dairy milk intake ( $\beta=-2.3 ; 95 \% \mathrm{CI}:-3.12,-1.5 ; \mathrm{p}<0.001$; F statistic=32.1). There was some evidence that the physical activity intervention modified the self-reported mean number of steps taken per day ( $\beta$ (per 1,000 steps) $=1,350 ; 95 \% \mathrm{CI}$ : 42-2658; $p=0.04$; F-statistic: 4.23) and percentage of days with more than 10,000 steps ( $\beta$ $=0.16 ; 95 \% \mathrm{CI}: 0.004-0.23 ; \mathrm{p}=0.04 ;$ F-statistic $=4.25$ ).

The results of the unadjusted linear regression of the instrumented exposures on the individual metabolite levels (second step of the IV analysis) are presented in Appendix C, Tables C10, C11 and C12. Generally, the results of the IV were consistent with those from the ITT. The metabolites with the largest effect sizes from the ITT also had the largest effects in IV analyses. None of the associations surpassed my conservative significance threshold. The same patterns were seen after adjusting for baseline metabolites in the dietary advice arm and baseline metabolites and age in the IV to ITT results (Appendix C, Tables C13, C14, C15).

Table 6.3: Linear regression results of exposures and intervention arms (Instrumental variable instruments) from the first step of the Instrumental variable analysis.

| Exposure measure | N | F-statistic | R 2 |
| :--- | :---: | :---: | :---: |
| Serum lycopene (umol/l) | 48 | 0.57 | 0.12 |
| Fruit and vegetable intake (servings per day) | 44 | 4.9 | 0.11 |
| Mean steps (thousands of steps / day) | 74 | 4.23 | 0.06 |
| Percentage of days with more than 10,000 steps (percentage of <br> total reported days) | 74 | 4.25 | 0.06 |
| Dairy milk intake (100mL/day) | 44 | 32.1 | 0.43 |

### 6.4.4. Mendelian Randomisation analysis

Twenty-three metabolites were taken forward from the PrEvENT trial into the MR analysis with the list of metabolites and the F-statistic for the instruments presented in Appendix C Tables C16 and C17, respectively. The causal estimates of the metabolites on PCa mortality, using the inverse variance weighted method are presented in Figure 6.4 and Table 6.4. Generally, I found evidence of a causal association between cholesterol esters in medium VLDL (OR=1.16; 95\%CI:1.01-1.34; $\mathrm{p}=0.03$ ) and phospholipids in chylomicrons and extremely large VLDL (OR=1.14; 95\%CI:1.01-1.29; $\mathrm{p}=0.04$ ) with increased PCa mortality. The largest causal effect sizes for the metabolites taken forward were observed for lactate ( $\mathrm{OR}=1.49 ; 95 \% \mathrm{CI}: 0.95-2.34 ; \mathrm{p}=0.09$ ) and alanine ( $\mathrm{OR}=1.24$; $95 \% \mathrm{CI}: 0.99-1.55 ; \mathrm{p}=0.06$ ). In the MR sensitivity analysis, the causal estimates of the metabolites in relation to PCa mortality had similar effect size and the same direction of association when using different estimation methods (MR Egger, simple mode, weighted median and mode) (Table C18, Appendix C).

When selecting metabolites to be taken forward into the MR analysis from the PrEvENT trial based on their effect sizes (top 10\% in the unadjusted ITT in each arm), rather than p -value ( $\mathrm{p}<0.05$ ), 32 metabolites were identified. Approximately 16 metabolites per intervention arm were selected for the dietary and physical activity intervention. I did not conduct this sensitivity analysis in the lycopene arm, due to the lack of effect the intervention had on what it set out to achieve (i.e., elevate serum lycopene). All
metabolites, except for cholesterol esters in large VLDL in the brisk walking arm, which were selected using p-value were also selected when applying the $10 \%$ effect size cut-off. The extra 6 metabolites selected using the $10 \%$ effect size cut-off were all from the dietary intervention arm. The list of the 6 extra metabolites selected in this sensitivity analysis, the F-statistic of the instruments and MR results are presented in Appendix C Tables C19 and C20. Generally, I found weak evidence of a causal link between the additional metabolites identified in this sensitivity analysis and PCa mortality, with small effect sizes centred around the null.

Figure 6.4: Causal effects (odds ratio) of the altered metabolites in the Prostate Cancer Evidence of Exercise and Nutrition trial on Prostate Cancer mortality using data from the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome consortium.


LA_pct=Percentage of linoleic to total fatty acid; PUFA_pct=Percentage of polyunsaturated fatty acids to total fatty acids;
S_HDL_P=Concentration of small high-density lipoprotein particles; SFA_pct= Percentage of saturated fatty acids to total fatty
 low-density lipoprotein; XL_HDL_TG=Triglycerides in very large high-density lipoprotein; XXL_VLDL_CE=Cholesterol esters in chylomicrons and extremely large very low-density lipoprotein; $X L \_V L D L \_P L=$ Phospholipids in very large very low-density lipoprotein; XXL_VLDL_P=Concentration of chylomicrons and extremely large very low-density lipoprotein particles; SFA=; XXL_VLDL_C= Total cholesterol in chylomicrons and extremely large very low-density lipoprotein particles ; XXL_VLDL_FC= Free cholesterol in chylomicrons and extremely large very low-density lipoprotein particles; XXL_VLDL_L= Total lipids in chylomicrons and extremely large very low-density lipoprotein particles; $X L_{-} V L D L \_F C=$ Free cholesterol in very large very lowdensity lipoprotein particles ; L_VLDL_CE= Cholesterol esters in large very low-density lipoprotein particles; XXL_VLDL_PL= Phospholipids in chylomicrons and extremely large very low-density lipoprotein particles; M_VLDL_CE=Cholesterol esters in medium very low-density lipoprotein particles; Ala=Alanine

Causal estimates were calculated using inverse variance weighted, except for pyruvate and acetate where the Wald method was used. Acetate has an upper $95 \%$ CI of 5.54 , however for graphical aesthetics it was set to 3 .

Table 6.4: Causal effect estimates of altered metabolites from the Prostate Cancer Evidence of Exercise and Nutrition trial on Prostate Cancer mortality in the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome consortium, in the inverse variance weighted or Wald ratio methods.

| Metabolite ${ }^{*}$ | Intervention |  |  |  |
| :--- | :--- | :--- | :--- | :--- |
| arm | ITT effect <br> size and 95\% <br> CI | ITT adjusted <br> effect size and <br> $95 \%$ CI | MR PCa death <br> OR** estimate <br> and 95\% CI |  |
| Pyruvate | Dietary |  |  |  |
| advice | -0.68 | $(-1.28 ;-0.08)$ | $(-1.15 ;-0.06)$ | $(0.73 ; 1.24)$ |
| Cholesterol esters in <br> medium VLDL | Brisk walking | -0.47 | $(-0.93 ;-0.01)$ | $(-0.63 ; 0.02)$ |


| Metabolite* | Intervention arm | ITT effect size and $95 \%$ CI | ITT adjusted ${ }^{* *}$ effect size and 95\%CI | MR PCa death OR ${ }^{* * *}$ estimate and 95\% CI |
| :---: | :---: | :---: | :---: | :---: |
| Free cholesterol very large VLDL | Brisk walking | $\begin{aligned} & -0.48 \\ & (-0.94 ;-0.02) \end{aligned}$ | $\begin{aligned} & -0.36 \\ & (-0.67 ;-0.05) \end{aligned}$ | $\begin{aligned} & 1.12 \\ & 0.99 ; 1.27) \end{aligned}$ |
| Phospholipids in very large VLDL | Brisk walking | $\begin{aligned} & -0.46 \\ & (-0.92 ;-0.005) \end{aligned}$ | $\begin{aligned} & -0.35 \\ & (-0.66 ;-0.04) \end{aligned}$ | $\begin{aligned} & 1.09 \\ & (0.97 ; 1.24) \end{aligned}$ |
| Cholesterol in extremely large VLDL | Brisk walking | $\begin{aligned} & -0.51 \\ & (-0.96 ;-0.05) \end{aligned}$ | $\begin{aligned} & -0.38 \\ & (-0.70 ;-0.06) \end{aligned}$ | $\begin{aligned} & 1.11 \\ & 0.98 ; 1.25) \end{aligned}$ |
| Cholesterol esters in chylomicrons and extremely large VLDL | Brisk walking | $\begin{aligned} & -0.53 \\ & (-0.98 ;-0.07) \end{aligned}$ | $\begin{aligned} & -0.38 \\ & (-0.71 ;-0.05) \end{aligned}$ | $\begin{aligned} & 1.08 \\ & (0.97 ; 1.22) \end{aligned}$ |
| Free cholesterol in chylomicrons and extremely large VLDL | Brisk walking | $\begin{aligned} & -0.47 \\ & (-0.93 ;-0.01) \end{aligned}$ | $\begin{aligned} & -0.37 \\ & (-0.68 ;-0.07) \end{aligned}$ | $\begin{aligned} & 1.11 \\ & (0.98 ; 1.25) \end{aligned}$ |
| Total lipids in chylomicrons and extremely large VLDL | Brisk walking | $\begin{aligned} & -0.47 \\ & (-0.93 ;-0.01) \end{aligned}$ | $\begin{aligned} & -0.37 \\ & (-0.68 ;-0.06) \end{aligned}$ | $\begin{aligned} & 1.11 \\ & (0.99 ; 1.25)) \end{aligned}$ |
| Concentration of chylomicrons and extremely large VLDL | Brisk walking | $\begin{aligned} & -0.47 \\ & (-0.93 ;-0.006) \end{aligned}$ | $\begin{aligned} & -0.37 \\ & (-0.68 ;-0.05) \end{aligned}$ | $\begin{aligned} & 1.10 \\ & 0.97 ; 1.24) \end{aligned}$ |
| Phospholipids in chylomicrons and extremely large VLDL | Brisk walking | $\begin{aligned} & -0.47 \\ & (-0.93 ;-0.008) \end{aligned}$ | $\begin{aligned} & -0.38 \\ & (0.769-0.07) \end{aligned}$ | $\begin{aligned} & 1.14 \\ & (1.01 ; 1.29) \end{aligned}$ |
| Ratio of polysaturated fatty acids to total fatty acids | Brisk <br> Walking | $\begin{aligned} & 0.49 \\ & (0.03 ; 0.94) \end{aligned}$ | $\begin{aligned} & 0.31 \\ & (-0.01 ; 0.62) \end{aligned}$ | $\begin{aligned} & 0.97 \\ & (0.84 ; 1.16) \end{aligned}$ |

ITT=Intention-to-treat; CI=Confidence interval; MR=Mendelian randomisation; $\mathrm{PCa}=$ Prostate Cancer; VLDL=very lowdensity lipoprotein; HDL=high-density lipoprotein
*Metabolites taken forward from the unadjusted ITT analysis ( $\mathrm{p}<0.05$ )
**Adjusted for baseline metabolites and smoking in the lycopene arm, baseline metabolites in the dietary advice, and baseline metabolites and age in the brisk walking arm.
****Odds ratios estimates were obtained from MR of PCa survival in the PRACTICAL consortium, per 1 standard deviation increase in genetically instrumented levels of metabolites, using the IVW method

### 6.5. Discussion

In this Chapter, I used the PrEvENT feasibility trial to investigate the effects of dietary and lifestyle interventions on the metabolite levels of men with PCa and estimate their causal effect on PCa mortality. In the dietary (lycopene supplementation or increasing fruit and vegetable consumption and decreasing dairy milk intake) interventions, general patterns showed a reduction of circulating serum levels of lipoprotein particles, cholesterol, glycerides, phospholipids, apolipoproteins, most fatty acids, amino acids and glycolysis related metabolites, ratio of saturated fatty acid to total fatty acids, albumin and glycoprotein acetyls (dietary advice only). I also only found weak evidence of an increase in circulating large and very large HDL related measures, LDL and HDL particle size, degree of unsaturation, fatty acid ratios, citrate (dietary advice only), glutamine (dietary advice only), glycine and acetate with both interventions for the dietary interventions. The strongest evidence $(0.004<\mathrm{p}<0.05)$ showed a reduction of small HDL concentration, lactate, pyruvate, saturated fatty acid to fatty acid ratio and an increase in some fatty acid ratios (linoleic, omega-6, saturated fatty acids to total fatty acids) in the dietary advice arm. In the lycopene arm, alanine showed some reduction ( $0.004<\mathrm{p}<0.05$ ).

In the physical activity intervention, there was some, but not strong evidence of reduction of lipoproteins, VLDL and LDL cholesterol measures, fatty acids, glycolysis and amino acids. I also found weak evidence for an increase in HDL related molecules, degree of unsaturation, glycerol and ketone bodies in the physical activity arm. The strongest evidence $(0.004<\mathrm{p}<0.05)$ was for a reduction of VLDL measures, triglycerides in HDL and saturated fatty acids and for an increase of some ratios of fatty acids (linoleic, omega-6, polyunsaturated fatty acids to total fatty acids).

The MR analysis found causal evidence, that genetically instrumented levels of phospholipids in chylomicrons and extremely large VLDL, cholesterol esters in medium VLDL, may increase the risk of PCa death. There was also some evidence that alanine and lactate may be causally associated with increased PCa mortality. Alanine was found to be decreased by the lycopene intervention, lactate decreased by the dietary intervention and the two VLDL measures decreased by the brisk walking intervention. My findings suggest that lifestyle and dietary interventions may lower some VLDL measures, alanine and lactate, for which there is some evidence that they are causally
associated with increased PCa mortality, in men with localised PCa who underwent prostatectomy. However, it is important to note that none of the observed associations in the unadjusted ITT analysis surpassed the statistical thresholds of 0.004. In addition, the lycopene supplementation intervention failed to raise serum lycopene levels and thus the findings on alanine must be interpreted with caution.

### 6.5.1. Main findings

### 6.5.1.1. Lycopene supplementation

In this study, an intervention of oral lycopene supplementation did not increase serum lycopene. I found weak evidence that the intervention altered the metabolome of men with PCa. I found some evidence that randomisation to the lycopene supplementation lowered circulating levels of alanine in the unadjusted ITT, with the association becoming strong after adjusting for baseline metabolites and smoking. Alanine is a precursor of sarcosine, an amino acid that has been previously linked to PC progression, by altering the amino acid metabolism ( $157,191,211,212,215$ ). The MR analysis found some evidence that alanine is causally linked to increased risk of PCa progression. An MR of PCa risk found some evidence that alanine is linked with an increase in risk of developing PCa $(326,350)$. However, it is unlikely that the protective effect observed in the survival MR could be an artefact of collider bias alone since the present MR analysis was adjusted for collider bias. In addition to these limitations, the lycopene supplement did not increase serum lycopene in the present feasibility RCT, whilst lycopene supplementation has been previously associated with a 2-fold increase in serum levels in other studies (376). The lack of effect observed on the metabolome despite very good adherence for the lycopene intervention, could be due to the fact the lycopene supplementation failed to achieve its aim to elevate lycopene serum level. Therefore, any evidence from this intervention arm should be carefully interpreted as suggestive evidence.

### 6.5.1.2. Dietary advice intervention

I found that a dietary advice intervention to increase fruit and vegetable, and decrease dairy milk consumption, altered self-reported measures of fruit and vegetable and dairy milk intake. Changes in lipid profiles were observed in the dietary advice arm compared to the control arm, and are consistent with previously published evidence of the lipid
lowering effects of a plant-based diet $(377,378)$. In the IV analysis, I found that, assuming the causal effect of the interventions on the metabolites (identified by the ITT), and taking into account compliance, the metabolites were generally altered by decreased dairy milk intake and increased portions of fruits and vegetables in the same direction as shown in the ITT (complier adjusted causal effect). When adjusting for smoking status and metabolite measures, the associations from the unadjusted analysis were generally maintained, with similar effect sizes and direction, the exception being some fatty acid ratios (linoleic acid and omega-6 to total fatty acids).

In the ITT, I found some, but not strong evidence that increased fruit and vegetable and decreased dairy milk intake, lowered circulating pyruvate, lactate, concentration of small HDL measures and ratio of saturated fatty acids to total fatty acids. The MR analysis found some evidence that lactate, which was decreased by the dietary intervention may be causally linked to increased PCa progression. In the ITT, there was some evidence that the ratio of omega- 6 to total fatty acids increased in the intervention arm, while the MR analysis found weak evidence of association of this metabolite and decreased PCa mortality.

It is established that lipids play an important role in cancer metabolism, particularly in supporting the metabolic alterations that are required in the rapid proliferation of cancerous cells $(379,380)$. The hypothesis described in Chapter 3, section 3.4.5, states that PCa cells, unlike most other cancer cells, use active glycolysis and oxidative phosphorylation to generate the required increase in energy for cell proliferation and only use the aerobic glycolysis in later stages of the disease (221-226). Therefore, taken together the observational and causal analyses provide suggestive evidence that an intervention of dietary advice to increase fruit and vegetable and decrease dairy milk intake may lower levels of lactate which may be causally linked with increased PCa mortality. This is supported by the current evidence, since it is hypothesized that glycolysis metabolites (lactate and pyruvate) are required to generate the extra energy required for proliferation in later stage PCa. The findings thus suggest a potential role for interventions of increased fruit and vegetable and reduced dairy consumption in PCa progression prevention.

However, an adequately powered RCT must be conducted to establish the effect of fruit and vegetable and dairy milk consumption on the metabolome of men with localised disease. An important limitation is that glycolysis can occur as soon as the sample is
collected, if the sample is not kept at very low temperatures and plasma and red blood cells are not separated immediately $(381,382)$. Whilst of great interest in cancer metabolism, the glycolysis measures I found to be altered by the dietary advice intervention could be in part or even completely due to sample handling, although this is unlikely due to the randomised design of the trial, which would expect the samples from the control participants to undergo the same changes.

### 6.5.1.3. Physical activity intervention

Changes in lipid profiles were observed in the brisk walking group and were consistent with previously published evidence of the effects of physical activity on lipids $(383,384)$. In the unadjusted ITT, I found that ratios of omega-6, linoleic acid and polyunsaturated fatty acids to total fatty acids were increased and saturated fatty acids decreased in the physical activity intervention group. However, the estimates were imprecise (wide 95\% confidence intervals). I also found 12 measures of VLDL and one measure of HDL cholesterol to be lowered by the physical activity intervention, again with imprecise estimates which did not reach my pre-specified threshold of significance ( $\mathrm{p}>0.004$ ). After adjusting for baseline metabolic measures and age, of the metabolites identified in the unadjusted ITT, only saturated fatty acids surpassed the statistical threshold, with a comparable effect size to that in the unadjusted ITT analysis.

The MR analysis found evidence that two of the metabolites which were decreased by the intervention (phospholipids in chylomicrons and extremely large VLDL and cholesterol esters in medium VLDL) increased PCa mortality, by $14 \%$ and $16 \%$, respectively. The MR sensitivity analyses (MR Egger, weighted median, weighted mode, simple mode), which used different estimation methods, found similar effect sizes in the same direction of association. Taken together, the findings suggest that an intervention of brisk walking, may lower levels of phospholipids in chylomicrons and extremely large VLDL and cholesterol esters in medium VLDL, which the MR analysis found to be causally linked to increased PCa mortality.

Recent evidence suggests that cholesterol esterification and free cholesterol may play a role in cholesterol homeostasis. This is hypothesized to occur through the inhibition of cholesterol esterification which in turn may suppress the synthesis of androgens and cell migration, thus delaying PCa metastases via the $\beta$-catenin and the mevalonate pathways $(379,380,385)$. However, the mechanism through which cholesterol affects cancer
progression is yet to be established. Despite a large number of studies aiming to provide mechanistic insight in the observational association between cholesterol and PCa progression, it is still uncertain whether cholesterol affects PCa progression by increasing cell proliferation, acting on signalling molecules and cell-cycle regulators or synthesis of androgens within the tumours, which leads to castration-resistant PCa (62).

Although there was weak evidence to suggest that ratios of various types of fatty acids to total fatty acids (saturated fatty acids, omega-6 fatty acids and linoleic acids to total fatty acids) causally altered PCa mortality in this MR study, it has been previously suggested that such measures may prevent progression. Studies that investigated the link between PCa cells that are hormone independent or castrate-resistant have shown that progression is inhibited by fatty acid and enzymes related to fatty acid synthesis (386-389). Hormone independent cells are seen in advanced cases, and in this MR study I could not differentiate between localised and advanced PCa cases. The ratio of different types of fatty acids to total fatty acids may affect PCa phenotypes differently and this may be one reason for the lack of causal evidence observed. My results suggest that a brisk walking intervention may alter the ratio of saturated, omega-6 and linoleic fatty acids to total fatty acid in men with localised PCa, but no causal link was observed between these measures and overall PCa mortality. However, further research is required to understand the effects of saturated, omega-6 and linoleic fatty acids on PCa phenotypes. In addition, a recent study of PCa progression in mice has shown that fatty acids and cholesterol in combination activate pathways that enhance development of prostate cancer stem cells, which have been linked to disease progression $(390,391)$.

### 6.5.2. Strengths and Limitations

My study has several strengths. Firstly, the adherence to the lycopene intervention was high and most measured confounders were balanced when compared to the randomised control group. The adherence to the dietary advice and brisk walking interventions were previously reported to be lower than other diet and physical activity interventions in cancer patients (127). However, the definitions used in the PrEvENT trial compared to other studies were stricter, expecting men to follow the intervention $90 \%$ of the time vs $70 \%$, in other studies. Therefore, the differences in metabolic traits observed in the ITT and IV analyses are likely to be due to the intervention itself rather than confounders. Secondly, to measure the metabolic traits, I used a highly reproducible and
quantitatively accurate, high throughput NMR platform, covering a wide range of metabolic pathways. Thirdly, using a 2-sample MR approach, I leveraged the statistical power of this small feasibility trial by using data from the largest GWAS study of metabolites to date ( $\mathrm{n}=115,078$ ) and the largest PCa consortium with over $75,672 \mathrm{PCa}$ cases and 7,914 PCa deaths ( $23,35,392$ ).

One of the main limitations is that the lycopene capsule administered in the lycopene arm did not elevate serum lycopene in the intervention group. The aim of the PrEvENT trial which was a feasibility trial was to measure adherence to the intervention and not the change in serum lycopene. Previous studies have generally found that oral lycopene supplementation with capsules increases serum lycopene levels $(80,393,394)$. However, most of these studies used a soft-gel capsule containing 15 mg of lycopene (Lyc-OMato®), while the PrEvENT trial used a Holland and Barrett hard supplement capsule containing 10 mg lycopene ( $80,127,330,393,394$ ). In addition, the Lyc-O-Mato® supplement contains small amounts of other compounds, such as $ß$-carotene, which has been suggested to increase lycopene absorption (395). NMR analysis, conducted by the University of Bristol, Department of Chemistry, of the current and Lyc-O-Mato® supplements found that both tablets contained lycopene compounds of similar chemical structure (all-Elycopene.). The analysis also found that the red pigment, which is a measure of the amount of lycopene contained, was darker in the red Lyc-O-Mato® supplement compared to the Holland and Barret tablets, suggesting a smaller concentration was present, although no formal quantification was performed. Given the lack of serum lycopene elevation at follow-up post supplementation, any evidence or lack of, observed in the lycopene arm should be interpreted with caution, since the effect of a lycopene supplementation intervention cannot be generalised from the PrEvENT trial.

Whilst adherence to the physical activity intervention $90 \%$ of the time was $53.8 \%$, compensatory mechanisms cannot be dismissed(127). Compensatory mechanisms in physical activity interventions refers to the fact that participants may change their physical activity patterns as a result of enrolling in the study. These potential mechanisms could be due to fatigue, lack of time and motivation to perform additional physical activity, perceptions of overexertion and drive to be less active(396). In the $\operatorname{PrEvENT}$ trial for example, participants may have decreased or changes their physical activity patterns already existent before enrolling. This could bias the results, given that
even though participants in the physical activity active arm adhered to the intervention, the observed effects of brisk walking intervention on the metabolome may be altered by both the intervention, but also the change in other physical activity. Therefore, when comparing the physical activity group to the control group, we cannot be certain that compensatory mechanisms were not in action, leading to biased estimates towards the null. Although one of the PrEvENT trial's exclusion criteria was participation in high level of regular physical activity (5 or more times a week), the results could still be subject to compensatory mechanisms, through other type of physical activity that did not fall into this category such as active commuting and strenuous exercise for less than 5 days per week.

Another limitation of my study is the use of an RCT with a feasibility design, which does not allow for a definitive inference to be made with regards to the effect of the intervention on metabolic traits since they do not set out to achieve this. The lower adherence to some of the interventions and the poor performance to raise serum lycopene of the supplement, raised concerns of the validity of the ITT analysis. To overcome this, I employed an IV analysis to look at the effects of the exposures altered by the intervention on the follow-up metabolic traits; however, this was not possible for lycopene. In addition, despite its RCT design, there was evidence of some unbalanced measured confounders, for which I employed sensitivity analyses. Lastly, as expected in an 'omics study, I assessed a large number of metabolic traits, some of which were highly correlated. To overcome this I employed a previously developed principal component analysis technique to reduce the multiple testing burden and improve power for detecting true associations $(244,248)$. However, since the power was limited, I evaluated patterns in the metabolites altered with the largest effect sizes, rather than employing a strict statistical significance threshold. This helps avoid taking an overly conservative approach which could lead to low power in identifying true associations, particularly in my small study $(244,248,347)$.

One key assumption of MR analysis is that the genetic variant should only be associated with the outcome (PCa mortality) through the exposure of instrument - also known as the no-horizontal pleiotropy or exclusion restriction assumption. There is evidence of non-specificity of genetic variants used to instrument some of the metabolites, where some correlated metabolites share common instruments. While this limits causal
inference of individual metabolites, it can still allow the identification of relevant classes of metabolites. Nonetheless, where possible I conducted sensitivity analyses by applying MR Egger and weighted median approach to assess the presence of horizontal pleiotropy.

A specific assumption of the two-sample MR approach is that the two samples represent the same underlying population. While the two GWAS studies used in this analysis were restricted to individuals of European ancestry, there were differences in the average ages of the study samples, and the UKBB metabolite GWAS study included both males and females whereas in PRACTICAL all participants were male. One limitation of this analysis is that the samples included in the two-sample MR have some overlap, which has been shown to introduce bias into the MR analysis (303). However, only $10 \%$ of the total number of participants in the PRACTICAL consortium originated from UKBB.

Finally, it is important to highlight that absence of metabolites related to both intervention and PCa survival does not mean that the intervention cannot influence PCa via other mechanism which I was unable to explore in this study. Well designed and conducted RCTs can help minimise the bias that would otherwise occur in observational studies assessing lifestyle and dietary factors, while MR studies can decrease confounding and reverse causation and help predict long-term outcomes in intervention studies of short durations. By integrating both these approaches in a two-step randomisation design, my results suggest that interventions aimed to increase serum lycopene, reduce dairy and increase fruit and vegetable consumption, and increase brisk walking may alter the metabolome of men with localised PCa, by particularly improving the lipid profiles, with evidence that lowering measures of cholesterol, may decrease PCa mortality.

### 6.5.3. Implications for future research

Future studies are required to establish definitive effects of dairy, fruit and vegetable, lycopene and brisk walking in the survival of men with PCa and to further understand mechanisms of action. Most of the studies looking at the effect of dietary and physical activity interventions on PCa progression did not have clinical progression outcomes, such as metastases. Instead, they used proxy measures, such as PSA increases, which could lead to classifying men as having 'PCa progression' that may not be clinically
important. A recent RCT which followed up men with localised PCa over three years, found no difference in progression when men were randomised to an intervention to increase vegetable consumption (104). However, $50 \%$ of the patients were classed as progressed using PSA measures, which suggests that study conclusions are not generalisable given that most men with PCa have a much lower rate of clinically defined progression (e.g. metastases, castrate resistance), particularly over the study's follow-up period of three years $(2,19)$. Generally, the RCTs conducted to assess the effects of physical activity on localised PCa progression had short follow-up times (up to 2 years) which given the indolent nature of the disease would not comprehensively capture progression patterns. Despite recent advances in treatment and diagnosis of PCa, castration resistant PCa remains a lethal disease with poor clinical outcomes and thus better understanding the aetiology and prediction factors of disease progression as well as therapeutic targets are needed (397-399).

The findings in my study warrant further investigation to establish whether lifestyle and dietary interventions could alter cholesterol, fatty acid, lactate and alanine levels and thus act as new therapeutic targets for the prevention of PCa progression to fatal disease. Future diet and lifestyle studies, appropriately powered to detect changes in metabolites in localised PCa, along with an MR analysis performed in localised PCa phenotypes, could provide triangulation of evidence. This would allow more generalisable inferences to be drawn in the first instance, without requiring RCTs with long follow-up.

From a practical perspective, when starting a large-scale study which aims to assess the effects of lycopene supplementation in PCa patients, it is advisable to investigate the efficacy of the supplement in increasing serum lycopene levels in a subset of the study population, using a pilot or a feasibility study design. This should apply even when using an over-the-counter supplement from a high street store, given the lack of efficacy at elevating serum lycopene observed in the PrEvENT trial. In addition, assumptions on efficacy are hard to make given other factors in the manufacturing and presentation of the supplement. For example, supplements can contain additional compounds, have a different isomer and presentation and it is uncertain what the effect of these aspects is on bioavailability and absorption into the bloodstream.

### 6.6. Chapter Summary

In this Chapter, I used ITT and IV analyses of a factorial feasibility RCT to investigate the effects of diet and lifestyle interventions on metabolomic measures in men with localised PCa. I then used MR analysis to estimate the potential causal effect of the altered metabolites on PCa mortality. In a factorial feasibility RCT, there was some, but not strong evidence that an intervention of lycopene supplementation, an intervention of dietary advice to increase fruit and vegetable intake and decrease dairy milk consumption, and an intervention of brisk walking, may alter VLDL, HDL and fatty acid measures, acetate, lactate, pyruvate and alanine, in men with PCa who underwent radical prostatectomy. In turn, my causal analysis found some evidence of an effect of raised levels of some of these metabolites (phospholipids and cholesterol esters in medium VLDL, alanine and lactate) and increased risk of PCa mortality. My findings suggest that lifestyle interventions (increased fruit and vegetable and decreased dairy milk intake, brisk walking) may decrease some metabolites in men with PCa which are causally linked to increased PCa mortality. Further studies are required to directly determine the effects of diet and physical activity interventions on PCa mortality or intermediate biomarkers of PCa progression.

# Chapter 7. Evaluating the performance of multimetabolomic risk scores to 

 predict PCa mortality
### 7.1. Chapter overview

In this Chapter, I aimed to explore the link between individual metabolites and Prostate Cancer (PCa) specific and all-cause mortality and characterise and validate a PCaspecific mortality multi-metabolite risk score (PCa mRS) developed in a large population cohort (UK Biobank) (344). I assessed the predictive performance of the score compared to a previously developed all-cause mortality multi-metabolite risk score (all-cause mRS) as well as against the established prognostic Gleason scoring system $(271,400)$. The PCa mRS had been previously trained by investigators from Nightingale Health in the UK Biobank (UKBB) cohort of 24,895 men with PCa of whom 1,234 died of PCa. The allcause mortality mRS had been previously trained Deelen et. al using data from 12 cohorts (44,168 men and 5,512 all-cause deaths). To get a detailed understanding of the PCa mRS clinical value, I evaluated the capacity of the three scores (PCa mRS, all-cause mRS, and Gleason score) to predict PCa mortality in a clinical setting of PCa men, diagnosed with localised and locally advanced disease, through the PSA testing intervention in the Prostate Testing for Cancer and Treatment (ProtecT) cohort.

The Chapter opens with a methodological summary, detailing the models which were investigated, with the statistical analysis methodology being described in detail in Chapter 4. First, I present the associations of individual metabolites in relation to PCa and all-cause mortality in ProtecT ( $\mathrm{n}=2,093$ ), a case-only cohort. I then present the descriptive and main analyses of the evaluation of the two metabolomic risk scores in the case only and a case-control study design, with cases and controls from the ProtecT cohort. The Chapter ends with a discussion of the main findings and their implications for future research, and the strengths and weaknesses of the study.

### 7.2. Research questions

1) Are circulating metabolites associated with PCa specific and all-cause mortality in men with PSA detected PCa?
2) Does the PCa mRS model predict PCa or all-cause mortality in men with PSA detected PCa ?
3) Does the all-cause mRS model predict PCa death or all-cause mortality in men with PSA detected PCa ?

### 7.3. Methods

In this section I introduce the two clinical prediction models based on measures of circulating metabolites and briefly present the methods I used to analyse their performance at predicting PCa mortality. The full methodology used in this Chapter and details of the study population and measured variables in the Protec $T$ cohort are described in detail in Chapter 4.

### 7.3.1. Study population

In this Chapter the main analyses in the case-only study were conducted using a subset of participants from the ProtecT cohort (33). These included participants who were diagnosed with localised and locally advanced PCa as part of PSA testing to identify men with localised prostate cancer for the ProtecT randomised controlled trial (RCT). Only men with serum metabolomics data who were followed up for mortality through the ProtecT cohort were included in this analysis ( $\mathrm{n}=2,093$ ). There is some overlap of these men with those selected in Chapter 5. The main difference is that in Chapter 5, I selected men with localised disease who were followed-up for progression (clinical progression, metastases or PCa death) through the ProtecT RCT. In this Chapter, to maximise the number of deaths which was the outcome of interest, I selected all participants for which mortality data was available. Some of the men were therefore not followed-up through the ProtecT RCT, but rather as part of the comprehensive ProtecT cohort, for which only death (PCa specific or all-cause) was available as progression
indicator (33). In sensitivity analyses, I restricted the cohort to men with localised PCa (i.e., clinical stages T 1 and T 2 ).

NMR metabolomic studies that assessed the links between metabolites and all-cause mortality have been previously investigated in multiple large cohorts of participants and found similar patterns $(271,401,402)$. Associations between metabolites and all-cause mortality in PCa cases and controls should therefore show similar patterns. By comparing the findings, in cases and controls, to those of other studies, I examined whether unexpected patterns are observed between the cases and controls. This could indicate issues such as sample quality and prevalent PCa being unevenly distributed between cases and controls, which could bias the analyses proposed. Therefore, to address these issues, I investigated the link between all-cause mortality in both cases and controls in the ProtecT study. I used the same cohort for cases as described above, and a previously selected set of controls from the comprehensive ProtecT cohort, from the same stratum as the cases, matched on a 5-year age-band (age at PSA test) and GP practice for which I had metabolomic data $(\mathrm{n}=2,167)(33,350,403)$.

### 7.3.2. Measures

The variables from the ProtecT cohort which were used in this Chapter's analysis are: age, centre, PSA at baseline, clinical stage, Gleason grade, and PCa- and all-cause death. The metabolomic risk models which were evaluated in this Chapter were the all-cause mRS and PCa mRS, which are described in detail in Chapter 4, section 4.3.5.4.

### 7.3.2.1 Metabolomic measures

Two hundred and twenty-seven circulating metabolites were identified and quantified using the Nightingale NMR platform as described in Chapter 3 (section 3.2.2.1) and 4 (section 4.2.1.7.2). The samples were collected at the ProtecT cohort study baseline (the participant's first visit) for both cases and controls.

### 7.3.2.2 Outcomes

In this analysis, the outcomes used were PCa and all-cause mortality as obtained through the procedures of the ProtecT cohort study. Cause of death was generated by an independent committee for both studies which assessed each individual death record (404). Details on data linkage of the participant study records with mortality and
committee procedures are described in Chapter 4 section 4.2.1.6. Briefly, the cause of death was recorded as PCa if this was definite or probable death due to PCa during the median 10-year follow-up, as assessed by the Independent Cause of Death committee blind to the death certificate recorded cause of death.

### 7.3.3. Statistical analyses

For the descriptive statistics, I computed mean and standard deviation for continuous variables and percentages for categorical data. To establish differences within PCa and all-cause death groups, I used $t$-test and $\chi 2$ tests for continuous and categorical variables, respectively.

I used Cox proportional hazards regression to examine the association between metabolomic measures (both individual serum metabolite levels and metabolomic risk scores) and risk of PCa-specific and all-cause mortality in cases. The minimally adjusted model included age and centre as covariates. Additional covariates included in the fully adjusted model were age, centre, PSA at baseline, clinical stage and Gleason score. I used Cox proportional hazard regression to investigate the individual metabolomic measures and the metabolomic risk scores in cases, controls and the cases and controls pooled. I evaluated the correlation between the metabolites used in the metabolomic risk scores using Pearson's correlation coefficient (405). In addition, I regressed both age and BMI against individual metabolite measures as a quality control check, since both factors have been extensively characterised in relation to metabolite levels in the literature.

To evaluate the performance of the metabolomic risk scores at predicting PCa death and all-cause mortality, I used area under the receiver operating curve (AUROC) in i) all PCa cases, ii) localised PCa only, and iii) all cases and controls (pooled). To assess the discrimination ability of the models, I calculated positive and negative predictive values and accuracy. I also used the Cohen's Kappa statistic to evaluate discrimination performance corrected for outcome class imbalances (number of cases vs controls) (261). I used calibration curves to assess how the predicted risk for PCa or all cause death matched the observed outcome proportion for each of the models (266).

### 7.4. Results

### 7.4.1. Baseline descriptive analyses

The median follow-up time for the Protec T cases was 9.31 years (interquartile range 7.98 10.8 ) and 9.7 years (interquartile range: $8.28-11.25$ ) for the controls. The baseline characteristics for the cases and controls, broken down by their mortality status (PCa specific or all-cause death) are presented in Table 7.1. On average, men diagnosed with PCa who subsequently died from PCa were older (mean difference $=2.0$ years, $95 \% \mathrm{CI}=.33-3.69 ; \mathrm{p}=0.02$ ), had a higher baseline PSA (mean difference $=1.49 \mathrm{ng} / \mathrm{ml}$, $95 \% \mathrm{CI}: 39-2.59 ; \mathrm{p}<0.001$ ) and were more likely to have Gleason score higher than 7 compared to PCa survivors ( $\mathrm{p}<0.001$ ). Men with PCa who died of any cause were older (mean difference $=2.17$ years; $95 \%$ CI:1.49-2.84; $\mathrm{p}<0.001$ ), had a higher baseline PSA (mean difference $=0.78 \mathrm{ng} / \mathrm{ml}, 95 \% \mathrm{CI}=.34=1.23 ; \mathrm{p}=0.001$ ), were more likely to have a Gleason score higher than 7 ( $\mathrm{p}<0.001$ ), be current or former smokers( $\mathrm{p}<0.001$ ) and hazardous drinkers ( $\mathrm{p}=0.03$ ) compared to survivors. Within men with no PCa (controls), men who died of any cause were older (mean difference $=2.2$ years; $95 \% \mathrm{CI}: 1.48-2.90$; $\mathrm{p}<0.001$ ) and were more likely to be current or former smokers ( $\mathrm{p}<0.001$ ) and hazardous drinkers ( $\mathrm{p}=0.07$ ).

Table 7.1: Baseline characteristics for: i) men with a Prostate Cancer specific death vs. survivors in cases; and ii) all-cause deaths vs survivors in controls in the Prostate Testing for Cancer and Treatment cohort.

| PCa cases |  |  |  |  |  |  |  |  |  |  | Controls (PCa free) |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Variable | PCa specific survivors and deaths |  |  |  |  | All-cause survivors and deaths |  |  |  |  | All-cause survivors and deaths |  |  |  |  |
|  | Survivors |  | PCa deaths |  | P-value* | Survivors |  | All-cause deaths |  | P-value* | Survivors |  | All-cause deaths |  |  |
|  | N | Mean (SD) | N | Mean <br> (SD) |  | N | Mean (SD) | N | Mean (SD) |  | N | Mean(SD) | N | Mean(SD) | P-value* |
| Age (years) | 2,059 | 62.3 (4.96) | 34 | 64.3 (4.32) | 0.02 | 1,864 | 62.07 (4.96) | 229 | 64.23 (4.51) | <0.001 | 1960 | 62.2 (4.96) | 207 | 64.4 (4.71) | <0.001 |
| Body mass index ( $\mathrm{kg} / \mathrm{m} 2$ ) | 1,398 | 27.2 (3.61) | 27 | 27.2 (3.81) | 0.98 | 1,280 | 27.17 (3.51) | 145 | 27.56 (4.44) | 0.22 | 1,285 | 27.4 (3.71) | 110 | 27.42 (4.06) | 0.85 |
| Overweight and obese $(\% \text { BMI }>25 \text { ) }$ | 1,024 | 73.3 | 17 | 63.0 | 0.23 | 939 | 73.4 | 102 | 70.3 | 0.44 | 924 | 71.9\% | 78 | 70.9 | 0.82 |
| PSA baseline ( $\mathrm{ng} / \mathrm{ml}$ ) | 2,059 | 5.93 (3.22) | 34 | 7.42 (4.18) | 0.008 | 1,864 | 5.87 (3.17) | 229 | 6.65 (3.73) | 0.001 | 1960 | 1.42 (1.32) | 207 | 1.43 (1.49) | 0.93 |
| Gleason score         <br> Gleason $<=7(\%)$ 1,993 96.8 23 67.7 1,809 97.1 207 90.4 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Gleason >=8 (\%) | 66 | 3.21 | 11 | 32.4 | <0.001 | 66 | 2.95 | 22 | 9.61 | <0.001 |  |  |  |  |  |
| Alcohol Intake |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Hazardous drinker (\%) | 63 | 4.15 | * | * |  | 63 | 3.87 | 13 | 8.44 |  | 73 | 5.22 | 13 | 10.8 |  |
| Moderate drinker (\%) | 1,139 | 91.9 | 25 | 83.3 |  | 1,139 | 92.1 | 136 | 88.3 |  | 1,268 | 90.4 | 111 | 86.1 |  |
| Non-drinker (\%) | 60 | 3.95 | * | * | 0.05 | 60 | 4.02 | 5 | 3.25 | 0.03 | 58 | 4.2 | 5 | 3.9 | 0.07 |
| Smoking Status |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Current smoker (\%) | 187 | 12.3 | 5 | 16.1 |  | 187 | 11.1 | 37 | 23.7 |  | 175 | 12.5 | 32 | 24.8 |  |
| Former smoker (\%) | 759 | 49.8 | 16 | 51.6 |  | 759 | 49.3 | 86 | 55.1 |  | 733 | 52.4 | 70 | 54.3 |  |
| Never smoker (\%) | 577 | 37.9 | 10 | 32.3 | 0.73 | 577 | 39.6 | 33 | 21.1 | <0.001 | 492 | 35.1 | 27 | 20.9 | <0.001 |

*The p-value tests the null hypothesis that there is no difference between the survivors and deaths (PCa or all-cause) groups
$P C a=$ Prostate Cancer; $S D=$ standard deviation; $P S A=$ Prostate specific antigen; $B M I=$ Body mass index

Body mass index was only used for the sensitivity analyses, for regressing individual metabolites on BMI as a quality control check.

### 7.4.2. Individual metabolites

### 7.4.2.1. Metabolite - survival associations among Prostate specific antigen-detected Prostate Cancer cases-only

To establish how strongly metabolite levels were associated with survival over the follow-up period of ProtecT, I used Cox regression to calculate the hazard ratios (HR) for each metabolite in relation to all-cause and PCa-specific mortality in men detected with PCa. Figures 7.1 and 7.2 present the estimates of the associations (HR per 1SD increase in each metabolite) between individual metabolite levels and PCa-specific and all-cause mortality, respectively, using the minimally and fully adjusted models (Tables D1 and D 2 , Appendix D). Isoleucine ( $\mathrm{HR}=1.56,95 \% \mathrm{CI}$ : 1.18-2.01, $\mathrm{p}=0.002$ ) and leucine ( $\mathrm{HR}=1.53$, $95 \% \mathrm{CI}: 1.16-0.2 .03, \mathrm{p}=0.002$ ) were positively associated with PCa specific death (Supplementary Table D1, Appendix D). Generally, there was some evidence that very low-density (LDL) lipoprotein measures, branched-chain amino acids, ketone bodies and glycoprotein acetyls were associated with increased risk of PCa specific death in the minimally adjusted model. However, the statistical evidence supporting the associations was weak to modest and the p-values did not reach the pre-defined conservative significance threshold ( $\mathrm{p}<0.003$ ).

For the all-cause mortality the direction and magnitude of effect with individual metabolites were for the most part consistent with those for PCa specific death (Figure 7.2 and Supplementary table D2, Appendix D). The exceptions were: i) LDL particle size and phenylalanine which showed a positive association with all-cause mortality in contrast to a negative association with PCa-specific death; and ii) docosahexaenoic acid and omega- 3 fatty acids which were inversely associated with all-cause mortality in contrast to a positive association with PCa-specific mortality, iii) amino acids (isoleucine, leucine and valine) which showed some evidence of association with increased PCa mortality and weak evidence of association with all-cause mortality. The estimates of the effect sizes for the all-cause analysis were precise, with smaller $p$-values and narrower confidence intervals compared to the PCa-specific mortality analysis (to be expected given the much greater number of all-cause versus PCa-specific deaths). There was strong evidence that degree of unsaturation, cholesterol esters in small HDL and multiple fatty acid ratios (omega3,omega-6, linoleic acid, polyunsaturated fatty acids to
total fatty acids) were inversely associated with all-cause mortality. I also found strong evidence of association for the ratio of saturated fatty acids and monounsaturated fatty acids to total fatty acids, glucose, acetoacetate, acetate, LDL particle size and phospholipids in small HDL particles with increased all-cause mortality. The largest effect sizes were observed for LDL particle size (HR=1.26; 95\%CI:1.11-1.42; $\mathrm{p}=0.0004$ ) and the ratio of saturated fatty acids to total fatty acids ( $\mathrm{HR}=1.25 ; 95 \% \mathrm{CI}: 1.11-1.42$; $\mathrm{p}=0.0004$ ). The fully adjusted model found that the direction of the association and effect size were maintained; however, the effect estimates had wider confidence intervals.

### 7.4.2.2 Metabolite - survival associations among Prostate specific antigen-detected Prostate Cancer cases vs controls

To investigate any potential effects that prevalent PCa or sample quality could have on the metabolite measures in the ProtecT cohort, I compared the associations of all-cause mortality in cases versus controls. Generally, associations of the individual biomarkers with all-cause mortality in the controls are in the same direction and had attenuated effects compared to the PCa cases (Supplementary Table D5 and Figure D1, Appendix D). Some exceptions were VLDL and triglyceride measures where the effect sizes were larger in PCa cases compared to controls. Concentration of VLDL measures were negatively associated with all-cause mortality in the controls, and positively associated in the cases. However, the effect sizes for most of these instances were imprecise with confidence intervals crossing 1 . Acetate, phenylalanine and some fatty acid measures had associations in the same direction for cases, controls and in a pooled analysis that combined cases and controls, with precise estimates and narrower confidence intervals. LDL particle size was strongly positively associated with both cases ( $\mathrm{HR}=1.25$, $95 \% \mathrm{CI}: 1.41-1.82 ; \mathrm{p}<0.0001$ ) and controls ( $\mathrm{HR}=1.26,95 \% \mathrm{CI}: 1.11-1.43 ; \mathrm{p}<0.0001$ ), with the effect size larger in the controls.

I investigated the links between BMI and age with metabolite levels and found that age and BMI were generally associated with individual metabolites in cases, controls and in a pooled analysis of cases and controls, with the effect sizes in the same direction and of similar in magnitude (Supplementary Figures D2 and D3, and Tables D7 and D8, Appendix D). There was evidence that the effect sizes for the associations of the metabolite levels with age were higher for measures of cholesterol (total, remnant, LDL,
esterified, and free), phosphoglycerides, phosphatidylcholine, sphongomyelins, alipoproteins A and B and fatty acid measures (total, linoleic acid, omega-6 fatty acids, polyunsaturated and mono fatty acids).

Figure 7.1: Hazard ratios (per 1 standard deviation) and their 95\% confidence intervals for associations between individual metabolites and Prostate Cancer specific mortality (2,059 survivors, 34 deaths) in minimally* and fully adjusted** Cox regression models in the Prostate Testing for Cancer and Treatment cohort.

## Lipoprotein subclasses

| Extremely large VLDL | $\square$ |
| :---: | :---: |
| Very large VLDL | $\square$ |
| Large VLDL | $\square-0$ |
| Medium VLDL | $\square-\mathrm{O}$ |
| Small VLDL | $\square 0$ |
| Very Small VLDL | 8- |
| IDL | $=0$ |
| Large LDL | $=0$ |
| Medium LDL | $=-$ |
| Small LDL | $=0$ |
| Very large HDL | 0 |
| Large HDL | $=0$ |
| Medium HDL | - |
| Small HDL | $\square$ |

Lipoprotein particle size


Glycerides and phospholipids


Apolipoproteins


Fatty acids


Fatty acids ratios
DHA (\%)
Linoleic acid (\%)
n-3 FA (\%)
n-6 FA (\%)
PUFA (\%)
MUFA (\%)
SFA (\%)


Glycolysis related metabolites


Amino acids


Branched-chain amino acids


Aromatic amino acids


Ketone bodies


Fluid balance


Inflammation


Open symbols: $\mathrm{P}>=0.003$
$V L D L=$ Very low-density lipoprotein; $L D L=$ low-density lipoprotein; $H D L=h i g h$-density lipoprotein; $C=$ Cholesterol; $F A=$ fatty acids; DHA = docosahexaenoic acid; MUFA=monounsaturated fatty acids; PUFA=polyunsaturated fatty acids; n3=omega; n6=omega 6

Closed symbols: $\mathrm{p}<0.003$; open symbols: $\mathrm{p}>=0.003$
*Minimally adjusted regression: adjusted for age and centre;
**Fully adjusted regression: adjusted for age, centre, PSA at baseline, stage, Gleason

Figure 7.2: Hazard ratios (per 1SD) and their $95 \%$ confidence intervals for the associations between individual biomarkers and all-cause mortality ( 1,864 survivors and 229 deaths) in minimally* and fully adjusted ${ }^{* *}$ Cox regression models.

Lipoprotein subclasses


## Lipoprotein particle size



## Cholesterol



Glycerides and phospholipids


## Fatty acids



Fatty acids ratios


Glycolysis related metabolites


## Amino acids



Branched-chain amino acids

| Isoleucine |  |
| :--- | :--- |
| Leucine |  |
| Valine |  |

Aromatic amino acids
Phenylalanine
Tyrosine


B-hydroxybutyrate
Fluid balance

$\rightarrow$ Death Fully Adjusted
$V L D L=V e r y$ low-density lipoprotein; $L D L=$ low-density lipoprotein; $H D L=$ high-density lipoprotein; $C=$ Cholesterol; $F A=$ fatty acids; DHA = docosahexaenoic acid; MUFA=monounsaturated fatty acids; PUFA=polyunsaturated fatty acids; n3=omega; n6=omega 6

Closed symbols: $\mathrm{p}<0.003$; open symbols: $\mathrm{p}>=0.003$
*Minimally adjusted regression: adjusted for age and centre;
**Fully adjusted regression: adjusted for age, centre, PSA at baseline, stage, Gleason

### 7.4.3. Multi-metabolite risk scores

In PCa cases the correlation coefficient between the PCa mRS and all-cause mRS was $58 \%$ and the correlation between the individual metabolites included in each score ranged from 0 to 85\% (Supplementary Table D4, Appendix D).

### 7.4.3.1 Associations of multi-metabolite risk scores with Prostate

 Cancer and all-cause mortalityI ran minimally and fully adjusted Cox regression models in PCa cases in ProtecT to investigate the associations between the metabolite risk scores (PCa mRS and all-cause mRS ) and subsequent PCa-specific and all-cause mortality, respectively (Figures 7.3A and 7.3B). The PCa mRS showed little evidence of association with PCa specific mortality in the minimally adjusted model (HR = 1.16; 95\% CI: $0.83-1.61$ ) or the fully adjusted model (HR=1.32; 95\% CI: 0.91-1.90) (Figure 7.3 and Supplementary Table D3, Appendix D). The PCa mRS model was positively associated with all-cause mortality, and the effect size was similar to that for PCa specific death, but with smaller confidence intervals (HR=1.17, 95\% CI:1.02-1.33). There was little evidence of an association between all-cause mRS and PCa mortality in the minimally adjusted model $(\mathrm{HR}=0.95$, $95 \% \mathrm{CI}: 0.67-1.40$ ) or the fully adjusted model (HR $=1.03,95 \% \mathrm{CI}$ : $0.70-1.51$ ) (Figure 7.3 and Supplementary Table D3, Appendix D). All-cause mRS was positively associated with all-cause mortality in the main model (HR=1.34, $95 \% \mathrm{CI}: 1.19-1.51$ ) and in the opposite direction compared to the effect size observed for PCa mortality.

Since prevalent PCa could affect metabolic profiles in men with PCa, I investigated the associations between PCa mRS and all-cause mRS and all-cause mortality, separately in cases and controls. In the analysis of cases and controls, the associations of the two biomarker scores and all-cause mortality are similar in the controls only and pooled cases and controls analyses (Figure 7.3C). The PCa mRS model showed a similar effect size to all-cause mortality in the minimally adjusted model, in the cases and controls analysis compared to the case only analysis (HR (pooled) $=1.16,95 \% \mathrm{CI}: 1.06-1.27$ ) (, Supplementary Table D6, Appendix D). The all-cause mRS model showed a slightly increased effect size of association with all-cause mortality in cases and controls, compared to the case only analysis, with an HR (pooled) $=1.36$ ( $954 \% \mathrm{CI} 11.25-1.48$ ) (Figure 7.3C and Supplementary Table D6, Appendix D).

### 7.4.3.2 Predictive performance of metabolomic-risk scores and

## Prostate Cancer and all-cause mortality

I investigated the performance of the risk scores (PCa and all-cause mRS) at discriminating PCa and all-cause mortality using positive predictive values (PPV) and negative predictive values (NPV), accuracy and Cohen's Kappa metrics.

To evaluate the ability of the PCa mRS and all-cause mRS risk scores to predict PCa and all-cause mortality, I used ROCs and calculated AUROCs in all men. In addition, I included the ROC curve generated by using Gleason score as sole predictor to compare the performance of the two risk scores against Gleason score (Figures 7.4A and 7.4C). Figures 7.4B and 7.4D present the performance of both metabolomic risk scores and Gleason model when the dataset was restricted to localised PCa only (stage T1/T2 or PSA<20 and Gleason score <8) to assess whether the risk scores better predict localised disease. To investigate how the models' predicted risk matched the observed outcome proportion, I constructed calibration plots (Figure 7.5 A, B, C and D).

Both metabolomic risk-scores showed modest performance at discriminating all-cause and PCa mortality, with similar results when restricting the data to men with localised disease only. All-cause mRS performed slightly better at predicting all-cause mortality compared to PCa mRS, while PCa mRS performed somewhat better than all-cause mRS at predicting PCa mortality. However, none of the metabolite risk scores performed substantially better, with AUROCs with overlapping confidence intervals. Calibration was good for both metabolomic risk scores when comparing to the observed outcomes (PCa and all-cause mortality). Overall, the highest PPV and accuracy were observed for all-cause mRS (PPV=16.6\%; 95\%CI:13.7-19.9\%; accuracy=70.1\%; 95\%CI: 68.1\%-72.1\%) and the highest NPV was observed for PCa mRS (NPV=98.8\%; 95\%CI: 98.1\%-99.3\%). The highest AUROC was seen for all-cause mRS at discriminating all-cause mortality in men with localised disease (AUROC $=0.61 ; 95 \% \mathrm{CI}: 0.56-0.65$ ).
7.4.3.2.1. Prostate Cancer mortality metabolomic risk score model, Prostate

## Cancer and all-cause mortality in Prosate Cancer cases

For PCa mortality, the PCa mRS had a PPV and NPV of $2.4 \%$ ( $95 \%$ CI: $1.43 \%-3.76 \%$ ) and $98.8 \%$ ( $95 \%$ CI: $98.1 \%-99.3 \%$ ), respectively. This means that only $2.4 \%$ of men who were predicted to have PCa death using the PCa mRS truly experienced PCa death (observed). $98.8 \%$ of men who were predicted to not experience PCa death using the PCa mRS, did not truly experience PCa death (observed). PCa mRS had Kappa of 0.015 ( $95 \% \mathrm{CI}$ : -0.041-0.071) suggesting that there was little evidence of agreement between the predicted and observed categories. PCa mRS had accuracy of $64.2 \%$ ( $95 \% \mathrm{CI}$ : $62.1 \%$ $66.3 \%$ ), therefore correctly assigning PCa mortality status in $64.2 \%$ of participants (Table 7.2). In the ROC analysis, the PCa mRS risk score showed limited ability at predicting PCa mortality, with an AUROC $=0.57$ ( $95 \%$ CI: $0.47-0.66$ ), but with a confidence interval that did not exclude chance performance (Figure 7.4). When restricting the analysis to participants with localised PCa only, the model performance was similar (AUROC $=0.60 \%, 95 \%$ CI: $0.47-0.72$ ). The calibration was good for PCa mRS when compared to the observed PCa death (Figure 7.6 A).

PCa mRs performed similarly in showing limited ability to predict all-cause mortality in all PCa cases, with an AUROC $=0.57$ ( $95 \%$ CI $0.53-0.61$ ) and similar model discrimination statistics (Table 7.2). In the analysis of localised PCa, the performance of the model at predicting all-cause mortality was similar to that for the all-case analysis, with an AUROC $=0.56$ ( $95 \%$ CI: 0.52-0.61).

As a positive control, I also considered the discriminative capacity of Gleason score, which is expected to track closely with PCa progression. Here, I observed strong discriminative performance, with an AUROC $=0.82 \%$ ( $95 \%$ CI: $0.75-0.89$ ) for PCa mortality and modest performance, and AUROC $=0.56 \%$ ( $95 \% \mathrm{CI}: 0.52 .2-0.59$ ) for allcause mortality. For PCa mortality, the Gleason score had a PPV of 0.05 ( $95 \% \mathrm{CI}: 0.03$ 0.07 ), NPP of $1.0(0.99,1.00)$, an accuracy of 0.73 ( $95 \%$ CI: $0.71-0.75$ ) and a Kappa of 0.06 . The calibration was good for PCa mRS when compared to the observed all-cause mortality (Figure 7-6 B).

### 7.4.3.2.2. All-cause mortality multi-metabolite risk score, Prostate Cancer and all-cause mortality in Prostate Cancer cases

All-cause mRS showed a PPV of 1.23\% (95\%CI: 0.62\% - 2.2\%) and NPV of 98.1\% (95\%CI:97.1\% - 98.8\%), an accuracy of 56.5\% (95\%CI:54.3\% - 58.6\%) and Kappa of -0.008
(95\%CI: -0.057-0.04) for PCa death (Table 7.2). In the ROC analysis, all-cause mRS performed worse than the PCa mRS at predicting PCa mortality, with AUROC $=0.51$ ( $95 \%$ CI: $0.42-0.61$ ) (Figure 7.4). When restricting the analysis to localised cases only, the all-cause mRS still performed poorly at predicting PCa mortality with an AUROC= 0.55 ( $95 \%$ CI: $0.42-0.67$ ). The calibration was good for all-cause mRS compared with the observed PCa death (Figure 7.6 C).

The all-cause mRS model performed similarly at predicting all-cause mortality compared to PCa mortality, with an AUROC $=0.60(95 \% \mathrm{CI}: 0.56-0.64)$ and comparable model discrimination statistics (Table 7.2). In the analysis of cases with localised disease only, the model performed similarly to the all-case analysis, with an $\mathrm{AUROC}=0.61$ ( $95 \%$ CI: $0.56-0.65$ ). The calibration was good for all-cause mRS compared with the observed all-cause mortality (Figure 7.6 D).
7.4.3.2.3. Prostate Cancer mortality metabolomic risk score model and all-cause mortality multi-metabolite risk score in Prostate Cancer cases and controls The PCa mRS and all-cause mRS models performed similarly at discriminating all-cause mortality, in the pooled cases and controls analysis compared to the case only analysis, with an $\mathrm{AUROC}=0.62$ ( $95 \% \mathrm{CI}: 0.58-0.66$ ) for the all-cause mRS score and an AUROC $=0.56(95 \%$ CI0.51-0.60) for the PCa mRS (Figure 7.6).

Table 7.2: Model discrimination statistics for the Prostate Cancer mortality metabolomic risk score model and all-cause mortality multi-metabolite risk score for Prostate Cancer and all-cause mortality in the ProtecT cohort.

|  | PCa mRS |  | All-cause mRS |  |
| :--- | :--- | :--- | :--- | :--- |
| Measure | PCa death | All-cause <br> death | PCa death | All-cause death |
| PPV |  | $14.7 \%$ |  |  |
|  | $2.4 \%(1.43 \%-$ | $(12.4 \%-$ | $1.23 \%(0.62 \%-$ | $16.6 \%(13.7 \%-$ |
|  | $3.76 \%)$ | $17.3 \%)$ | $2.2 \%)$ | $19.9 \%)$ |
| NPV | $98.8 \%$ | $91.6 \%$ |  |  |
|  | $(98.1 \%-$ | $(89.9 \%-$ | $98.1 \%(97.1 \%-$ | $91.3 \%(89.8 \%-$ |
|  | $99.3 \%)$ | $93.1 \%)$ | $98.8 \%)$ | $92.7 \%)$ |
| Accuracy | $64.2 \%$ | $60.6 \%$ |  |  |
|  | $(62.1 \%-$ | $(58.5 \%-$ | $56.5 \%(54.3 \%-$ | $70.1 \%(68.1 \%-$ |
|  | $66.3 \%)$ | $62.7 \%)$ | $58.6 \%)$ | $72.1 \%)$ |
| Cohen's | 0.015 | 0.071 | -0.008 | 0.098 |
| Kappa | $(-0.041-0.071)$ | $(0.022-0.12)$ | $(-0.057-0.04)$ | $(0.039-0.16)$ |

PCa mRS = Prostate Cancer mortality multi-metabolite risk score ; All-cause mRS = all-cause mortality multi-metabolite risk score; $P C a=$ Prostate Cancer; $P P V=$ Positive predictive value; $N P V=$ Negative predictive value

Figure 7.3: Hazard ratios (per 1 standard deviation) and $95 \%$ confidence intervals for the association of Prostate Cancer mortality metabolomic risk score model and all-cause mortality multi-metabolite risk score in cases with Prostate Cancer specific (A) and all-cause mortality (B) in the minimally* and fully adjusted** models and in controls and pooled participants (C).


PCa $m R S=$ Prostate cancer mortality multi-metabolite risk score ; All-cause $m R S=$ all-cause mortality multi-metabolite risk score
Hazard ratios (per 1SD) for the association between the PCa mRS and all-cause mRS and:
(A) PCa specific mortality (2,042 survivors, 34 deaths) in minimally* and fully adjusted** Cox regression models, in all eligible to be randomised participants (cases).
(B)All-cause mortality ( 1,846 survivors and 229 deaths) in minimally* and fully adjusted** Cox regression models, in all eligible to be randomised participants (cases).
(C) All-cause mortality in minimally* adjusted Cox regression models, in controls ( 1,960 survivors and 207 deaths) and pooled cases and controls (3,824 survivors and 436 deaths)

Figure 7.4: Receiver operating curves for the Prostate Cancer mortality metabolomic risk score model and all-cause mortality multi-metabolite risk score at discriminating PCa-specific and all-cause mortality in all eligible to be randomised

participants and participants with localised disease only.

ROC curves for the PCa mRS, all-cause mRS, and Gleason scores at discriminating:
(A) PCa specific deaths ( $\mathrm{N}=34$ ) from survivors among all eligible to be randomised ProtecT participants ( $\mathrm{N}=2,076$ ).
(B) PCa specific deaths $(\mathrm{N}=20)$ from survivors among eligible to be randomised ProtecT participants with localised disease only ( $\mathrm{N}=1,962$ ).
(C) All-cause mortality $(\mathrm{N}=229)$ from survivors among all eligible to be randomised ProtecT participants $(\mathrm{N}=2,076)$.
(D) All-cause mortality $(\mathrm{N}=202)$ from survivors among all eligible to be randomised ProtecT participants with localised disease only ( $\mathrm{N}=1,962$ ).

Figure 7.5: Calibration slopes for the Prostate Cancer mortality metabolomic risk score model and all-cause mortality multi-metabolite risk score models and Prostate Canceer and all-cause death.

(A) Calibration slope for the PCa mRS predicted probabilities and the observed PCa death outcomes.
(B) Calibration slope for the PCa mRS predicted probabilities and the observed all-cause death outcomes.
(C) Calibration slope for the all-cause mRS predicted probabilities and the observed PCa death outcomes.
(D) Calibration slope for the all-cause mRS predicted probabilities and the observed all-cause death outcomes.

Figure 7.6: Receiver operating curves for the Prostate Cancer mortality metabolomic risk score model and all-cause mortality multi-metabolite risk score at discriminating all-cause mortality in all Prostate Testing for Cancer and Treatment controls $(\mathrm{N}=2,150)$ and Prostate Testing for Cancer and Treatment cases and controls ( $\mathrm{N}=4,225$ ).


AUC: area under the curve
(A) ROC Curves for the PCa mRS and all-cause mRS biomarker models, at discriminating all-cause mortality ( $\mathrm{N}=205$ ) in controls in the CAP study $(\mathrm{N}=2,150)$.
(B) ROC Curves for the PCa mRS and all-cause mRS biomarker models, at discriminating all-cause mortality ( $\mathrm{N}=434$ ) in pooled participants (cases and controls) in the CAP study ( $\mathrm{N}=4,225$ ).

### 7.5. Discussion

### 7.5.1. Summary of findings

In this Chapter, I investigated the predictive utility of serum NMR metabolite levels and two metabolite risk scores previously developed in the general population to predict PCaspecific and all-cause mortality. The population of men used in this study originated though a study of population-wide PSA testing for PCa (the ProtecT cohort). I found evidence of
association of NMR metabolite subclasses with PCa and all-cause mortality in this cohort of PSA detected PCa participants. In the same cohort, I found evidence for association between PCa mRS and and all-cause mRS and all-cause mortality. However, I did not find robust evidence that the two metabolite risk scores predicted PCa or all-cause mortality. When comparing the associations of individual level metabolites and the two metabolite risk scores, in participants with and without PCa, the findings were similar.

I found some evidence that individual metabolites levels, such as very low-density lipoprotein and HDL measures, branch-chain amino acids, ketone bodies and glycoprotein acetyls were associated with PCa mortality. However, only leucine and isoleucine were positively associated with PCa mortality.

All-cause mortality showed similar patterns of association in PCa cases to those between metabolites and PCa mortality, with strong positive associations for LDL particle size, acetate, acetoacetate, glucose, ratio of monounsaturated fatty acids to total fatty acids and saturated fatty acids to total fatty acids, and phospholipids in small HDL particles. Negative associations with all-cause mortality were observed for cholesterol esters in small HDL, degree of unsaturation, ratios of omega-3, omega-6, linoleic acid, polyunsaturated fatty acids to total fatty acids. To explore the potential effect that prevalent PCa may have on individual metabolites, I used the control population in the ProtecT. The associations between metabolites and all-cause mortality in PCa-free men generally followed the same direction as cases, but with smaller effect sizes.

Both multi-metabolite risk scores had a poor prediction performance, with AUROC point estimates (AUROCs $\sim 0.6$ ) having confidence intervals that included chance prediction. This suggests that as constituted, the scores could not accurately discriminate PCa-specific mortality and all-cause mortality in this cohort of participants. In contrast, Gleason score, a measure known to robustly track PCa aggressiveness and which had been included as a positive control in our analysis, showed improved prediction performance compared to the two multi-metabolite risk scores. The predictive performance of the Gleason score in localised only cases was marginally lower than in the all-cases pooled analyses, which is to be expected since there were fewer events in the localised only cohort. Comparing the performance of the multi-metabolite risk scores and the clinically validated predictor of PCa aggressiveness and progression (Gleason score) suggests that the multi-metabolite risk score is a less suitable candidate as predictor of PCa mortality.

Possible explanations for why the PCa mRS model may not have discriminated PCa mortality include that this score was trained in a general population setting, as a prediagnostic tool to predict PCa-specific mortality, and included all cases of PCa, while the current study evaluated its performance at predicting PCa specific death, in locally and locally advanced PCa cases, diagnosed through PSA testing. Further, the PCa mRS performance in ProtecT could be due to the heterogeneity of PCa aggressiveness, since the dataset included men with localised and locally advanced PCa only, with few events ( $\mathrm{N}=34$ ) whose metabolic trait patterns may substantially differ to those on which the score was developed, which may have contained more advanced disease. The mortality rates in the ProtecT trial were much lower than in the UKBB dataset on which the score was developed, $2 \%$ vs $10 \%$, respectively. This could indicate that the model was trained to predict more aggressive cancers rather than localised and locally advanced cases. Interestingly, the PCa mRS model performs slightly better in localised cases than the locally advanced cases who would have been more advanced at diagnosis, so differences in predictors of outcome may be expected.

The all-cause mRS score was developed in the general population, aiming to predict allcause mortality rather than PCa specific death, both of which could explain the lack of replication I found in the current dataset. In addition, the all-cause mRS score was derived in 12 cohorts of both male and female participants and of a wide age range (18-109), while the ProtecT study only included males with ages between 50 and 71, which could constitute another reason for the lack of replication. The similar performance of the PCa mRS and allcause mRS could be due to the fact that the two scores were highly correlated. This probably led to the PCa mRS picking up some of the all-cause signal of the all-cause mRS, with the extra noise slightly dropping the performance of PCa mRS.

### 7.5.2. Findings in context

The evidence I found suggests that metabolites are associated with both PCa and all-cause mortality. For example, I found some amino acids and ketone bodies were associated with PCa and all-cause mortality, which was expected since studies that assessed the link between PCa aggressiveness, defined using Gleason and TNM stage, found that amino acids, such as sarcosine, valine, leucine were associated with PCa aggressiveness
$(19,35,36)$. One study found that three metabolites (cystathionine, cysteine, homocysteine) were associated with biochemical recurrence and each metabolite had a good AUROC (0.71-
0.80 ) at predicting the event (176). However, the model was not externally validated, the sample size was small ( $n=58$ participants with serum samples) and the follow-up period was short (2 years) (176). Although I found metabolites to be associated with all-cause and PCa mortality, further work was required to ascertain whether they are also predictive of these outcomes in men with already diagnosed PCa.

A large number of predictive models for PCa progression have been introduced, with a wide range of predictors, form anthropometric to imaging factors, however many of these have not been externally validated and most have not been taken up in clinical practice (406). One model using clinical information on comorbidities which showed great performance at predicting 10-year life expectancy, with an AUROC $=0.82$ was only internally validated, and used all-cause mortality rather than PCa- specific death as outcome (407). Another model, using biopsy grade, PSA, clinical T-stage and age, which used PCa specific mortality was found to have a good performance power, with an AUROC $=0.78$, however this has only been internally validated and requires invasive procedures (408). This study is one of the first studies to use a metabolomic predictive model, developed in a large cohort of general population and externally validated, and test its predictive performance for use in clinical settings.

Although the performance observed in the ProtecT study population was modest, the findings are important for the field. First, it shows that published performance even from external validation populations may not be reproducible in other populations if the disease characteristics are different to those in which the model was trained and tested. This is particularly important for localised PCa, which is still a poorly characterised disease when it comes to predicting survival and differentiating between cases that are more likely to advance compared to indolent cases $(409,410)$. Second, it is important consider the settings in which risk stratification approaches can most appropriately be applied. For example, whilst I found the score had limited performance in a clinical setting of already diagnosed localised PCa cases, the predictive model can still be a great asset in clinical settings, by using it as a of pre-diagnostic predictive tool of PCa lethality. This could help identify PCa patients who are most at risk of developing lethal PCa, whilst avoiding the over diagnosing of cases that are likely to not have clinically significant disease.

More generally, while current NMR metabolomics technology can provide a quick, highly reproducible, and cost-effective way of measuring proxies of metabolic footprints, in order to use these in prediction of PCa mortality in clinical practice, clearly defined and precise
study populations and must be used when aiming to translate from general population settings. Future studies undertaking prediction work in general populations settings should try and predict mortality in patients already diagnosed with PCa, in order to provide a useful tool for disease management in clinical settings. In addition, developing predictive models by subgrouping PCa cases on staging (localised vs advanced) may be beneficial to prediction of PCa-mortality since localised cases tend to have very different clinical characteristics and survival patterns to more advanced cases.

### 7.5.3. Strengths and weaknesses

This study had several strengths. First, it used data from the ProtecT cohort, a large PSA screening-based cohort where follow-up is very good and death outcomes are linked through national death registers. Second, the NMR platform which was used to identify and quantify the metabolite concentrations used in the predictive models is highly reproducible, making the models easy to replicate in other studies or in practice. Third, the predictive models evaluated were developed in large and generalisable datasets which have previously been extensively analysed, thus the findings are likely to be generalisable to many other populations. The main weakness of the current study was the small number of events, particularly PCa deaths, which lead to low power in ascertaining the ability of the models to robustly predict the events of interest. Also, while the metabolomic predictive models were developed in the general population, the current assessment was performed in a screenbased cohort, thus potentially using two noncomparable groups. Lastly, due to the lack of data availability of staging and cancer aggressiveness in the UKBB, when developing the predictive model, the researchers were unable to distinguish between localised and advanced PCa cases, and so the predictive model may only be suitable for cohorts similar to UKBB in terms of cancer aggressiveness and staging.

### 7.6. Chapter summary

In this chapter I investigated the link between individual NMR metabolites and PCa and allcause mortality and evaluated the performance of two multi-metabolomic risk scores at predicting PCa and all-cause mortality in a screen-detected cohort of men. I found that NMR metabolites were associated with PCa mortality, in particular glycine and glycoprotein acetyls which were positively and negatively associated with PCa mortality, respectively.

All-cause mortality was positively associations with cholesterol in LDL, degree of unsaturation and fatty acid ratios, and negatively associated for glucose and acetate. There was evidence of association for two metabolomic risk scores trained in general population with increased PCa and all-cause mortality; however, the two metabolomic risk scores did not predict all-cause or PCa specific mortality in a clinical setting of screen detected localised or locally advanced PCs.

## Chapter 8. Discussion

### 8.1. Chapter introduction

In this Chapter, I summarise the main findings of this thesis, its strengths and limitations and the future work which could be undertaken.

### 8.2. Summary of key findings

This thesis aimed to use circulating metabolomic data measured using NMR to: i) better understand mechanisms of PCa progression to identify new potential therapeutic and behavioural targets for interventions; and ii) identify novel clinical prediction biomarkers which could, in clinical practice, distinguish indolent from lethal disease. In the observational analysis in Chapter 5, I found only weak evidence that some measures of lipoproteins, cholesterol, triglycerides, fatty acids, leucine, citrate and glycoprotein acetyls may be observationally associated with PCa progression (clinical progression and metastases or PCa death) (Table 8.1).

The Mendelian randomisation (MR) analysis in the second part of Chapter 5 found causal evidence that PCa mortality was increased with higher levels of total, free and esterified cholesterol and some measures of lipoprotein subclasses (very low-density lipoprotein (VLDL), intermediate-density lipoprotein (IDL) and low-density lipoprotein (LDL)), sphingomyelins, apolipoprotein B, total omega-3 fatty acids, docosahexaenoic acid, and valine. It also found evidence of decreased PCa mortality with increasing levels of histidine, phospholipids to total lipids ratio in small and large HDL cholesterol, ratio of triglyceride to total lipids in very small VLDL and ratio of omega-6 to omega-3.

In Chapter 6, I found weak evidence that a randomised intervention to increase fruit and vegetable, and decreased dairy consumption, altered the lipid and glycolysis metabolic profiles of men with PCa. Within the same trial, I also found that a lycopene supplementation intervention did not increase serum lycopene and there was only weak evidence of the intervention altering levels of alanine. In the physical activity intervention of the trial, I found weak evidence to support that a brisk walking intervention altered the lipid profiles of men with PCa. The MR analysis found some causal evidence to support a link
between the metabolites decreased by the dietary advice and brisk walking interventions (pyruvate and cholesterol measures, respectively) and increased PCa death.

In Chapter 7, I found strong evidence that acetate was positively, and glycine was negatively associated with PCa mortality. Also, there was strong evidence for positive associations of cholesterol in LDL and degree of unsaturation and fatty acid ratios, and negative associations for glucose and acetate, in relation to all-cause mortality. Also in Chapter 7, I found that two multi-metabolite risk scores developed in the general population to predict PCa-specific and an all-cause mortality did not produce the same level of performance when predicting PCa-specific and all-cause mortality in a clinical setting amongst screen-detected PCa cases.

When the evidence is taken together, my findings suggest that metabolites may be involved in the progression of PCa. These metabolites could serve as potential interventional targets, whether behavioural or therapeutic.

However, studies that are adequately powered, have a large follow-up time and precise progression definitions are required to further investigate the link between metabolites and PCa progression. My findings do not exclude the value of multi-metabolite risk scores in predicting PCa progression, rather they suggest that risk scores available to date which have been developed in general populations are not appropriate for use in clinical settings for screen-detected PCa. The metabolite risk scores investigated in this thesis should be further investigated in other general population cohorts to confirm their predictive performance in those settings. Multi-metabolite risk scores that are designed to be used for clinical use should be developed in specific and generalisable clinical settings.

Table 8.1: Summary of the main findings of this thesis

| Research question | N | Main Findings | Chapter |
| :---: | :---: | :---: | :---: |
| Are circulating metabolite levels associated with risk of clinical progression in men with localised PCa ? | 1,827 participants <br> of which 201 <br> clinically <br> progressed | - Suggestive evidence ( $p>0.003$ ) for clinical progression being positively associated with measures of very low-density lipoprotein (VLDL) and triglycerides and inversely associated with non-triglyceride measures of high-density lipoprotein (HDL) (triglycerides in medium HDL) (HR $=1.14,95 \%$ CI: $1.00-1.30 ; p=0.05)$ <br> - Suggestive evidence for citrate being inversely associated with clinical progression ( $\mathrm{HR}=0.87$, $95 \%$ CI: $0.75-1.00, p=0.05$ ) <br> - Suggestive evidence for glycoprotein acetyls which showed a positive association with clinical progression (HR=1.18; 95\% CL:1.03-1.35; $p=0.01$ ) | 5 |
| Are there associations between circulating metabolite levels and risk of developing metastases or PCa-specific mortality in men with localised disease? | 1,827 participants of which 64 progressed to metastases or PCa death | - Suggestive evidence for medium, very large and extremely large VLDL, some triglyceride measures and particle diameter for low-density lipoprotein (LDL) (concentration of chylomicrons and extremely large VLDL particle: $\mathrm{HR}=1.26 ; 95 \% \mathrm{CI}: 1.04-1.53 ; \mathrm{p}=0.02$ ) with increased metastases or PCa mortality. There was suggestive evidence for inverse associations of metastases and PCa specific mortality with small HDL measures, small, medium, and large LDL, and multiple intermediate density lipoprotein measures. <br> - Suggestive evidence for the ratio of saturated fatty acid to total fatty acid with increased risk of metastases or PCa death (HR=1.20;95\%CI: 0.93-1.55, $\mathrm{p}=0.16$ ) and degree of unsaturation with decreased risk ( $\mathrm{HR}=0.82,95 \% \mathrm{CI}: 0.64-1.06, \mathrm{p}=0.13$ ) <br> - Suggestive evidence for leucine with increased risk of metastases or PCa death (HR=1.19; $95 \%$ CI: $0.94-1.50 ; p=0.14$ ) <br> - Suggestive evidence for glycoprotein acetyls and increased risk of metastases or PCa death ( $\mathrm{HR}=1.31, \mathrm{CI}: 1.04-1.64, \mathrm{p}=0.02$ ) | 5 |


| Research question | N | Main Findings | Chapter |
| :---: | :---: | :---: | :---: |
| Are there causal links between circulating metabolite levels and PCa-specific mortality? | 75,672 participants of which 7,914 died of PCa | - Evidence of association ( $\mathrm{p}<0.05$ ) of total, free and esterified cholesterol, some measures of IDL, LDL (in total, small, medium and large particles) and VLDL (in small, medium and extremely large particles), sphingomyelins, apolipoprotein B, total omega-3 fatty acids, docosahexaenoic acid, and valine with increased PCa mortality. Additionally, weaker evidence ( $\mathrm{p} \geq 0.05$ with large effect size) for the concentration of branched-chain amino acids (leucine, isoleucine and valine), lactate, beta-hydroxybutyrate and alanine with increased PCa mortality. <br> - Evidence of causal association of histidine, phospholipids to total lipids ratio in small and large HDL cholesterol, ratio of triglyceride to total lipids in very small VLDL and ratio of omega- 6 to omega-3 with decreased PCa mortality. | 5 |
| Does a dietary intervention comprising either i) lycopene supplementation; or ii) dietary advice to increase fruit and vegetable and reduce dairy milk consumption; or iii) a physical activity intervention of brisk walking alter the metabolome of men with PCa ? | 74 participants | - i) Weak evidence that lycopene intervention reduced alanine levels. <br> - ii) Weak evidence that the dietary advice intervention decreased levels of pyruvate, lactate, concentration of small HDL and ratio of saturated to total fatty acids, and increased levels of linoleic and omega-6 fatty acids to total fatty acids. <br> - iii) Weak evidence that the physical activity intervention increased ratios of omega-6, linoleic acid and polyunsaturated fatty acids to total fatty acids and decreased multiple measures of VLDL. | 6 |
| What is the causal effect of the metabolites altered by the above dietary and lifestyle interventions (i), ii) and iii)) on PCa mortality? | 75,672 participants of which 7,914 died of PCa | - i) Weak causal evidence of alanine being linked to reduced PCa death <br> - ii) Weak causal evidence of pyruvate, concentration of small HDL, ratio of saturated fatty acids to total fatty acids to decrease and ratio of omega-6 fatty acids and linoleic acid to total fatty acids and acetate to increase PCa death. <br> - iii) Suggestive causal evidence of association of cholesterol esters in medium VLDL, and phospholipids in chylomicrons and extremely large VLDL with increased PCa death. | 6 |


| Research question | N | Main Findings | Chapter |
| :---: | :---: | :---: | :---: |
| Are metabolites associated with PCa specific and allcause mortality in a clinical setting? | 2,093 of which 229 <br> all-cause deaths <br> and 34 PCa deaths | - Strong evidence of positive association of leucine ( $\mathrm{HR}=1.53,95 \% \mathrm{CI}: 1.16-0.2 .03, \mathrm{p}=0.002$ ) and isoleucine ( $\mathrm{HR}=1.56,95 \% \mathrm{CI}: 1.18-2.01, \mathrm{p}=0.002$ ) with increased PCa mortality. <br> - Strong associations of acetate, acetoacetate, glucose, ratio of monounsaturated fatty acids to total fatty acids and saturated fatty acids to total fatty acids, LDL particle size, phospholipids in small HDL particles, with increased risk of all-cause mortality. <br> - Strong associations of cholesterol esters in small HDL, degree of unsaturation, ratios of omega3 , omega-6, linoleic acid, polyunsaturated fatty acids to total fatty acids with reduced risk of allcause mortality. | 7 |
| Does a risk score that predicts PCa in a general population also predict PCa or all-cause mortality in a clinical setting? | 2,093 of which 229 <br> all-cause deaths and 34 PCa deaths | - Both PCa and all-cause multi-metabolite risk scores developed in the general population, were associated with all-cause mortality, but not with PCa mortality. <br> - The PCa multi-metabolite risk score had a poor predictive performance for PCa-specific and allcause mortality in a clinical setting. | 7 |
| Does a risk score that predicts all-cause mortality also predict PCa death or all-cause mortality in a clinical setting? | 2,093 of which 229 <br> all-cause deaths and 34 PCa deaths | - The all-cause multi-metabolite risk score had a poor predictive performance for PCa-specific and all-cause mortality in a clinical setting. | 7 |

$P C a=$ Prostate Cancer; VLDL= very low-density lipoprotein; $H D L=$ high-density lipoprotein; $L D L=$ low-density lipoprotein

### 8.3. Strengths and weaknesses of this research

### 8.3.1. Data sources

One of the strengths of this thesis is the use of multiple and varied data sources such as large consortia (PRACTICAL), UKBB and clinical trial data (PrEvENT and ProtecT) to triangulate the evidence. The UKBB and PRACTICAL data sources provided large genetic data on which causal evidence was assessed. The UKBB is the largest GWAS of NMR metabolites and was conducted in a population representative of the general population. The PRACTICAL PCa mortality GWAS is the largest GWAS of PCa mortality and includes data from a wide variety of genetic studies from across the world. Having a large number of PCa deaths on which genomic data is generated is a key strength of this thesis, given PCa is a slow progressing disease with a low mortality rate. Using these large datasets, I was able to conduct an adequately powered two-sample MR study and assess causality between metabolomic data and PCa mortality.

The ProtecT trial, the embedded trial of the CAP trial, is one of the UK's largest RCT of PCa progression, with follow-up data available at 10 years. It also has progression outcomes which are well-defined and assessed by an independent committee. In addition, for participants who were not followed up through the ProtecT trial, data was available by linking the data to the mortality datasets from the Office of National Statistics through the CAP trial, thus minimising the potential of misclassification of non-progressed cases. The PrEvENT feasibility RCT used a factorial design with three dietary and physical activity interventions and had good adherence and low loss to follow up. By using the above the sources together, I was able to i) better characterise the metabolome of men with $\mathrm{PCa}, \mathrm{ii}$ ) assess causality between metabolites and PCa death and iii) leverage statistical power and provide suggestive evidence of the effects of a short interventional RCT (PrEvENT) on long term clinical outcomes (PCa death).

### 8.3.2. Randomised controlled trial design

Although I used the ProtecT trial as a prospective cohort study, the data used was collected as part of a RCT. Using a RCT design has multiple benefits. Firstly, the inclusion criteria are very precise and implemented though standard operating procedures, therefore leading to
the study having an accurately and clearly defined population. Progression outcome indicators were evaluated using predefined and specific definitions. This is particularly important as seen in Chapter 3, given the different progression definitions and proxy measures used in the literature. In the ProtecT RCT, participants were randomised to three types of treatment which were followed throughout the trial. This is an advantage over other non-RCT progression studies where treatment is not randomly allocated, and where there can be variation on the type and delivery of the treatment received. To assess effects of dietary and physical activity interventions on the metabolome of men with PCa I used data from the PrEvENT feasibility RCT using intention to treat analysis, a gold standard analysis in RCTs. In addition, a well-designed and implemented RCT will have confounders randomly distributed across the intervention arms, thus issues of confounding did not constitute a major issue in my analysis of the PrEvENT trial.

### 8.3.3. Mendelian randomisation

As described in Chapter 4, MR is a powerful tool to assess causality in epidemiological research and overcomes issues common in traditional epidemiology such as reverse causation. However, there are also limitations associated with conducting MR analyses. In this thesis, one of the main limitations in my MR analyses was that of horizontal pleiotropy, which has been described in detail throughout the thesis. Invalid instruments can bias the results through horizontal pleiotropy, where they have pleiotropic effects on pathways independent of the exposure being investigated $(287,295)$. I employed sensitivity analyses such as MR Egger, which generally generated similar results to the main estimation method (inverse variance weighted).

However, the tools to investigate pleiotropy are still limited. The field of metabolomics is particularly prone to this type of bias in MR studies given that metabolomics reflects a wide variety of metabolic processes and pathways. One other potential bias in my MR analysis is that of collider bias, which has been described in Chapter 4 as well as within the discussion sections of Chapters 5 and 6 . Collider bias in the context of this thesis refers to the bias generated by metabolites being associated with both the risk of developing PCa and PCa mortality. For each of the metabolites where I found causal evidence of association with PCa mortality, I investigated their association with risk. I presented this as a limitation where a metabolite was associated with both as this could be an artefact of collider bias rather than a true causal link between the metabolite and PCa mortality. This issue warrants further
investigation and is presented later in this Chapter. The PRACTICAL consortium did not have PCa staging information, and thus the causal estimates for each of the metabolites in relation to PCa mortality were for all stages in a cumulative analysis. Since this thesis focused on localised or locally advanced PCa and it could be that the causal estimates observed for the metabolites in relation to PCa mortality do not necessarily apply in the localised PCa population. However, there is no genetic wide association study (GWAS) of localised PCa mortality to date, making the MR analysis impossible to be conducted in localised cases only.

### 8.3.4. Interventional feasibility randomised controlled trial

Using metabolomic data from the PrEvENT trial, allowed me to investigate the effect of each of the interventions (lycopene supplementation, dietary advice and physical activity) on the metabolome of men with PCa. By using a RCT design, I was able to disregard the effects that confounders could have on the link between the interventions and metabolite levels. This is an issue which would be a substantial limitation in a non-interventional study, particularly given the broad effects that would generally be triggered by a change in dietary and lifestyle patterns. However, the PrEvENT trial was a feasibility RCT and thus did not set out to detect changes in metabolites, therefore leading to a lack of power in my study to detect effects of the interventions on the metabolome. The evidence from the PrEvENT trial is suggestive in nature, and larger trials, adequately powered to detect the effects of interventions on the metabolome of men with PCa should be undertaken to provide definitive evidence.

### 8.3.5. Triangulation of evidence

In Chapters 5 and 6, I used MR to complement my findings from a cohort study (using ProtecT trial data) and a feasibility RCT (PrEvENT). The triangulation of evidence represents a strength of my thesis, allowing for evidence from various data sources, using multiple methods (survival analysis, intention-to-treat, and MR) to contribute to the inferences. However, each of these techniques and data sources come with their own limitations and biases as described above. When triangulating the evidence, one of the major limitations is the lack of comparable effects since each method tends to estimate different measures. For
example, the MR technique estimates a lifelong causal effect between the metabolites and PCa mortality, unlike the estimates obtained from the survival analysis of the ProtecT trial data which provide the effect from diagnosis. When using MR to provide evidence on the long-term effects on PCa mortality of the metabolites altered by the PrEvENT trial's interventions had similar issue arises. The metabolites altered by the trial were for a 6 -month period, while the MR analysis provides an estimate for a lifelong of exposure of the metabolites. Therefore, whilst it is informative to establish whether the interventions have a causal effect on PCa mortality, through the metabolites' pathway, we cannot fully estimate the effect. Whilst other methods exist to try and estimate the effect, this was beyond the scope of this thesis.

### 8.3.6. Missing data

Missing data was generally not an issue in this thesis. The exception was the measurement of BMI in the ProtecT trial, which was incomplete for nearly $1 / 3$ of participants. This was mainly due to the fact that weight and height were not collected at the beginning of the study (the pilot study). This aspect, however, did not seem to introduce substantial bias in my analysis.

### 8.3.7. Bias and confounding

### 8.3.7.1 Bias resulting from metabolomic pre-analysis handling

There is suggestive evidence in the literature and in this thesis that glycolytic metabolites may play a role in PCa progression. However, a major limitation in analysing the glycolysis metabolites is sample handling. The glycolysis process starts immediately after the collection of serum samples in the vial (411). This can lead to the glycolysis metabolites being altered as a result of the sample handling rather than a biological reason. This could introduce information or measurement bias, particularly as the length of time the sample is left out would directly impact on the concentration of the glycolysis metabolite concentration. Whilst the Nightingale NMR platform has great reproducibility and low interlaboratory uncertainty, the samples in the ProtecT and PrEvENT study could still be affected by information bias, given that the samples were not consistently frozen at -80 degrees Celsius within three hours of collection which is the period after which measurements can be significantly altered (411). In the ProtecT trial, the samples were stored in a cool box after
collection and centrifugation until arrival at the laboratory where they were frozen, thus allowing enough time for glycolysis to occur. This could introduce bias, for a number of reasons when assessing the glycolysis metabolites. Firstly, the absolute values observed for the glycolytic metabolites in ProtecT may be inaccurate. Secondly, the observed patterns in glycolytic measures could be affected by bias of confounding by centre given there were multiple prostate clinic checks which may have had different processing timings and staffing. Whilst a standard operating procedure was in place for the samples to be processed and frozen as soon as possible and ideally the same day, external factors, such as staffing issues, distances between the prostate check clinic and availability of transport could have led to confounding. Since these aspects were not recorded, it is difficult to ascertain the level of bias that was introduced by the sample handling and shipping However, prostate check clinics had centrifuges and thus the samples were processed on site before shipment, stored in a cool box with ice and eventually stored at -80 degrees Celsius, make it unlikely for other measures other than glycolysis to be affected.

To a lesser extent, the PrEvENT trial could have also had their glycolytic metabolites affected by sample handling. While the PrEvENT trial was based at one location only and the laboratory was located in the same campus as the clinic, no formal procedure was in place to collect information on how long the sample took until it was frozen. The PrEvENT trial could not blind the participants or the healthcare professional in the lifestyle and dietary interventions and did not use a placebo for the lycopene supplementation. This could have introduced performance bias, since it may be that the healthcare professionals may have inconsistently handled the samples, for example rushing the samples of those in the intervention arms compared to controls. This could have induced biased when assessing the associations in the intention to treat analysis of glycolytic metabolites, given that the changes observed could be due to pre-analytical differences.

The stability of the other metabolomic measures, tend to be stable and thus the preanalytical processing techniques used in PrEvENT and ProtecT should not have affected the metabolite measured by much. The most common causes of inconsistencies in metabolite levels are a lower freezing temperature ( -20 instead of -80 degrees Celsius) and the thawing at room temperature (412). However, none of these were concerns in the ProtecT and PrEvENT trials.

### 8.3.7.2 Selection bias

The ProtecT trial, whilst representative of the UK population could still be prone to selection bias. Participants were invited to attend a PCa screening and only $54 \%$ responded. Participants who did not respond or attend the screening could systematically differ in disease incidence and staging, as well as in socio-economic and morbidity characteristics. This could make the finding from this thesis not generalisable to the UK population. In addition, of men who attended the ProtecT screening visit, only $67 \%$ also entered the associated ProMPT study which was responsible for the collection of other samples and questionnaires. While the samples were analysed via the ProtecT study, additional information which was collected via the ProMPT study, such as height, comorbidities, lifestyle and dietary patterns were not available for those men who did not consent to enter the ProMPT study. Thus, the fully adjusted analyses, which adjusted for body mass index (which was missing in participants with no height information), comorbidities, alcohol and smoking could be biased due to selection.

### 8.3.7.3 Confounding

As with any study, confounding may be an issue in this thesis. While in the PrEvENT trial the randomisation should have led to confounders being equally distributed across the intervention arms, there was some evidence of imbalances in confounders across the arms (age, smoking). Adjusted analyses were performed to investigate this, however in the case of smoking there could be some residual confounding. Smoking was measured only as a categorical variable (never, ever or current smoker) and thus the patterns of smoking (frequency or number of cigarettes smoked) may not accurately reflect the effects of smoking between the intervention arms.

Although the ProtecT trial had a randomised study design, this thesis analysed it as a cohort study, and thus was more prone to confounding than when analysed as an RCT in an intention to treat framework. Whilst a wide range of potential confounders were measures and were accounted for in the analyses (age, disease staging, Gleason score, Prostate specific antigen, body mass index, morbidity) these were not exhaustive. For example, morbidity was self-reported and only contained specific groups of comorbidities, rather than an exhaustive list or hospital admission data which could have led to residual confounding even in the fully adjusted analyses. Lifestyle and dietary questionnaires were available in the ProtecT study, but not investigated in this thesis. Unmeasured confounders such as use
of supplements, lifestyle patterns, genetic risk scores may lead to biased estimates in the observational analysis, given that these were not measured in the Protec T trial.

### 8.4. Future work

### 8.4.1. Prognostic biomarker development

As seen in Chapter 7, a metabolomic prognostic biomarker risk score which performed well in the general population, did not replicate well in a clinical setting. A logical extension of the analysis performed in this thesis would be the development of a metabolomic risk score developed in localised PCa cases only. This would allow the risk score to identify metabolomic changes specific to localised PCa, given that as PCa progresses, it is likely to have increasingly different metabolomic profiles. In addition, it may allow for better replication in a clinical setting, given the more specific population definition. Also, by including other 'omics in addition to the metabolomic dimension and developing a multiomics risk score could allow a better replication. As part of my PhD, I set-up and coordinated an epigenetic wide association study of PCa progression in men with localised PCa and Gleason grade lower than $4+3$ in the ProtecT trial. In this population, genetic, epigenetic and metabolomic data will be available which could provide a more comprehensive measure and could reproduce better in multiple settings.

### 8.4.2. Metformin and Prostate Cancer progression

Another project I took part in during my PhD , and which constituted the focus of my research for the first year was the development of a feasibility RCT which aimed to investigate the effects of metformin and physical activity in men with PCa undergoing prostatectomy, radiotherapy and active surveillance (Prostate cancer-Exercise and Metformin Trial). I helped set-up the study, contributed to the ethical application and took ownership of the metabolomic sample collection and processing. Working with the laboratory staff at the North Bristol Trust Southmead I aimed to ensure all samples collected as part of the trial were frozen at -80 Celsius degrees within two hours of blood collection. This was achieved for all samples collected in the trial. The NMR analysis of these samples will provide evidence on the link between glycolysis related metabolites and PCa progression. In addition, it will help understand the extent to which glycolytic metabolites
are affected by glycolysis. In the physical activity intervention group, I lead the delivery and monitoring of the Garmin devices which participants were give as motivational tools. The analysis of the data will allow more details of the exercise performed by the men (duration, pace, type of activity). In addition, it will provide strengths and limitations of using Garmin wristbands as interventional tool.

### 8.5. Concluding remarks

This thesis highlights the potential that circulating metabolomic markers have in PCa progression. The results suggest that circulating metabolites, particularly lipid and amino acid measures, may be linked to PCa mortality, thus opening an avenue for progression biomarker development and therapeutic targets. I also found that lifestyle and dietary interventions in men with PCa may alter metabolites which may be causally linked to PCa death. Lastly, this thesis found that metabolomic risk scores developed in the general population do not replicate in clinical settings in PSA detected localised PCa. The use of circulating metabolomic biomarkers should be further investigated to expand the current understanding of PCa progression and potential therapeutic targets.

### 8.6. References

1. Pernar CH, Ebot EM, Wilson KM, Mucci LA. The Epidemiology of Prostate Cancer. Cold Spring Harb Perspect Med TA - TT -. 2018;8(12):a030361.
2. Prostate Cancer incidence and survival statistics: Cancer Research UK (CRUK) [Internet]. 2017. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer/incidence
3. Hugosson J, Carlsson S, Aus G, Bergdahl S, Khatami A, Lodding P, et al. Mortality results from the Göteborg randomised population-based prostate-cancer screening trial. Lancet Oncol [Internet]. 2010;11(8):725-32. Available from:
https://www.sciencedirect.com/science/article/pii/S1470204510701467
4. Schröder FH, Hugosson J, Roobol MJ, Tammela TLJ, Zappa M, Nelen V, et al. Screening and prostate cancer mortality: results of the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. Lancet [Internet]. 2014;384(9959):2027-35. Available from:
https://www.sciencedirect.com/science/article/pii/S0140673614605250
5. Hamdy FC, Donovan JL, Lane JA, Mason M, Metcalfe C, Holding P, et al. 10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer. N Engl J Med [Internet]. 2016 Sep 14;375(15):1415-24. Available from: https://doi.org/10.1056/NEJMoa1606220
6. Woo S, Suh CH, Kim SY, Cho JY, Kim SH. Diagnostic Performance of Prostate Imaging Reporting and Data System Version 2 for Detection of Prostate Cancer: A Systematic Review and Diagnostic Meta-analysis. Eur Urol [Internet]. 2017;72(2):17788. Available from:
https://www.sciencedirect.com/science/article/pii/S0302283817300672
7. Litwin MS, Tan H-J. The Diagnosis and Treatment of Prostate Cancer: A Review. JAMA [Internet]. 2017 Jun 27;317(24):2532-42. Available from:
https://doi.org/10.1001/jama.2017.7248
8. Bjurlin MA, Carter HB, Schellhammer P, Cookson MS, Gomella LG, Troyer D, et al. Optimization of Initial Prostate Biopsy in Clinical Practice: Sampling, Labeling and Specimen Processing. J Urol [Internet]. 2013;189(6):2039-46. Available from: https://www.sciencedirect.com/science/article/pii/S0022534713003480
9. Nagana Gowda GA, Raftery D. Biomarker Discovery and Translation in Metabolomics. Curr Metabolomics [Internet]. 2013;1(3):227-40. Available from: https:// pubmed.ncbi.nlm.nih.gov/27134822
10. Jacob M, Lopata AL, Dasouki M, Abdel Rahman AM. Metabolomics toward personalized medicine. Mass Spectrom Rev [Internet]. 2019 May 1;38(3):221-38. Available from: https://doi.org/10.1002/mas. 21548
11. Alonso A, Marsal S, Julià A. Analytical Methods in Untargeted Metabolomics: State of the Art in 2015 [Internet]. Vol. 3, Frontiers in Bioengineering and Biotechnology . 2015. p. 23. Available from:
https://www.frontiersin.org/article/10.3389/fbioe.2015.00023
12. Fearnley LG, Inouye M. Metabolomics in epidemiology: from metabolite concentrations to integrative reaction networks. Int J Epidemiol [Internet]. 2016 Apr 26;45(5):1319-28. Available from: https://doi.org/10.1093/ije/dyw046
13. Wang X, Chen S, Jia W. Metabolomics in Cancer Biomarker Research. Curr Pharmacol Reports [Internet]. 2016;2(6):293-8. Available from: https:/ / doi.org/10.1007/s40495-016-0074-x
14. Shin SY, Fauman EB, Petersen AK, Krumsiek J, Santos R, Huang J, et al. An atlas of genetic influences on human blood metabolites. Nat Genet. 2014;
15. Vinayavekhin N, EA H, Saghatelian A. Exploring disease through metabolomics. ACS Chem Biol TA - TT -. 2010;5(1):91-103.
16. Huang J, Weinstein SJ, Moore SC, Derkach A, Hua X, Mondul AM, et al. Prediagnostic Serum Metabolomic Profiling of Prostate Cancer Survival. Journals Gerontol Ser A [Internet]. 2019 May 16;74(6):853-9. Available from: https://doi.org/10.1093/gerona/gly128
17. Center MM, Jemal A, Lortet-Tieulent J, Ward E, Ferlay J, Brawley O, et al. International Variation in Prostate Cancer Incidence and Mortality Rates•. Eur Urol TA - TT -. 2012;61(6).
18. Hirst CJ, Cabrera C, Kirby M. Epidemiology of castration resistant prostate cancer: A longitudinal analysis using a UK primary care database. Cancer Epidemiol [Internet]. 2012;36(6):e349-53. Available from: http://www.sciencedirect.com/science/article/ pii/S1877782112001099
19. Neal DE, Metcalfe C, Donovan JL, Lane JA, Davis M, Young GJ, et al. Ten-year Mortality, Disease Progression, and Treatment-related Side Effects in Men with Localised Prostate Cancer from the ProtecT Randomised Controlled Trial According to Treatment Received. Eur Urol [Internet]. 2020;77(3):320-30. Available from: https://www.sciencedirect.com/science/article/pii/S0302283819308371
20. Hackshaw-McGeagh LE, Perry RE, Leach VA, Qandil S, Jeffreys M, Martin RM, et al. A systematic review of dietary, nutritional, and physical activity interventions for the prevention of prostate cancer progression and mortality. Cancer Causes Control [Internet]. 2015;26(11):1521-50. Available from: https://doi.org/10.1007/s10552-015-0659-4
21. Lin P-H, Aronson W, Freedland SJ. Nutrition, dietary interventions and prostate cancer: the latest evidence. BMC Med [Internet]. 2015;13(1):3. Available from: https://doi.org/10.1186/s12916-014-0234-y
22. Peisch SF, Van Blarigan EL, Chan JM, Stampfer MJ, Kenfield SA. Prostate cancer progression and mortality: a review of diet and lifestyle factors. World J Urol [Internet]. 2017 Jun;35(6):867-74. Available from: https:// doi.org/10.1007/s00345-016-1914-3
23. Research WCRF. Welcome to PRACTICAL London: The Institute of Cancer Research.
24. World Cancer Research Fund. International/ American Institute for Cancer Report:Continuous Update Project Diet, Nutrition, Physical Activity, and Prostate Cancer. 2014; Available from: Www.wcrf.org/sites/default/files/Prostate-Cancer-2014-Report.pdf
25. Richman EL, Carroll PR, Chan JM. Vegetable and fruit intake after diagnosis and risk of prostate cancer progression. Int J Cancer [Internet]. 2012 Jul 1;131(1):201-10. Available from: https://doi.org/10.1002/ijc. 26348
26. Chan JM, Holick CN, Leitzmann MF, Rimm EB, Willett WC, Stampfer MJ, et al. Diet After Diagnosis and the Risk of Prostate Cancer Progression, Recurrence, and Death (United States). Cancer Causes Control [Internet]. 2006;17(2):199-208. Available from: https://doi.org/10.1007/ s10552-005-0413-4
27. Ma RW-L, Chapman K. A systematic review of the effect of diet in prostate cancer prevention and treatment. J Hum Nutr Diet [Internet]. 2009 Jun 1;22(3):187-99. Available from: https:/ /doi.org/10.1111/j.1365-277X.2009.00946.x
28. Downer MK, Batista JL, Mucci LA, Stampfer MJ, Epstein MM, Håkansson N, et al. Dairy intake in relation to prostate cancer survival. Int J Cancer [Internet]. 2017 May 1;140(9):2060-9. Available from: https://doi.org/10.1002/ijc. 30642
29. Haseen F, Cantwell MM, O'Sullivan JM, Murray LJ. Is there a benefit from lycopene supplementation in men with prostate cancer? A systematic review. Prostate Cancer Prostatic Dis [Internet]. 2009;12(4):325-32. Available from:
https://doi.org/10.1038/pcan. 2009.38
30. Rowles JL, Ranard KM, Smith JW, An R, Erdman JW. Increased dietary and circulating lycopene are associated with reduced prostate cancer risk: a systematic review and meta-analysis. Prostate Cancer Prostatic Dis [Internet]. 2017;20(4):361-77. Available from: https://doi.org/10.1038/pcan.2017.25
31. Ilic D, Misso M. Lycopene for the prevention and treatment of benign prostatic hyperplasia and prostate cancer: A systematic review. Maturitas [Internet]. 2012;72(4):269-76. Available from: http://www.sciencedirect.com/science/article/pii/S0378512212001600
32. Richman EL, Kenfield SA, Stampfer MJ, Paciorek A, Carroll PR, Chan JM. Physical Activity after Diagnosis and Risk of Prostate Cancer Progression: Data from the Cancer of the Prostate Strategic Urologic Research Endeavor. Cancer Res [Internet]. 2011 Jun 1;71(11):3889 LP - 3895. Available from: http:// cancerres.aacrjournals.org/content/71/11/3889.abstract
33. Donovan JL, Young GJ, Walsh EI, Metcalfe C, Lane JA, Martin RM, et al. A prospective cohort and extended comprehensive-cohort design provided insights about the generalizability of a pragmatic trial: the ProtecT prostate cancer trial. J Clin Epidemiol [Internet]. 2018;96:35-46. Available from: https://www.sciencedirect.com/science/article/pii/S0895435617304675
34. Hackshaw-McGeagh L, Lane JA, Persad R, Gillatt D, Holly JMP, Koupparis A, et al. Prostate cancer - evidence of exercise and nutrition trial (PrEvENT): study protocol
for a randomised controlled feasibility trial. Trials [Internet]. 2016;17(1):123. Available from: https:/ / doi.org/10.1186/s13063-016-1248-x
35. Borges MC, Haycock PC, Zheng J, Hemani G, Holmes M V, Davey Smith G, et al. Role of circulating polyunsaturated fatty acids on cardiovascular diseases risk: analysis using Mendelian randomization and fatty acid genetic association data from over 114,000 UK Biobank participants. BMC Med [Internet]. 2022;20(1):210. Available from: https:/ / doi.org/10.1186/s12916-022-02399-w
36. Culp MB, Soerjomataram I, Efstathiou JA, Bray F, Jemal A. Recent Global Patterns in Prostate Cancer Incidence and Mortality Rates. Eur Urol. 2020 Jan;77(1):38-52.
37. Rawla P. Epidemiology of Prostate Cancer. World J Oncol [Internet]. 2019/04/20. 2019 Apr;10(2):63-89. Available from: https:/ / pubmed.ncbi.nlm.nih.gov/31068988
38. Ferlay J EM, Lam F, ColombetM, Mery L, Piñeros M, Znaor A, Soerjomayaram I BF. Cancer today:International Agency for Research into Cancer (IARC), WHO; 2018 [Internet]. Available from: https:// gco.iarc.fr/today/home
39. Lloyd T, Hounsome L, Mehay A, Mee S, Verne J, Cooper A. Lifetime risk of being diagnosed with, or dying from, prostate cancer by major ethnic group in England 2008-2010. BMC Med [Internet]. 2015;13(1):171. Available from:
https://doi.org/10.1186/ s12916-015-0405-5
40. (NCRAS) NCR and AS. Cancer prevalence in England: 21 year prevalence by demographic and geographic measures.
41. Ferlay J, Ervik M, Lam F, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram I BF. Global Cancer Observatory: Cancer Today. Lyon, Fr Int Agency Res Cancer [Internet]. 2020; Available from: https://gco.iarc.fr/today, accessed
42. Compendium of Population Health Indicators Annex 3: Health and Social Care Information Centre [Internet]. 2015. Available from:
https://digital.nhs.uk/binaries/content/documents/corporate-website/publication-system/ci-hub/compendium-indicators/compendiumindicators/publicationsystem:cilandingasset\[3\]/publicationsystem:Attachme nts\%5B3\%5D/publicationsystem:attachmentResource
43. Age-Standardized Rates: Statistics Canada [Internet]. 2017. Available from: https://www.statcan.gc.ca/eng/dai/btd/asr
44. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018 Nov;68(6):394-424.
45. Yu H, Harris RE, Gao YT, Gao R, Wynder EL. Comparative epidemiology of cancers of the colon, rectum, prostate and breast in Shanghai, China versus the United States. Int J Epidemiol. 1991 Mar;20(1):76-81.
46. Shimizu H, Ross RK, Bernstein L, Yatani R, Henderson BE, Mack TM. Cancers of the prostate and breast among Japanese and white immigrants in Los Angeles County. Br J Cancer. 1991 Jun;63(6):963-6.
47. Draisma G, Etzioni R, Tsodikov A, Mariotto A, Wever E, Gulati R, et al. Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context. J Natl Cancer Inst TA - TT -. 2009;101(6):374-83.
48. Quinn M, Babb P. Patterns and trends in prostate cancer incidence, survival, prevalence and mortality. Part I: international comparisons INTERNATIONAL PROSTATE CANCER TRENDS: PART I. BJU Int TA - TT -. 2002;90(2):162-73.
49. RB H, RG Z, Gridley G, Swanson C, RS G, GM S, et al. Dietary factors and risks for prostate cancer among blacks and whites in the United States. LK -
https://bris.on.worldcat.org/oclc/118697466. Cancer Epidemiol biomarkers Prev a Publ Am Assoc Cancer Res cosponsored by Am Soc Prev Oncol TA - TT -. 1999;8(1):25-34.
50. AS W, LN K, AH W, EM J, RP G, GR H, et al. Prostate cancer in relation to diet, physical activity, and body size in blacks, whites, and Asians in the United States and Canada. LK - https:/ /bris.on.worldcat.org/oclc/118222722. J Natl Cancer Inst TA TT -. 1995;87(9):652-61.
51. Etzioni R, Penson DF, Legler JM, Di Tommaso D, Boer R, Gann PH, et al. Overdiagnosis due to prostate-specific antigen screening: lessons from US prostate cancer incidence trends. J Natl Cancer Inst. 2002;94(13):981-90.
52. Badal S, Aiken W, Morrison B, Valentine H, Bryan S, Gachii A, et al. Disparities in prostate cancer incidence and mortality rates: Solvable or not? Prostate [Internet]. 2020 Jan 1;80(1):3-16. Available from: https://doi.org/10.1002/pros. 23923
53. Chornokur G, Dalton K, Borysova ME, Kumar NB. Disparities at presentation, diagnosis, treatment, and survival in African American men, affected by prostate cancer. Prostate. 2011 Jun;71(9):985-97.
54. Graham-Steed T, Uchio E, Wells CK, Aslan M, Ko J, Concato J. 'Race' and Prostate Cancer Mortality in Equal-access Healthcare Systems. Am J Med [Internet]. 2013;126(12):1084-8. Available from:
https://www.sciencedirect.com/science/article/pii/S0002934313007638
55. Evans S, Metcalfe C, Ibrahim F, Persad R, Ben-Shlomo Y. Investigating Black-White differences in prostate cancer prognosis: A systematic review and meta-analysis. Int J Cancer [Internet]. 2008 Jul 15;123(2):430-5. Available from:
https://doi.org/10.1002/ijc. 23500
56. Ng K, Wilson P, Mutsvangwa K, Hounsome L, Shamash J. Overall survival of black and white men with metastatic castration-resistant prostate cancer (mCRPC): a 20year retrospective analysis in the largest healthcare trust in England. Prostate Cancer Prostatic Dis [Internet]. 2021;24(3):718-24. Available from: https://doi.org/10.1038/s41391-020-00316-x
57. Welch HG, Schwartz LM, Woloshin S. Are increasing 5-year survival rates evidence of success against cancer? JAMA. 2000 Jun;283(22):2975-8.
58. (NCRAS) NCR and AS. Prostate cancer survival [Internet]. Available from: https://www.cancerdata.nhs.uk/
59. National Institute for Health and Care Excellence. Prostate cancer: diagnosis and management (NG131). 2019.
60. Gold SA, Hale GR, Bloom JB, Smith CP, Rayn KN, Valera V, et al. Follow-up of negative MRI-targeted prostate biopsies: when are we missing cancer? World J Urol TA - TT -. 2019;37(2):235-41.
61. Excellence NI for H and C. Prostate cancer: diagnosis and management. 2019;May.
62. Humphrey PA. Gleason grading and prognostic factors in carcinoma of the prostate. Mod Pathol [Internet]. 2004;17(3):292-306. Available from: https://doi.org/10.1038/modpathol. 3800054
63. van Leenders GJLH, van der Kwast TH, Grignon DJ, Evans AJ, Kristiansen G, Kweldam CF, et al. The 2019 International Society of Urological Pathology (ISUP) Consensus Conference on Grading of Prostatic Carcinoma. Am J Surg Pathol [Internet]. 2020;44(8). Available from:
https://journals.lww.com/ajsp/Fulltext/2020/08000/The_2019_International_Societ y_of_Urological.1.aspx
64. (AJCC) AJC on C. What is cancer staging? 2018;
65. Gnanapragasam VJ, Bratt O, Muir K, Lee LS, Huang HH, Stattin P, et al. The Cambridge Prognostic Groups for improved prediction of disease mortality at diagnosis in primary non-metastatic prostate cancer: a validation study. BMC Med. 2018 Feb;16(1):31.
66. Partnership, National Prostate Cancer, , Audit Healthcare Quality Improvement, Partnership, National Prostate Cancer, , Audit Healthcare Quality Improvement. Using the Cambridge Prognostic Groups for risk stratification of prostate cancer in the National Prostate Cancer Audit: How could it impact our estimates of potential 'over-treatment'?
67. Rosario DJ, Lane JA, Metcalfe C, Donovan JL, Doble A, Goodwin L, et al. Short term outcomes of prostate biopsy in men tested for cancer by prostate specific antigen: prospective evaluation within ProtecT study. BMJ [Internet]. 2012;344. Available from: https://www.bmj.com/content/344/bmj.d7894
68. Donovan JL, Hamdy FC, Lane JA, Mason M, Metcalfe C, Walsh E, et al. PatientReported Outcomes after Monitoring, Surgery, or Radiotherapy for Prostate Cancer. N Engl J Med [Internet]. 2016 Sep 14;375(15):1425-37. Available from: https://doi.org/10.1056/NEJMoa1606221
69. Thorsen L, Courneya KS, Stevinson C, Fosså SD. A systematic review of physical activity in prostate cancer survivors: outcomes, prevalence, and determinants. Support Care Cancer [Internet]. 2008;16(9):987-97. Available from: https://doi.org/10.1007/s00520-008-0411-7
70. Sadeghian M, Asadi M, Rahmani S, Sadeghi N, Hosseini SA, Zare Javid A. Lycopene Does Not Affect Prostate-Specific Antigen in Men with Non-Metastatic Prostate Cancer: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. Nutr Cancer [Internet]. 2020 Dec 23;1-12. Available from:
https://doi.org/10.1080/01635581.2020.1862254
71. Song Y, Chavarro JE, Cao Y, Qiu W, Mucci L, Sesso HD, et al. Whole milk intake is associated with prostate cancer-specific mortality among U.S. male physicians. J Nutr [Internet]. 2012/12/19. 2013 Feb;143(2):189-96. Available from: https:/ / pubmed.ncbi.nlm.nih.gov/23256145
72. Mandair D, Rossi RE, Pericleous M, Whyand T, Caplin ME. Prostate cancer and the influence of dietary factors and supplements: a systematic review. Nutr Metab (Lond) [Internet]. 2014;11(1):30. Available from: https://doi.org/10.1186/1743-7075-11-30
73. Davies AA, Davey Smith G, Harbord R, Bekkering GE, Sterne JAC, Beynon R, et al. Nutritional Interventions and Outcome in Patients With Cancer or Preinvasive Lesions: Systematic Review LK - https://bris.on.worldcat.org/oclc/364374445. J Natl Cancer Inst TA - TT -. 2006;98(14):961-73.
74. Heber D, QY L. Overview of mechanisms of action of lycopene. LK https://bris.on.worldcat.org/oclc/110680845. Exp Biol Med (Maywood, NJ) TA - TT -. 2002;227(10):920-3.
75. P. C, W. Z, X. W, K. Z, D.S. N, L. Z, et al. Lycopene and risk of prostate cancer. Med (United States) TA - TT -. 2015;94(33).
76. Chen P, Zhang W, Wang X, Zhao K, Negi D, Zhuo L, et al. Lycopene and Risk of Prostate Cancer: A Systematic Review and Meta-Analysis. Med (United States) TA TT -. 2015;94(33):e1260.
77. Etminan M, Takkouche B, Caamaño-Isorna F. The Role of Tomato Products and Lycopene in the Prevention of Prostate Cancer: A Meta-Analysis of Observational Studies. Cancer Epidemiol Biomarkers \&amp; Prev [Internet]. 2004 Mar 1;13(3):340 LP - 345. Available from:
http://cebp.aacrjournals.org/content/13/3/340.abstract
78. Piwowarczyk K, Paw M, Ryszawy D, Rutkowska-Zapała M, Madeja Z, Siedlar M, et al. Connexin43high prostate cancer cells induce endothelial connexin 43 up-regulation through the activation of intercellular ERK1/2-dependent signaling axis. Eur J Cell Biol [Internet]. 2017;96(4):337-46. Available from: https://www.sciencedirect.com/science/article/pii/S0171933516302990
79. Boucher J, AC B, Debant M, Vix J, Harnois T, Bourmeyster N, et al. Cx43 Present at the Leading Edge Membrane Governs Promigratory Effects of OsteoblastConditioned Medium on Human Prostate Cancer Cells in the Context of Bone Metastasis. Cancers TA - TT -. 2020;12(10).
80. Kucuk O, Sarkar FH, Sakr W, Khachik F, Djuric Z, Banerjee M, et al. Lycopene in the treatment of prostate cancer. Pure Appl Chem [Internet]. 2002;74(8):1443-50.
Available from: https://www.degruyter.com/view/journals/ pac/74/8/articlep1443.xml
81. Rafi MM, Kanakasabai S, Reyes MD, Bright JJ. Lycopene modulates growth and survival associated genes in prostate cancer. J Nutr Biochem [Internet].
2013;24(10):1724-34. Available from:
https://www.sciencedirect.com/science/article/pii/S0955286313000727
82. Mansour M, Schwartz D, Judd R, Akingbemi B, Braden T, Morrison E, et al. Thiazolidinediones/PPAR $\gamma$ agonists and fatty acid synthase inhibitors as an experimental combination therapy for prostate cancer. Int J Oncol [Internet]. 2011;38(2):537-46. Available from: https://doi.org/10.3892/ijo.2010.877
83. Roddam AW, Allen NE, Appleby P, Key TJ, Ferrucci L, Carter HB, et al. Insulin-like Growth Factors, Their Binding Proteins, and Prostate Cancer Risk: Analysis of Individual Patient Data from 12 Prospective Studies. Ann Intern Med [Internet]. 2008 Oct 7;149(7):461-71. Available from: https://www.acpjournals.org/doi/abs/10.7326/0003-4819-149-7-200810070-00006
84. Wertz K, Siler U, Goralczyk R. Lycopene: modes of action to promote prostate health. Arch Biochem Biophys [Internet]. 2004;430(1):127-34. Available from: https://www.sciencedirect.com/science/article/pii/S0003986104002620
85. Pettersson A, Kasperzyk JL, Kenfield SA, Richman EL, Chan JM, Willett WC, et al. Milk and dairy consumption among men with prostate cancer and risk of metastases and prostate cancer death. Cancer Epidemiol Biomarkers Prev [Internet]. 2012/02/07. 2012 Mar;21(3):428-36. Available from: https:/ / pubmed.ncbi.nlm.nih.gov/22315365
86. Ahn J, Albanes D, Peters U, Schatzkin A, Lim U, Freedman M, et al. Dairy products, calcium intake, and risk of prostate cancer in the prostate, lung, colorectal, and ovarian cancer screening trial. Cancer Epidemiol biomarkers Prev a Publ Am Assoc Cancer Res cosponsored by Am Soc Prev Oncol. 2007 Dec;16(12):2623-30.
87. Allen NE, Key TJ, Appleby PN, Travis RC, Roddam AW, Tjønneland A, et al. Animal foods, protein, calcium and prostate cancer risk: the European Prospective Investigation into Cancer and Nutrition. Br J Cancer. 2008 May;98(9):1574-81.
88. Bosetti C, Micelotta S, Dal Maso L, Talamini R, Montella M, Negri E, et al. Food groups and risk of prostate cancer in Italy. Int J cancer. 2004 Jun;110(3):424-8.
89. Chan JM, Stampfer MJ, Ma J, Gann PH, Gaziano JM, Giovannucci EL. Dairy products, calcium, and prostate cancer risk in the Physicians' Health Study. Am J Clin Nutr. 2001 Oct;74(4):549-54.
90. Jain MG, Hislop GT, Howe GR, Ghadirian P. Plant foods, antioxidants, and prostate cancer risk: findings from case-control studies in Canada. Nutr Cancer. 1999;34(2):173-84.
91. Kristal AR, Cohen JH, Qu P, Stanford JL. Associations of energy, fat, calcium, and vitamin D with prostate cancer risk. Cancer Epidemiol biomarkers Prev a Publ Am Assoc Cancer Res cosponsored by Am Soc Prev Oncol. 2002 Aug;11(8):719-25.
92. Tseng M, Breslow RA, Graubard BI, Ziegler RG. Dairy, calcium, and vitamin D intakes and prostate cancer risk in the National Health and Nutrition Examination Epidemiologic Follow-up Study cohort. Am J Clin Nutr. 2005 May;81(5):1147-54.
93. Kurahashi N, Inoue M, Iwasaki M, Sasazuki S, Tsugane AS. Dairy product, saturated fatty acid, and calcium intake and prostate cancer in a prospective cohort of Japanese
men. Cancer Epidemiol biomarkers Prev a Publ Am Assoc Cancer Res cosponsored by Am Soc Prev Oncol. 2008 Apr;17(4):930-7.
94. Gao X, LaValley MP, Tucker KL. Prospective studies of dairy product and calcium intakes and prostate cancer risk: a meta-analysis. J Natl Cancer Inst. 2005 Dec;97(23):1768-77.
95. Yang M, Kenfield SA, Van Blarigan EL, Wilson KM, Batista JL, Sesso HD, et al. Dairy intake after prostate cancer diagnosis in relation to disease-specific and total mortality. Int J Cancer [Internet]. 2015 Nov 15;137(10):2462-9. Available from: https://doi.org/10.1002/ijc. 29608
96. Giovannucci E, Liu Y, Stampfer MJ, Willett WC. A Prospective Study of Calcium Intake and Incident and Fatal Prostate Cancer. Cancer Epidemiol Biomarkers \&amp; Prev [Internet]. 2006 Feb 1;15(2):203 LP - 210. Available from: http://cebp.aacrjournals.org/content/15/2/203.abstract
97. Schwartz GG. Prostate Cancer, Serum Parathyroid Hormone, and the Progression of Skeletal Metastases. Cancer Epidemiol Biomarkers \&amp; Prev [Internet]. 2008 Mar 1;17(3):478 LP - 483. Available from:
http://cebp.aacrjournals.org/content/17/3/478.abstract
98. Biernacka KM, Perks CM, Holly JMP. Role of the IGF axis in prostate cancer. Minerva Endocrinol. 2012 Jun;37(2):173-85.
99. Wu JD, Haugk K, Woodke L, Nelson P, Coleman I, Plymate SR. Interaction of IGF signaling and the androgen receptor in prostate cancer progression. J Cell Biochem [Internet]. 2006 Oct 1;99(2):392-401. Available from: https://doi.org/10.1002/jcb. 20929
100. MJ R, Lacchetti C, Bergman J, RJ H, KE H, TM K, et al. Prostate cancer survivorship care guideline: American Society of Clinical Oncology Clinical Practice Guideline endorsement. J Clin Oncol Off J Am Soc Clin Oncol TA - TT -. 2015;33(9):1078-85.
101. Research WCRFI for C. Diet,Nutrition, Physical Activity and Cancer: a Global Perspective. Contin Updat Proj Expert Rep. 2018;
102. Zuniga KB, Chan JM, Ryan CJ, Kenfield SA. Diet and lifestyle considerations for patients with prostate cancer. Urol Oncol Semin Orig Investig [Internet]. 2020;38(3):105-17. Available from: https://www.sciencedirect.com/science/article/pii/S1078143919302479
103. Parsons JK, Pierce JP, Mohler J, Paskett E, Jung S-H, Humphrey P, et al. A randomized trial of diet in men with early stage prostate cancer on active surveillance: Rationale and design of the Men's Eating and Living (MEAL) Study (CALGB 70807 [Alliance]). Contemp Clin Trials [Internet]. 2014;38(2):198-203. Available from:
https://www.sciencedirect.com/science/article/pii/S1551714414000597
104. Parsons JK, Zahrieh D, Mohler JL, Paskett E, Hansel DE, Kibel AS, et al. Effect of a Behavioral Intervention to Increase Vegetable Consumption on Cancer Progression Among Men With Early-Stage Prostate Cancer: The MEAL Randomized Clinical

Trial. JAMA [Internet]. 2020 Jan 14;323(2):140-8. Available from:
https:/ / doi.org/10.1001/jama.2019.20207
105. Imran M, Aslam Gondal T, Atif M, Shahbaz M, Batool Qaisarani T, Hanif Mughal M, et al. Apigenin as an anticancer agent. Phyther Res [Internet]. 2020 Aug 1;34(8):181228. Available from: https:/ / doi.org/10.1002/ptr. 6647
106. Ganai SA. Plant-derived flavone Apigenin: The small-molecule with promising activity against therapeutically resistant prostate cancer. Biomed Pharmacother [Internet]. 2017;85:47-56. Available from: https://www.sciencedirect.com/science/article/ pii/S0753332216314986
107. Shukla S, MacLennan GT, Fu P, Gupta S. Apigenin attenuates insulin-like growth factor-I signaling in an autochthonous mouse prostate cancer model. Pharm Res. 2012;29(6):1506-17.
108. Jarrard D, Filon M, Huang W, Havighurst T, DeShong K, Kim K, et al. A phase II randomized placebo-controlled trial of pomegranate fruit extract in men with localized prostate cancer undergoing active surveillance. Prostate [Internet]. 2021 Jan 1;81(1):41-9. Available from: https:/ /doi.org/10.1002/ pros. 24076
109. Prasad S, Gupta SC, Tyagi AK. Reactive oxygen species (ROS) and cancer: Role of antioxidative nutraceuticals. Cancer Lett [Internet]. 2017;387:95-105. Available from: https://www.sciencedirect.com/science/article/pii/S0304383516302026
110. Sharma P, McClees SF, Afaq F. Pomegranate for Prevention and Treatment of Cancer: An Update. Vol. 22, Molecules . 2017.
111. Rettig MB, Heber D, An J, Seeram NP, Rao JY, Liu H, et al. Pomegranate extract inhibits androgen-independent prostate cancer growth through a nuclear factor-кBdependent mechanism. Mol Cancer Ther [Internet]. 2008 Sep 1;7(9):2662 LP - 2671. Available from: http:/ /mct.aacrjournals.org/content/7/9/2662.abstract
112. Domingo-Domenech J, Mellado B, Ferrer B, Truan D, Codony-Servat J, Sauleda S, et al. Activation of nuclear factor- $\kappa B$ in human prostate carcinogenesis and association to biochemical relapse. Br J Cancer [Internet]. 2005;93(11):1285-94. Available from: https:/ /doi.org/10.1038/ sj.bjc. 6602851
113. Fradet V, Lessard L, Bégin LR, Karakiewicz P, Masson A-MM, Saad F. Nuclear FactorкВ Nuclear Localization Is Predictive of Biochemical Recurrence in Patients with Positive Margin Prostate Cancer. Clin Cancer Res [Internet]. 2004 Dec 15;10(24):8460 LP - 8464. Available from:
http:/ /clincancerres.aacrjournals.org/content/10/24/8460.abstract
114. Shin SY, Petersen AK, Wahl S, Zhai G, Römisch-Margl W, Small KS, et al. Interrogating causal pathways linking genetic variants, small molecule metabolites, and circulating lipids. Genome Med. 2014;
115. Paller CJ, Ye X, Wozniak PJ, Gillespie BK, Sieber PR, Greengold RH, et al. A randomized phase II study of pomegranate extract for men with rising PSA following initial therapy for localized prostate cancer. Prostate Cancer Prostatic Dis [Internet]. 2013;16(1):50-5. Available from: https:/ / doi.org/10.1038/pcan. 2012.20
116. Organization WH. WHO guidelines on physical activity and sedentary behaviour. 2020.
117. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep. 1985;100(2):126-31.
118. SE B, Sjolander A, YT L, Wiklund F, Stattin P, Holmberg E, et al. Physical activity and survival among men diagnosed with prostate cancer. Cancer Epidemiol biomarkers Prev a Publ Am Assoc Cancer Res cosponsored by Am Soc Prev Oncol TA - TT -2015;24(1):57-64.
119. JR G, PM L, SF F. Effects of exercise on treatment-related adverse effects for patients with prostate cancer receiving androgen-deprivation therapy: a systematic review. J Clin Oncol Off J Am Soc Clin Oncol TA - TT -. 2014;32(4):335-46.
120. Cormie P, Galvão DA, Spry N, Joseph D, Chee R, Taaffe DR, et al. Can supervised exercise prevent treatment toxicity in patients with prostate cancer initiating androgen-deprivation therapy: a randomised controlled trial Exercise programme to reduce treatment toxicity when initiating ADT. BJU Int TA - TT -. 2015;115(2):256-66.
121. Fong DYT, Ho JWC, Hui BPH, Lee AM, Macfarlane DJ, Leung SSK, et al. Physical activity for cancer survivors: meta-analysis of randomised controlled trials LK https://bris.on.worldcat.org/oclc/5964400793. BMJ Br Med J TA - TT -. 2012;344(7844):17.
122. Galvao DA, Spry N, Denham J, Taaffe DR, Cormie P, Joseph D, et al. A Multicentre Year-long Randomised Controlled Trial of Exercise Training Targeting Physical Functioning in Men with Prostate Cancer Previously Treated with Androgen Suppression and Radiation from TROG 03.04 RADAR LK https://bris.on.worldcat.org/oclc/55. Eur Urol TA - TT -. 2014;65(5):856-64.
123. Segal RJ, Reid RD, Courneya KS, Malone SC, Parliament MB, Scott CG, et al. Resistance Exercise in Men Receiving Androgen Deprivation Therapy for Prostate Cancer. J Clin Oncol. 2003;21(9):1653-9.
124. Segal RJ, Reid RD, Courneya KS, Sigal RJ, Kenny GP, Prud'Homme DG, et al. Randomized Controlled Trial of Resistance or Aerobic Exercise in Men Receiving Radiation Therapy for Prostate Cancer. J Clin Oncol [Internet]. 2009 Jan 20;27(3):34451. Available from: https://doi.org/10.1200/JCO.2007.15.4963
125. Kenfield SA, Stampfer MJ, Giovannucci E, Chan JM. Physical activity and survival after prostate cancer diagnosis in the health professionals follow-up study. J Clin Oncol [Internet]. 2011/01/04. 2011 Feb 20;29(6):726-32. Available from: https://pubmed.ncbi.nlm.nih.gov/21205749
126. CH P, Fall K, JR R, SC M, HO A, SO A, et al. A Walking Intervention Among Men With Prostate Cancer: A Pilot Study. Clin Genitourin cancer TA - TT -. 2017;15(6):e1021-8.
127. Hackshaw-McGeagh LE, Penfold C, Shingler E, Robles LA, Perks CM, Holly JMP, et al. Phase II randomised control feasibility trial of a nutrition and physical activity
intervention after radical prostatectomy for prostate cancer. BMJ Open [Internet]. 2019 Nov 1;9(11):e029480. Available from:
http://bmjopen.bmj.com/content/9/11/e029480.abstract
128. Rocha-Rodrigues S, Matos A, Afonso J, Mendes-Ferreira M, Abade E, Teixeira E, et al. Skeletal Muscle-Adipose Tissue-Tumor Axis: Molecular Mechanisms Linking Exercise Training in Prostate Cancer. Vol. 22, International Journal of Molecular Sciences. 2021.
129. McGee SL, Walder KR. Exercise and the skeletal muscle epigenome. Cold Spring Harb Perspect Med. 2017;7(9):a029876.
130. Dai JY, Wang B, Wang X, Cheng A, Kolb S, Stanford JL, et al. Vigorous Physical Activity Is Associated with Lower Risk of Metastatic-Lethal Progression in Prostate Cancer and Hypomethylation in the CRACR2A Gene. Cancer Epidemiol Biomarkers Prev [Internet]. 2019 Feb 4;28(2):258-64. Available from:
https://doi.org/10.1158/1055-9965.EPI-18-0622
131. Institute TEB. What is metabolomics? [Internet]. [cited 2021 Oct 15]. Available from: https://www.ebi.ac.uk/training/online/courses/metabolomics-introduction/whatis/
132. Lima AR, Bastos M de L, Carvalho M, Guedes de Pinho P. Biomarker Discovery in Human Prostate Cancer: an Update in Metabolomics Studies. Transl Oncol [Internet]. 2016;9(4):357-70. Available from:
http://www.sciencedirect.com/science/article/ pii/S1936523316300511
133. Trock BJ. Application of metabolomics to prostate cancer. Urol Oncol Semin Orig Investig [Internet]. 2011;29(5):572-81. Available from:
http://www.sciencedirect.com/science/article/ pii/S1078143911002444
134. Schmidt DR, Patel R, Kirsch DG, Lewis CA, Vander Heiden MG, Locasale JW. Metabolomics in cancer research and emerging applications in clinical oncology. CA Cancer J Clin [Internet]. 2021 Jul 1;71(4):333-58. Available from:
https://doi.org/10.3322/caac. 21670
135. Tsutsui H, Maeda T, Min JZ, Inagaki S, Higashi T, Kagawa Y, et al. Biomarker discovery in biological specimens (plasma, hair, liver and kidney) of diabetic mice based upon metabolite profiling using ultra-performance liquid chromatography with electrospray ionization time-of-flight mass spectrometry. Clin Chim Acta [Internet]. 2011;412(11):861-72. Available from: https://www.sciencedirect.com/science/article/pii/S0009898110007746
136. Zhang A, Sun H, Yan G, Wang P, Han Y, Wang X. Metabolomics in diagnosis and biomarker discovery of colorectal cancer. Cancer Lett [Internet]. 2014;345(1):17-20. Available from:
https://www.sciencedirect.com/science/article/pii/S0304383513008203
137. Spratlin JL, Serkova NJ, Eckhardt SG. Clinical Applications of Metabolomics in Oncology: A Review. Clin Cancer Res [Internet]. 2009 Jan 15;15(2):431 LP - 440. Available from: http://clincancerres.aacrjournals.org/content/15/2/431.abstract
138. Roberts LD, Souza AL, Gerszten RE, Clish CB. Targeted metabolomics. Curr Protoc Mol Biol [Internet]. 2012 Apr;Chapter 30:Unit30.2-30.2.24. Available from:
https:// pubmed.ncbi.nlm.nih.gov/22470063
139. Putri SP, Yamamoto S, Tsugawa H, Fukusaki E. Current metabolomics: Technological advances. J Biosci Bioeng [Internet]. 2013;116(1):9-16. Available from:
http://www.sciencedirect.com/science/article/pii/S1389172313000054
140. Patti GJ. Separation strategies for untargeted metabolomics. J Sep Sci [Internet]. 2011

Dec 1;34(24):3460-9. Available from: https://doi.org/10.1002/jssc. 201100532
141. Zhou B, Xiao JF, Tuli L, Ressom HW. LC-MS-based metabolomics. Mol Biosyst [Internet]. 2011/11/01. 2012 Feb;8(2):470-81. Available from:
https:// pubmed.ncbi.nlm.nih.gov/22041788
142. Dettmer K, Aronov PA, Hammock BD. Mass spectrometry-based metabolomics. Mass Spectrom Rev [Internet]. 2007;26(1):51-78. Available from:
https:// pubmed.ncbi.nlm.nih.gov/16921475
143. Villas-Bôas SG, Mas S, Åkesson M, Smedsgaard J, Nielsen J. Mass spectrometry in metabolome analysis. Mass Spectrom Rev [Internet]. 2005 Sep 1;24(5):613-46. Available from: https://doi.org/10.1002/mas. 20032
144. Griffiths WJ, Koal T, Wang Y, Kohl M, Enot DP, Deigner H. Targeted Metabolomics for Biomarker Discovery. Angew Chemie Int Ed. 2010;49(32):5426-45.
145. Wishart DS. NMR metabolomics: A look ahead. J Magn Reson [Internet]. 2019;306:155-61. Available from:
http://www.sciencedirect.com/science/article/pii/S109078071930134X
146. Wishart DS. Quantitative metabolomics using NMR. TrAC Trends Anal Chem [Internet]. 2008;27(3):228-37. Available from: http://www.sciencedirect.com/science/article/ pii/S0165993607002622
147. Reo N V. NMR-BASED METABOLOMICS. Drug Chem Toxicol [Internet]. 2002 Jan 1;25(4):375-82. Available from: https:/ / doi.org/10.1081/DCT-120014789
148. Soininen P, Kangas AJ, Würtz P, Tukiainen T, Tynkkynen T, Laatikainen R, et al. High-throughput serum NMR metabonomics for cost-effective holistic studies on systemic metabolism. Analyst [Internet]. 2009;134(9):1781-5. Available from: http://dx.doi.org/10.1039/B910205A
149. Würtz P, Kangas AJ, Soininen P, Lawlor DA, Davey Smith G, Ala-Korpela M. Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Large-Scale Epidemiology: A Primer on -Omic Technologies. Am J Epidemiol. 2017;
150. Soininen P, Kangas Antti J, Würtz P, Suna T, Ala-Korpela M. Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Cardiovascular Epidemiology and Genetics. Circ Cardiovasc Genet [Internet]. 2015;8(1):192-206. Available from: https://doi.org/10.1161/CIRCGENETICS.114.000216
151. Moses C, Garcia-Bloj B, Harvey AR, Blancafort P. Hallmarks of cancer: The CRISPR generation. Eur J Cancer TA - TT -. 2018;93:10-8.
152. Hanahan D, Weinberg RA. The Hallmarks of Cancer. Cell [Internet]. 2000;100(1):5770. Available from:
https://www.sciencedirect.com/science/article/pii/S0092867400816839
153. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. Cell TA - TT -. 2011;144(5):646-74.
154. Warburg O, Wind F, Negelein E. THE METABOLISM OF TUMORS IN THE BODY. J Gen Physiol. 1927 Mar;8(6):519-30.
155. Warburg O. Über den Stoffwechsel der Carcinomzelle. Naturwissenschaften [Internet]. 1924;12(50):1131-7. Available from: https:// doi.org/10.1007/BF01504608
156. Pavlova NN, Thompson CB. The Emerging Hallmarks of Cancer Metabolism. Cell Metab [Internet]. 2016;23(1):27-47. Available from: https://www.sciencedirect.com/science/article/ pii/S155041311500621X
157. Abate-Shen C, Shen MM. The prostate-cancer metabolome. Nature [Internet]. 2009;457(7231):799-800. Available from: https://doi.org/10.1038/457799a
158. Hebert JR, Miller DR. Methodologic considerations for investigating the diet-cancer link. Am J Clin Nutr [Internet]. 1988 Jun 1;47(6):1068-77. Available from: https://doi.org/10.1093/ajcn/47.6.1068
159. Rodríguez-Molinero J, Migueláñez-Medrán BD, Puente-Gutiérrez C, DelgadoSomolinos E, Martín Carreras-Presas C, Fernández-Farhall J, et al. Association between Oral Cancer and Diet: An Update. Vol. 13, Nutrients . 2021.
160. Guertin KA, Moore SC, Sampson JN, Huang W-Y, Xiao Q, Stolzenberg-Solomon RZ, et al. Metabolomics in nutritional epidemiology: identifying metabolites associated with diet and quantifying their potential to uncover diet-disease relations in populations. Am J Clin Nutr [Internet]. 2014 Jul 1;100(1):208-17. Available from: https://doi.org/10.3945/ajen.113.078758
161. Li J, Guasch-Ferré M, Chung W, Ruiz-Canela M, Toledo E, Corella D, et al. The Mediterranean diet, plasma metabolome, and cardiovascular disease risk. Eur Heart J [Internet]. 2020 Jul 21;41(28):2645-56. Available from: https://doi.org/10.1093/ eurheartj/ehaa209
162. Kelly RS, Kelly MP, Kelly P. Metabolomics, physical activity, exercise and health: A review of the current evidence. Biochim Biophys Acta - Mol Basis Dis [Internet]. 2020;1866(12):165936. Available from:
https://www.sciencedirect.com/science/article/pii/S0925443920302842
163. Al-Khelaifi F, Diboun I, Donati F, Botrè F, Alsayrafi M, Georgakopoulos C, et al. A pilot study comparing the metabolic profiles of elite-level athletes from different sporting disciplines. Sport Med - Open [Internet]. 2018;4(1):2. Available from: https://doi.org/10.1186/s40798-017-0114-z
164. Danaher J, Gerber T, Wellard RM, Stathis CG, Cooke MB. The use of metabolomics to monitor simultaneous changes in metabolic variables following supramaximal low volume high intensity exercise. Metabolomics [Internet]. 2015;12(1):7. Available from:
https://doi.org/10.1007/s11306-015-0883-7
165. Zafeiridis A, Chatziioannou AC, Sarivasiliou H, Kyparos A, Nikolaidis MG, Vrabas IS, et al. Global Metabolic Stress of Isoeffort Continuous and High Intensity Interval Aerobic Exercise: A Comparative 1H NMR Metabonomic Study. J Proteome Res [Internet]. 2016 Dec 2;15(12):4452-63. Available from: https://doi.org/10.1021/acs.jproteome.6b00545
166. Xiao Q, Moore SC, Keadle SK, Xiang Y-B, Zheng W, Peters TM, et al. Objectively measured physical activity and plasma metabolomics in the Shanghai Physical Activity Study. Int J Epidemiol [Internet]. 2016 Oct 1;45(5):1433-44. Available from: https://doi.org/10.1093/ije/dyw033
167. Ulaszewska MM, Weinert CH, Trimigno A, Portmann R, Andres Lacueva C, Badertscher R, et al. Nutrimetabolomics: An Integrative Action for Metabolomic Analyses in Human Nutritional Studies. Mol Nutr Food Res [Internet]. 2019 Jan 1;63(1):1800384. Available from: https://doi.org/10.1002/mnfr. 201800384
168. Lin H-M, Yeung N, Hastings JF, Croucher DR, Huynh K, Meikle TG, et al. Relationship between Circulating Lipids and Cytokines in Metastatic CastrationResistant Prostate Cancer. Cancers (Basel). 2021 Oct;13(19).
169. Zoni E, Minoli M, Bovet C, Wehrhan A, Piscuoglio S, Ng CKY, et al. Preoperative plasma fatty acid metabolites inform risk of prostate cancer progression and may be used for personalized patient stratification. BMC Cancer. 2019 Dec;19(1):1216.
170. Zheng H, Dong B, Ning J, Shao X, Zhao L, Jiang Q, et al. NMR-based metabolomics analysis identifies discriminatory metabolic disturbances in tissue and biofluid samples for progressive prostate cancer. Clin Chim Acta [Internet]. 2020;501:241-51. Available from:
https://www.sciencedirect.com/science/article/pii/S0009898119321102
171. Häggström C, Stocks T, Nagel G, Manjer J, Bjørge T, Hallmans G, et al. Prostate cancer, prostate cancer death, and death from other causes, among men with metabolic aberrations. Epidemiology. 2014;25(6):823.
172. Moreel X, Allaire J, Léger C, Caron A, Labonté M-È, Lamarche B, et al. Prostatic and Dietary Omega-3 Fatty Acids and Prostate Cancer Progression during Active Surveillance. Cancer Prev Res [Internet]. 2014 Jul 1;7(7):766-76. Available from: https://doi.org/10.1158/1940-6207.CAPR-13-0349
173. Rodríguez-Blanco G, Burgers PC, Dekker LJM, Ijzermans JJN, Wildhagen MF, Schenk-Braat EAM, et al. Serum levels of arachidonic acid metabolites change during prostate cancer progression. Prostate. 2014 May;74(6):618-27.
174. Huang J, Mondul AM, Weinstein SJ, Derkach A, Moore SC, Sampson JN, et al. Prospective serum metabolomic profiling of lethal prostate cancer. Int J Cancer [Internet]. 2019 Dec 15;145(12):3231-43. Available from: https://doi.org/10.1002/ijc. 32218
175. The alpha-tocopherol, beta-carotene lung cancer prevention study: design, methods, participant characteristics, and compliance. The ATBC Cancer Prevention Study

Group. Ann Epidemiol. 1994 Jan;4(1):1-10.
176. Stabler S, Koyama T, Zhao Z, Martinez-Ferrer M, Allen RH, Luka Z, et al. Serum Methionine Metabolites Are Risk Factors for Metastatic Prostate Cancer Progression. PLoS One [Internet]. 2011 Aug 10;6(8):e22486. Available from:
https://doi.org/10.1371/journal.pone. 0022486
177. Cook RJ, Coleman R, Brown J, Lipton A, Major P, Hei YJ, et al. Markers of bone metabolism and survival in men with hormone-refractory metastatic prostate cancer. Clin cancer Res an Off J Am Assoc Cancer Res. 2006 Jun;12(11 Pt 1):3361-7.
178. Lv W, Shang H, Pei X, Chen Y, Xie H, He D, et al. A simple prognostic model involving prostate-specific antigen, alkaline phosphatase and albumin for predicting the time required to progress to castration-resistant prostate cancer in patients who received androgen deprivation therapy. Int Urol Nephrol. 2017 Jan;49(1):61-7.
179. Sejima T, Iwamoto H, Masago T, Morizane S, Yao A, Isoyama T, et al. Low preoperative levels of serum albumin predict lymph node metastases and ultimately correlate with a biochemical recurrence of prostate cancer in radical prostatectomy patients. Cent Eur J Urol. 2013;66(2):126-32.
180. Wang N, Liu J-Y, Li X, Deng M-H, Long Z, Tang J, et al. Pretreatment serum albumin/globulin ratio as a prognostic biomarker in metastatic prostate cancer patients treated with maximal androgen blockade. Asian J Androl. 2018 Jul;21(1):5661.
181. Aydh A, Mori K, D'Andrea D, Motlagh RS, Abufaraj M, Pradere B, et al. Prognostic value of the pre-operative serum albumin to globulin ratio in patients with nonmetastatic prostate cancer undergoing radical prostatectomy. Int J Clin Oncol. 2021 Sep;26(9):1729-35.
182. Pan J, Wang J, Wei Y, Zhang T, Zhang S, Ye D, et al. Combination of body mass index and albumin predicts the survival in metastatic castration-resistant prostate cancer patients treated with abiraterone: A post hoc analysis of two randomized trials. Cancer Med. 2021 Oct;10(19):6697-704.
183. Rantaniemi L, Tammela TLJ, Kujala P, Murtola TJ. Blood cholesterol, tumor clinical characteristics and risk of prostate cancer progression after radical prostatectomy. Scand J Urol. 2018 Aug;52(4):269-76.
184. Mak B, Lin H-M, Kwan EM, Fettke H, Tran B, Davis ID, et al. Combined impact of lipidomic and genetic aberrations on clinical outcomes in metastatic castrationresistant prostate cancer. BMC Med. 2022 Mar;20(1):112.
185. Hirano H, Ide H, Lu Y, Inoue Y, Okada H, Horie S. Impact of Pretreatment Total Cholesterol Level Is Associated With Metastasis of Prostate Cancer. Am J Mens Health. 2020;14(2):1557988320918788.
186. Roberts MJ, Schirra HJ, Lavin MF, Gardiner RA. Metabolomics: A Novel Approach to Early and Noninvasive Prostate Cancer Detection. kju [Internet]. 2011 Feb 19;52(2):7989. Available from: http:/ / dx.doi.org/10.4111/kju.2011.52.2.79
187. Marín-Aguilera M, Pereira M V, Jiménez N, Reig Ò, Cuartero A, Victoria I, et al. Glutamine and Cholesterol Plasma Levels and Clinical Outcomes of Patients with Metastatic Castration-Resistant Prostate Cancer Treated with Taxanes. Vol. 13, Cancers. 2021.
188. Giunchi F, Fiorentino M, Loda M. The Metabolic Landscape of Prostate Cancer. Eur Urol Oncol [Internet]. 2019;2(1):28-36. Available from: https://www.sciencedirect.com/science/article/pii/S2588931118300956
189. Kumar D, Gupta A, Mandhani A, Sankhwar SN. Metabolomics-Derived Prostate Cancer Biomarkers: Fact or Fiction? J Proteome Res [Internet]. 2015 Mar 6;14(3):145564. Available from: https://doi.org/10.1021/pr5011108
190. Kumar D, Gupta A, Mandhani A, Sankhwar SN. NMR spectroscopy of filtered serum of prostate cancer: A new frontier in metabolomics. Prostate [Internet]. 2016 Sep 1;76(12):1106-19. Available from: https://doi.org/10.1002/pros. 23198
191. Sreekumar A, Poisson LM, Rajendiran TM, Khan AP, Cao Q, Yu J, et al. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. Nature [Internet]. 2009;457(7231):910-4. Available from: https://doi.org/10.1038/nature07762
192. Soga T. Cancer metabolism: Key players in metabolic reprogramming. Cancer Sci [Internet]. 2013 Mar 1;104(3):275-81. Available from: https://doi.org/10.1111/cas. 12085
193. Pértega-Gomes N, Baltazar F. Lactate Transporters in the Context of Prostate Cancer Metabolism: What Do We Know? Vol. 15, International Journal of Molecular Sciences . 2014.
194. Costello LC, Franklin RB. The clinical relevance of the metabolism of prostate cancer; zinc and tumor suppression: connecting the dots. Mol Cancer [Internet]. 2006 May 15;5:17. Available from: https:/ / pubmed.ncbi.nlm.nih.gov/16700911
195. Swinnen J V, Heemers H, de Sande T Van, Schrijver E De, Brusselmans K, Heyns W, et al. Androgens, lipogenesis and prostate cancer. J Steroid Biochem Mol Biol [Internet]. 2004;92(4):273-9. Available from: https://www.sciencedirect.com/science/article/pii/S096007600400367X
196. Ettinger SL, Sobel R, Whitmore TG, Akbari M, Bradley DR, Gleave ME, et al. Dysregulation of Sterol Response Element-Binding Proteins and Downstream Effectors in Prostate Cancer during Progression to Androgen Independence. Cancer Res [Internet]. 2004 Mar 15;64(6):2212 LP - 2221. Available from: http://cancerres.aacrjournals.org/content/64/6/2212.abstract
197. Pomare EW, Branch WJ, Cummings JH. Carbohydrate fermentation in the human colon and its relation to acetate concentrations in venous blood. J Clin Invest [Internet]. 1985 May;75(5):1448-54. Available from: https:// pubmed.ncbi.nlm.nih.gov/3998144
198. Ackerman D, Simon MC. Hypoxia, lipids, and cancer: surviving the harsh tumor microenvironment. Trends Cell Biol [Internet]. 2014;24(8):472-8. Available from:
https://www.sciencedirect.com/science/article/pii/S0962892414000877
199. Nomura DK, Lombardi DP, Chang JW, Niessen S, Ward AM, Long JZ, et al. Monoacylglycerol Lipase Exerts Dual Control over Endocannabinoid and Fatty Acid Pathways to Support Prostate Cancer. Chem Biol [Internet]. 2011;18(7):846-56. Available from:
https://www.sciencedirect.com/science/article/pii/S1074552111002006
200. Rysman E, Brusselmans K, Scheys K, Timmermans L, Derua R, Munck S, et al. De novo lipogenesis protects cancer cells from free radicals and chemotherapeutics by promoting membrane lipid saturation. Cancer Res [Internet]. 2010 Oct 15;70(20):8117 LP - 8126. Available from:
http://cancerres.aacrjournals.org/content/70/20/8117.abstract
201. Wu X, Daniels G, Lee P, Monaco ME. Lipid metabolism in prostate cancer. Am J Clin Exp Urol [Internet]. 2014 Jul 12;2(2):111-20. Available from:
https:// pubmed.ncbi.nlm.nih.gov/25374912
202. Zadra G, Photopoulos C, Tyekucheva S, Heidari P, Weng QP, Fedele G, et al. A novel direct activator of AMPK inhibits prostate cancer growth by blocking lipogenesis. EMBO Mol Med [Internet]. 2014 Apr 1;6(4):519-38. Available from:
https://doi.org/10.1002/emmm. 201302734
203. Massie CE, Lynch A, Ramos-Montoya A, Boren J, Stark R, Fazli L, et al. The androgen receptor fuels prostate cancer by regulating central metabolism and biosynthesis.
EMBO J [Internet]. 2011 Jul 6;30(13):2719-33. Available from:
https://doi.org/10.1038/emboj.2011.158
204. Pelton K, Freeman MR, Solomon KR. Cholesterol and prostate cancer. Curr Opin Pharmacol [Internet]. 2012;12(6):751-9. Available from: https://www.sciencedirect.com/science/article/pii/S1471489212001178
205. Freeman MR, Solomon KR. Cholesterol and prostate cancer. J Cell Biochem [Internet]. 2004 Jan 1;91(1):54-69. Available from: https:// doi.org/10.1002/jcb. 10724
206. Di Vizio D, Solomon KR, Freeman MR. Cholesterol and Cholesterol-Rich Membranes in Prostate Cancer: An Update. Tumori J [Internet]. 2008 Sep 1;94(5):633-9. Available from: https:/ / doi.org/10.1177/030089160809400501
207. Chen M, Zhang J, Sampieri K, Clohessy JG, Mendez L, Gonzalez-Billalabeitia E, et al. An aberrant SREBP-dependent lipogenic program promotes metastatic prostate cancer. Nat Genet [Internet]. 2018;50(2):206-18. Available from:
https://doi.org/10.1038/s41588-017-0027-2
208. Nowak DG, Cho H, Herzka T, Watrud K, DeMarco D V, Wang VMY, et al. MYC Drives Pten/Trp-53-deficient proliferation and metastasis due to IL6 secretion and AKT suppression via PHLPP2. Cancer Discov [Internet]. 2015 Jun 1;5(6):636 LP - 651. Available from: http://cancerdiscovery.aacrjournals.org/content/5/6/636.abstract
209. Yue S, Li J, Lee S-Y, Lee HJ, Shao T, Song B, et al. Cholesteryl Ester Accumulation Induced by PTEN Loss and PI3K/AKT Activation Underlies Human Prostate Cancer Aggressiveness. Cell Metab [Internet]. 2014;19(3):393-406. Available from:
https://www.sciencedirect.com/science/article/pii/S1550413114000254
210. Prabhu N, Kapur N, Catalona W, Leikin R, Helenowski I, Jovanovich B, et al. Statin use and risk of prostate cancer biochemical recurrence after radical prostatectomy. Urol Oncol Semin Orig Investig [Internet]. 2021;39(2):130.e9-130.e15. Available from: https://www.sciencedirect.com/science/article/pii/S1078143920304646
211. Heger Z, Merlos Rodrigo MA, Michalek P, Polanska H, Masarik M, Vit V, et al. Sarcosine Up-Regulates Expression of Genes Involved in Cell Cycle Progression of Metastatic Models of Prostate Cancer. PLoS One [Internet]. 2016 Nov 8;11(11):e0165830. Available from: https:/ /doi.org/10.1371/journal.pone.0165830
212. Khan AP, Rajendiran TM, Bushra A, Asangani IA, Athanikar JN, Yocum AK, et al. The Role of Sarcosine Metabolism in Prostate Cancer Progression. Neoplasia [Internet]. 2013;15(5):491-IN13. Available from: https://www.sciencedirect.com/science/article/pii/S1476558613800358
213. Yousefi M, Qujeq D, Shafi H, Tilaki KH. Serum and Urine Levels of Sarcosine in Benign Prostatic Hyperplasia and Newly Diagnosed Prostate Cancer Patients. J Kermanshah Univ Med Sci [Internet]. 2020;24(1):e97000. Available from: https://sites.kowsarpub.com/jkums/articles/97000.html
214. Markin PA, Brito A, Moskaleva N, Fodor M, Lartsova E V, Shpot Y V, et al. Plasma Sarcosine Measured by Gas Chromatography-Mass Spectrometry Distinguishes Prostatic Intraepithelial Neoplasia and Prostate Cancer from Benign Prostate Hyperplasia. Lab Med [Internet]. 2020 Nov 2;51(6):566-73. Available from: https://doi.org/10.1093/labmed/lmaa008
215. Koutros S, Meyer TE, Fox SD, Issaq HJ, Veenstra TD, Huang W-Y, et al. Prospective evaluation of serum sarcosine and risk of prostate cancer in the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial. Carcinogenesis [Internet]. 2013 Oct 1;34(10):2281-5. Available from: https://doi.org/10.1093/carcin/bgt176
216. Lucarelli G, Fanelli M, Larocca AMV, Germinario CA, Rutigliano M, Vavallo A, et al. Serum sarcosine increases the accuracy of prostate cancer detection in patients with total serum PSA less than $4.0 \mathrm{ng} / \mathrm{ml}$. Prostate [Internet]. 2012;72(15):1611-21. Available from: https://doi.org/10.1002/pros. 22514
217. Jentzmik F, Stephan C, Miller K, Schrader M, Erbersdobler A, Kristiansen G, et al. Sarcosine in Urine after Digital Rectal Examination Fails as a Marker in Prostate Cancer Detection and Identification of Aggressive Tumours. Eur Urol [Internet]. 2010;58(1):12-8. Available from:
https://www.sciencedirect.com/science/article/pii/S0302283810000849
218. Issaq HJ, Veenstra TD. Is sarcosine a biomarker for prostate cancer? J Sep Sci [Internet]. 2011;34(24):3619-21. Available from:
https://doi.org/10.1002/jssc. 201100572
219. Bott AJ, Peng I-C, Fan Y, Faubert B, Zhao L, Li J, et al. Oncogenic Myc Induces Expression of Glutamine Synthetase through Promoter Demethylation. Cell Metab [Internet]. 2015;22(6):1068-77. Available from:
https://www.sciencedirect.com/science/article/pii/S1550413115004829
220. Koochekpour S, Majumdar S, Azabdaftari G, Attwood K, Scioneaux R, Subramani D, et al. Serum Glutamate Levels Correlate with Gleason Score and Glutamate Blockade Decreases Proliferation, Migration, and Invasion and Induces Apoptosis in Prostate Cancer Cells. Clin Cancer Res [Internet]. 2012 Nov 1;18(21):5888 LP - 5901. Available from: http://clincancerres.aacrjournals.org/content/18/21/5888.abstract
221. Vaz C V, Alves MG, Marques R, Moreira PI, Oliveira PF, Maia CJ, et al. Androgenresponsive and nonresponsive prostate cancer cells present a distinct glycolytic metabolism profile. Int J Biochem Cell Biol [Internet]. 2012;44(11):2077-84. Available from: https:/ / www.sciencedirect.com/science/article/ pii/S1357272512002919
222. LIU Y, ZUCKIER LS, GHESANI N V. Dominant Uptake of Fatty Acid over Glucose by Prostate Cells: A Potential New Diagnostic and Therapeutic Approach. Anticancer Res [Internet]. 2010 Feb 1;30(2):369 LP - 374. Available from:
http://ar.iiarjournals.org/content/30/2/369.abstract
223. Sadeghi RN, Karami-Tehrani F, Salami S. Targeting prostate cancer cell metabolism: impact of hexokinase and CPT-1 enzymes. Tumor Biol [Internet]. 2015;36(4):2893-905. Available from: https://doi.org/10.1007/s13277-014-2919-4
224. Jadvar H. PET of Glucose Metabolism and Cellular Proliferation in Prostate Cancer. J Nucl Med [Internet]. 2016 Oct 1;57(Supplement 3):25S LP-29S. Available from: http://jnm.snmjournals.org/content/57/Supplement_3/25S.abstract
225. Twum-Ampofo J, Fu D-X, Passaniti A, Hussain A, Siddiqui MM. Metabolic targets for potential prostate cancer therapeutics. Curr Opin Oncol TA - TT -. 2016;28(3):241-7.
226. Cardoso HJ, Carvalho TMA, Fonseca LRS, Figueira MI, Vaz C V, Socorro S. Revisiting prostate cancer metabolism: From metabolites to disease and therapy. Med Res Rev [Internet]. 2021 May 1;41(3):1499-538. Available from:
https://doi.org/10.1002/med. 21766
227. Rezende LP, Galheigo MRU, Landim BC, Cruz AR, Botelho FV, Zanon RG, et al. Effect of glucose and palmitate environment on proliferation and migration of PC3prostate cancer cells. Cell Biol Int [Internet]. 2019 Apr 1;43(4):373-83. Available from: https://doi.org/10.1002/cbin. 11066
228. Gonzalez-Menendez P, Hevia D, Alonso-Arias R, Alvarez-Artime A, RodriguezGarcia A, Kinet S, et al. GLUT1 protects prostate cancer cells from glucose deprivation-induced oxidative stress. Redox Biol [Internet]. 2018;17:112-27. Available from: https://www.sciencedirect.com/science/article/pii/S2213231718300703
229. Andrade RG, Ayala FR, Begnami MD, Cajaíba MM, Carvalho KC, Cunha IW, et al. GLUT1 expression in malignant tumors and its use as an immunodiagnostic marker. Clin TA - TT -. 2011;66(6):965-72.
230. Li Z, Liu H, Ju W, Xing Y, Zhang X, Yang J. LncRNA GASL1 inhibits growth and promotes expression of apoptosis-associated proteins in prostate carcinoma cells through GLUT-1. Oncol Lett [Internet]. 2019;17(6):5327-34. Available from: https://doi.org/10.3892/ol.2019.10244
231. Tian J, Guo F, Chen Y, Li Y, Yu B, Li Y. Nanoliposomal formulation encapsulating celecoxib and genistein inhibiting COX-2 pathway and Glut-1 receptors to prevent prostate cancer cell proliferation. Cancer Lett [Internet]. 2019;448:1-10. Available from: https://www.sciencedirect.com/science/article/pii/S030438351930014X
232. van der Laan SW, Fall T, Soumaré A, Teumer A, Sedaghat S, Baumert J, et al. Cystatin C and Cardiovascular Disease: A Mendelian Randomization Study. J Am Coll Cardiol. 2016;
233. Bok R, Lee J, Sriram R, Keshari K, Sukumar S, Daneshmandi S, et al. The Role of Lactate Metabolism in Prostate Cancer Progression and Metastases Revealed by DualAgent Hyperpolarized 13C MRSI. Vol. 11, Cancers . 2019.
234. van der Mijn JC, Kuiper MJ, Siegert CEH, Wassenaar AE, van Noesel CJM, Ogilvie AC. Lactic Acidosis in Prostate Cancer: Consider the Warburg Effect. Case Rep Oncol [Internet]. 2017;10(3):1085-91. Available from:
https://www.karger.com/DOI/10.1159/000485242
235. Sanità P, Capulli M, Teti A, Galatioto GP, Vicentini C, Chiarugi P, et al. Tumor-stroma metabolic relationship based on lactate shuttle can sustain prostate cancer progression. BMC Cancer [Internet]. 2014;14(1):154. Available from: https://doi.org/10.1186/1471-2407-14-154
236. Xian Z-Y, Liu J-M, Chen Q-K, Chen H-Z, Ye C-J, Xue J, et al. Inhibition of LDHA suppresses tumor progression in prostate cancer. Tumor Biol [Internet]. 2015;36(10):8093-100. Available from: https:/ / doi.org/ 10.1007/s13277-015-3540-x
237. Turner EL, Metcalfe C, Donovan JL, Noble S, Sterne JAC, Lane JA, et al. Design and preliminary recruitment results of the Cluster randomised triAl of PSA testing for Prostate cancer (CAP). Br J Cancer. 2014 Jun;110(12):2829-36.
238. Donovan J, Hamdy F, Neal D, Peters T, Oliver S, Brindle L, et al. Prostate testing for cancer and treatment (ProtecT) feasibility study. 2003;
239. Lane JA, Donovan JL, Davis M, Walsh E, Dedman D, Down L, et al. Active monitoring, radical prostatectomy, or radiotherapy for localised prostate cancer: study design and diagnostic and baseline results of the ProtecT randomised phase 3 trial. Lancet Oncol [Internet]. 2014;15(10):1109-18. Available from: https://www.sciencedirect.com/science/article/pii/S1470204514703614
240. Lane JA, Hamdy FC, Martin RM, Turner EL, Neal DE, Donovan JL. Latest results from the UK trials evaluating prostate cancer screening and treatment: the CAP and ProtecT studies. Eur J Cancer. 2010 Nov;46(17):3095-101.
241. Pasi S, J. KA, Peter W, Teemu S, Mika A-K, Soininen P, et al. Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Cardiovascular Epidemiology and Genetics. Circ Cardiovasc Genet [Internet]. 2015 Feb 1;8(1):192-206. Available from: https://doi.org/10.1161/CIRCGENETICS.114.000216
242. Julkunen H, Cichońska A, Slagboom PE, Würtz P, Initiative NHUKB. Metabolic biomarker profiling for identification of susceptibility to severe pneumonia and COVID-19 in the general population. Elife. 2021;10:e63033.
243. Würtz P, Wang Q, Kangas AJ, Richmond RC, Skarp J, Tiainen M, et al. Metabolic signatures of adiposity in young adults: Mendelian randomization analysis and effects of weight change. PLoS Med [Internet]. 2014 Dec 9;11(12):e1001765-e1001765. Available from: https:// pubmed.ncbi.nlm.nih.gov/25490400
244. Beynon R, Richmond RC, Santos Ferreira DL, Ness AR, May M, Davey Smith G, et al. Investigating the effects of lycopene and green tea on the metabolome of men at risk of prostate cancer: The ProDiet randomised controlled trial. Int J Cancer [Internet]. 2018;0(ja). Available from: https:/ / doi.org/10.1002/ijc. 31929
245. Everitt BS, Dunn G. Applied multivariate data analysis. Vol. 2. Wiley Online Library; 2001.
246. Mardia K V, Kent JT, Bibby JM, Bibby J. Multivariate analysis LK https://bris.on.worldcat.org/oclc/ 6164035 [Internet]. Probability and mathematical statistics TA - TT -. London ; SE - xv, 521 pages: illustrations, graphs, tables ; 23 cm : Academic Press; 1979. Available from:
http://catdir.loc.gov/catdir/toc/els031/79040922.html
247. Holmes E, Loo RL, Stamler J, Bictash M, Yap IKS, Chan Q, et al. Human metabolic phenotype diversity and its association with diet and blood pressure. Nature. 2008;
248. Gao X, Starmer J, Martin ER. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. Genet Epidemiol [Internet]. 2008 May 1;32(4):361-9. Available from:
https://doi.org/10.1002/gepi. 20310
249. Steuer R, Morgenthal K, Weckwerth W, Selbig J. A Gentle Guide to the Analysis of Metabolomic Data BT - Metabolomics: Methods and Protocols. In: Weckwerth W, editor. Totowa, NJ: Humana Press; 2007. p. 105-26. Available from:
https://doi.org/10.1007/978-1-59745-244-1_7
250. Wang Q, Kangas AJ, Soininen P, Tiainen M, Tynkkynen T, Puukka K, et al. Sex hormone-binding globulin associations with circulating lipids and metabolites and the risk for type 2 diabetes: observational and causal effect estimates. Int J Epidemiol [Internet]. 2015 Apr 1;44(2):623-37. Available from: https://doi.org/10.1093/ije/dyv093
251. Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis part I: basic concepts and first analyses. Br J Cancer [Internet]. 2003 Jul 21;89(2):232-8. Available from: https://pubmed.ncbi.nlm.nih.gov/12865907
252. Hosmer DW, Lemeshow S, May STA-TT-. Applied survival analysis : regression modeling of time-to-event data LK - https://bris.on.worldcat.org/oclc/761318476 [Internet]. 2nd ed. NV. Hoboken, N.J.: Wiley-Interscience; 2008. (Wiley series in probability and statistics). Available from:
http:// public.eblib.com/choice/publicfullrecord.aspx?p=700085_0
253. Goel MK, Khanna P, Kishore J. Understanding survival analysis: Kaplan-Meier estimate. Int J Ayurveda Res [Internet]. 2010 Oct;1(4):274-8. Available from: https:// pubmed.ncbi.nlm.nih.gov/21455458
254. Altman DG. Practical statistics for medical research LK https://bris.on.worldcat.org/oclc/ 23255428 [Internet]. 1st ed. Statistics texts. TA - TT -. London ; SE - xii, 611 pages : illustrations ; 24 cm : Chapman and Hall; 1991.
Available from: http://catdir.loc.gov/catdir/enhancements/ fy0646/91124890-d.html
255. Peto R, Peto J. Asymptotically Efficient Rank Invariant Test Procedures. J R Stat Soc Ser A [Internet]. 1972 Mar 2;135(2):185-207. Available from:
http://www.jstor.org/stable/2344317
256. In J, Lee DK. Survival analysis: part II - applied clinical data analysis. Korean J Anesthesiol [Internet]. 2019/05/17. 2019 Oct;72(5):441-57. Available from: https:// pubmed.ncbi.nlm.nih.gov/31096731
257. Kleinbaum DG, Klein MTA-TT-. Survival analysis : a self-learning text [Internet]. 2nd ed. NV. New York, NY: Springer; 2005. (Statistics for biology and health). Available from: http:// site.ebrary.com/id/10229012
258. Hancock MJ, Maher CG, Costa L da CM, Williams CM. A guide to survival analysis for manual therapy clinicians and researchers. Man Ther. 2014 Dec;19(6):511-6.
259. GRAMBSCH PM, THERNEAU TM. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika [Internet]. 1994 Sep 1;81(3):515-26. Available from: https://doi.org/10.1093/biomet/81.3.515
260. Yousefi PD, Suderman M, Langdon R, Whitehurst O, Davey Smith G, Relton CL. DNA methylation-based predictors of health: applications and statistical considerations. Nat Rev Genet. 2022;23(6):369-83.
261. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960;20(1):37-46.
262. Ferri C, Hernández-Orallo J, Modroiu R. An experimental comparison of performance measures for classification. Pattern Recognit Lett [Internet]. 2009;30(1):27-38. Available from:
https://www.sciencedirect.com/science/article/pii/S0167865508002687
263. Hajian-Tilaki K. Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. Casp J Intern Med [Internet]. 2013;4(2):627-35. Available from: https:/ / pubmed.ncbi.nlm.nih.gov/24009950
264. Van Calster B, Nieboer D, Vergouwe Y, De Cock B, Pencina MJ, Steyerberg EW. A calibration hierarchy for risk models was defined: from utopia to empirical data. J Clin Epidemiol. 2016;74:167-76.
265. McBride N, Yousefi P, Sovio U, Taylor K, Vafai Y, Yang T, et al. Do Mass Spectrometry-Derived Metabolomics Improve the Prediction of Pregnancy-Related Disorders? Findings from a UK Birth Cohort with Independent Validation. Vol. 11, Metabolites. 2021.
266. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW, Bossuyt P, et al. Calibration: the Achilles heel of predictive analytics. BMC Med [Internet]. 2019;17(1):230. Available from: https://doi.org/10.1186/s12916-019-1466-7
267. Davis SE, Lasko TA, Chen G, Siew ED, Matheny ME. Calibration drift in regression and machine learning models for acute kidney injury. J Am Med Informatics Assoc. 2017;24(6):1052-61.
268. Leijdekkers JA, Eijkemans MJC, Van Tilborg TC, Oudshoorn SC, McLernon DJ, Bhattacharya S, et al. Predicting the cumulative chance of live birth over multiple complete cycles of in vitro fertilization: an external validation study. Hum Reprod. 2018;33(9):1684-95.
269. Te Velde ER, Nieboer D, Lintsen AM, Braat DDM, Eijkemans MJC, Habbema JDF, et al. Comparison of two models predicting IVF success; the effect of time trends on model performance. Hum Reprod. 2014;29(1):57-64.
270. Thai TN, Ebell MH. Prospective validation of the Good Outcome Following Attempted Resuscitation (GO-FAR) score for in-hospital cardiac arrest prognosis. Resuscitation. 2019;140:2-8.
271. Deelen J, Kettunen J, Fischer K, van der Spek A, Trompet S, Kastenmüller G, et al. A metabolic profile of all-cause mortality risk identified in an observational study of 44,168 individuals. Nat Commun [Internet]. 2019;10(1):3346. Available from: https://doi.org/10.1038/s41467-019-11311-9
272. Peace KE. Statistical issues in drug research and development. Vol. 106. CRC Press; 1989.
273. Gupta SK. Intention-to-treat concept: A review. Perspect Clin Res [Internet]. 2011 Jul;2(3):109-12. Available from: https:// pubmed.ncbi.nlm.nih.gov/21897887
274. Sussman JB, Hayward RA. An IV for the RCT: using instrumental variables to adjust for treatment contamination in randomised controlled trials. BMJ [Internet]. 2010;340. Available from: https://www.bmj.com/content/340/bmj.c2073
275. Hadley J, Polsky D, Mandelblatt JS, Mitchell JM, Weeks JC, Wang Q, et al. An exploratory instrumental variable analysis of the outcomes of localized breast cancer treatments in a medicare population. Health Econ [Internet]. 2003 Mar 1;12(3):171-86. Available from: https://doi.org/10.1002/hec. 710
276. $\mathrm{Ng}^{\prime}$ Andu NH . An empirical comparison of statistical tests for assessing the proportional hazard assumption of Cox's model. Stat Med. 1997;16(6):611-26.
277. Bond SJ, White IR, Sarah Walker A. Instrumental variables and interactions in the causal analysis of a complex clinical trial. Stat Med [Internet]. 2007 Mar 30;26(7):147396. Available from: https://doi.org/10.1002/sim. 2644
278. Greenland S. An introduction to instrumental variables for epidemiologists. Int J Epidemiol [Internet]. 2000 Aug 1;29(4):722-9. Available from: https://doi.org/10.1093/ije/29.4.722
279. Davies NM, Holmes M V., Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. BMJ [Internet]. 2018 Jul 12;362:k601. Available from: http://www.bmj.com/content/362/bmj.k601
280. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology
contribute to understanding environmental determinants of disease? Int J Epidemiol [Internet]. 2003 Feb 1;32(1):1-22. Available from: https:// doi.org/10.1093/ije/dyg070
281. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med [Internet]. 2009;36(5):452-7. Available from: https://www.ncbi.nlm.nih.gov/pubmed/19362699
282. Smith GD, Ebrahim $S$. What can mendelian randomisation tell us about modifiable behavioural and environmental exposures? [Internet]. Vol. 330, BMJ. 2005. p. 1076-9. Available from: https:/ / doi.org/10.1136/bmj.330.7499.1076
283. Paternoster L, Tilling K, Davey Smith G. Genetic epidemiology and Mendelian randomization for informing disease therapeutics: Conceptual and methodological challenges. PLOS Genet [Internet]. 2017 Oct 5;13(10):e1006944. Available from: https://doi.org/10.1371/journal.pgen. 1006944
284. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. Stat Med [Internet]. 2008 Apr 15;27(8):1133-63. Available from: https:/ / doi.org/10.1002/ sim. 3034
285. Glymour MM, Tchetgen Tchetgen EJ, Robins JM. Credible Mendelian randomization studies: approaches for evaluating the instrumental variable assumptions. Am J Epidemiol [Internet]. 2012/01/12. 2012 Feb 15;175(4):332-9. Available from: https:// pubmed.ncbi.nlm.nih.gov/22247045
286. Lousdal ML. An introduction to instrumental variable assumptions, validation and estimation. Emerg Themes Epidemiol [Internet]. 2018;15:1. Available from: https://www.ncbi.nlm.nih.gov/pubmed/29387137
287. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol [Internet]. 2015;44(2):512-25. Available from: https:// pubmed.ncbi.nlm.nih.gov/26050253
288. Davey Smith G, Davies N, Dimou N, Egger M, Gallo V, Golub R, et al. STROBE-MR: Guidelines for strengthening the reporting of Mendelian randomization studies [Internet]. PeerJ Preprints; 2019. Available from:
http://europepmc.org/abstract/PPR/PPR85720
289. Sanderson E, Richardson TG, Hemani G, Davey Smith G. The use of negative control outcomes in Mendelian randomization to detect potential population stratification.
Int J Epidemiol [Internet]. 2021 Feb 11; Available from:
https://doi.org/10.1093/ije/dyaa288
290. Bowden J, Del Greco MF, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I2 statistic. Int J Epidemiol [Internet]. 2016;45(6):1961-74. Available from: https://doi.org/10.1093/ije/dyw220
291. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. Int J Epidemiol [Internet].

2017 Dec 1;46(6):1985-98. Available from: https://doi.org/10.1093/ije/dyx102
292. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol [Internet]. 2016;40(4):304-14. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4849733/
293. Burgess S, Thompson SG. Bias in causal estimates from Mendelian randomization studies with weak instruments. Stat Med TA - TT -. 2011;30(11):1312-23.
294. Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan N, Thompson J. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. Stat Med [Internet]. 2017 May 20;36(11):1783-802. Available from: https://doi.org/10.1002/sim. 7221
295. Hemani G, Bowden J, Davey Smith G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. Hum Mol Genet [Internet]. 2018 Aug 1;27(R2):R195-208. Available from: https://doi.org/10.1093/hmg/ddy163
296. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. Eur J Epidemiol [Internet]. 2017;32(5):377-89. Available from: https://doi.org/10.1007/ s10654-017-0255-x
297. Gkatzionis A, Burgess S. Contextualizing selection bias in Mendelian randomization: how bad is it likely to be? Int J Epidemiol [Internet]. 2019 Jun 1;48(3):691-701. Available from: https:/ / pubmed.ncbi.nlm.nih.gov/30325422
298. Smith GD, Ebrahim S. What can mendelian randomisation tell us about modifiable behavioural and environmental exposures? BMJ. 2005;330:1076-9.
299. Caramaschi D, Taylor AE, Richmond RC, Havdahl KA, Golding J, Relton CL, et al. Maternal smoking during pregnancy and autism: using causal inference methods in a birth cohort study. Transl Psychiatry. 2018 Nov;8(1):262.
300. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. Loos R, editor. Elife [Internet]. 2018;7:e34408. Available from: https://doi.org/10.7554/eLife. 34408
301. Lawlor DA. Commentary: Two-sample Mendelian randomization: opportunities and challenges. Int J Epidemiol [Internet]. 2016/07/17. 2016;45(3):908-15. Available from: https://www.ncbi.nlm.nih.gov/pubmed/27427429
302. VanderWeele TJ, Tchetgen Tchetgen EJ, Cornelis M, Kraft P. Methodological challenges in mendelian randomization. Epidemiology [Internet]. 2014 May;25(3):427-35. Available from: https:/ / pubmed.ncbi.nlm.nih.gov/24681576
303. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. Genet Epidemiol [Internet]. 2016 Nov 1;40(7):597-608. Available from: https://doi.org/10.1002/gepi. 21998
304. Smith GD, Timpson N, Ebrahim S. Strengthening causal inference in cardiovascular epidemiology through Mendelian randomization. Ann Med [Internet]. 2008 Jan

1;40(7):524-41. Available from: https://doi.org/10.1080/07853890802010709
305. Jacobs EJ, Stevens VL, Newton CC, Gapstur SM. Plasma total, LDL, and HDL cholesterol and risk of aggressive prostate cancer in the Cancer Prevention Study II Nutrition Cohort. Cancer Causes Control [Internet]. 2012;23(8):1289-96. Available from: https:// doi.org/10.1007/s10552-012-0006-y
306. Platz EA, Till C, Goodman PJ, Parnes HL, Figg WD, Albanes D, et al. Men with Low Serum Cholesterol Have a Lower Risk of High-Grade Prostate Cancer in the Placebo Arm of the Prostate Cancer Prevention Trial. Cancer Epidemiol Biomarkers \&amp; Prev [Internet]. 2009 Nov 1;18(11):2807 LP - 2813. Available from: http://cebp.aacrjournals.org/content/18/11/2807.abstract
307. Shafique K, McLoone P, Qureshi K, Leung H, Hart C, Morrison DS. Cholesterol and the risk of grade-specific prostate cancer incidence: evidence from two large prospective cohort studies with up to 37 years' follow up. BMC Cancer [Internet]. 2012;12(1):25. Available from: https://doi.org/10.1186/1471-2407-12-25
308. Bull CJ, Bonilla C, Holly JMP, Perks CM, Davies N, Haycock P, et al. Blood lipids and prostate cancer: a Mendelian randomization analysis. Cancer Med [Internet]. 2016 Jun 1;5(6):1125-36. Available from: https:/ / doi.org/10.1002/cam4.695
309. Miller GJ. Vitamin D and Prostate Cancer: Biologic Interactions and Clinical Potentials. Cancer Metastasis Rev [Internet]. 1998;17(4):353-60. Available from: https://doi.org/10.1023/A:1006102124548
310. Blutt SE, Weigel NL. Vitamin D and Prostate Cancer. Proc Soc Exp Biol Med [Internet]. 1999 Jun 1;221(2):89-98. Available from: https://doi.org/10.1046/j.1525-1373.1999.d01-60.x
311. Peehl DM, Feldman D. The role of vitamin D and retinoids in controlling prostate cancer progression. Endocrine-related cancer Endocr Relat Cancer Endocr Relat Cancer [Internet]. 2003;10(2):131-40. Available from: https://erc.bioscientifica.com/view/journals/erc/10/2/12790775.xml
312. Xu Y, Shibata A, McNeal JE, Stamey TA, Feldman D, Peehl DM. Vitamin D Receptor Start Codon Polymorphism (FokI) and Prostate Cancer Progression. Cancer Epidemiol Biomarkers \&amp; Prev [Internet]. 2003 Jan 1;12(1):23 LP - 27. Available from: http://cebp.aacrjournals.org/content/12/1/23.abstract
313. Trummer O, Langsenlehner U, Krenn-Pilko S, Pieber TR, Obermayer-Pietsch B, Gerger A, et al. Vitamin D and prostate cancer prognosis: a Mendelian randomization study. World J Urol [Internet]. 2016;34(4):607-11. Available from:
https://doi.org/10.1007/ s00345-015-1646-9
314. Zuccolo L, Lewis SJ, Donovan JL, Hamdy FC, Neal DE, Smith GD. Alcohol consumption and PSA-detected prostate cancer risk--a case-control nested in the ProtecT study. Int J cancer [Internet]. 2012/10/25. 2013 May 1;132(9):2176-85. Available from: https:// pubmed.ncbi.nlm.nih.gov/23024014
315. Sawada N, Inoue M, Iwasaki M, Sasazuki S, Yamaji T, Shimazu T, et al. Alcohol and smoking and subsequent risk of prostate cancer in Japanese men: the Japan Public

Health Center-based prospective study. Int J cancer. 2014 Feb;134(4):971-8.
316. Watters JL, Park Y, Hollenbeck A, Schatzkin A, Albanes D. Alcoholic beverages and prostate cancer in a prospective US cohort study. Am J Epidemiol. 2010 Oct;172(7):773-80.
317. Chao C, Haque R, Van Den Eeden SK, Caan BJ, Poon K-YT, Quinn VP. Red wine consumption and risk of prostate cancer: the California men's health study. Int J cancer. 2010 Jan;126(1):171-9.
318. Gong Z, Kristal AR, Schenk JM, Tangen CM, Goodman PJ, Thompson IM. Alcohol consumption, finasteride, and prostate cancer risk: results from the Prostate Cancer Prevention Trial. Cancer. 2009 Aug;115(16):3661-9.
319. Middleton Fillmore K, Chikritzhs T, Stockwell T, Bostrom A, Pascal R. Alcohol use and prostate cancer: a meta-analysis. Mol Nutr Food Res. 2009 Feb;53(2):240-55.
320. Brunner C, Davies NM, Martin RM, Eeles R, Easton D, Kote-Jarai Z, et al. Alcohol consumption and prostate cancer incidence and progression: A Mendelian randomisation study. Int J Cancer [Internet]. 2017 Jan 1;140(1):75-85. Available from: https://doi.org/10.1002/ijc. 30436
321. Collaboration Emerging Risk Factors. Adult height and the risk of cause-specific death and vascular morbidity in 1 million people: individual participant metaanalysis. Int J Epidemiol. 2012 Oct;41(5):1419-33.
322. Cao Y, Ma J. Body mass index, prostate cancer-specific mortality, and biochemical recurrence: a systematic review and meta-analysis. Cancer Prev Res (Phila). 2011 Apr;4(4):486-501.
323. Davies NM, Gaunt TR, Lewis SJ, Holly J, Donovan JL, Hamdy FC, et al. The effects of height and BMI on prostate cancer incidence and mortality: a Mendelian randomization study in 20,848 cases and 20,214 controls from the PRACTICAL consortium. Cancer Causes Control [Internet]. 2015 Oct 20;26(11):1603-16. Available from: http://www.jstor.org.bris.idm.oclc.org/stable/24716203
324. Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. Int J Epidemiol [Internet]. 2018 Feb 1;47(1):226-35. Available from: https://pubmed.ncbi.nlm.nih.gov/29040562
325. Watson JA, Leopold SJ, Simpson JA, Day NPJ, Dondorp AM, White NJ. Collider bias and the apparent protective effect of glucose-6-phosphate dehydrogenase deficiency on cerebral malaria. Lipsitch M, Ferguson NM, Murray E, editors. Elife [Internet]. 2019;8:e43154. Available from: https:/ / doi.org/10.7554/eLife. 43154
326. Mahmoud O, Dudbridge F, Smith GD, Munafo M, Tilling K. Slope-Hunter: A robust method for index-event bias correction in genome-wide association studies of subsequent traits. bioRxiv [Internet]. 2020 Jan 1;2020.01.31.928077. Available from: http://biorxiv.org/content/early/2020/01/31/2020.01.31.928077.abstract
327. Clarke GM, Rockett K, Kivinen K, Hubbart C, Jeffreys AE, Rowlands K, et al.

Characterisation of the opposing effects of G6PD deficiency on cerebral malaria and severe malarial anaemia. Elife. 2017;6:e15085.
328. Rockett KA, Clarke GM, Fitzpatrick K, Hubbart C, Jeffreys AE, Rowlands K, et al. Reappraisal of known malaria resistance loci in a large multicenter study. Nat Genet [Internet]. 2014;46(11):1197-204. Available from: https://doi.org/10.1038/ng. 3107
329. Dudbridge F, Allen RJ, Sheehan NA, Schmidt AF, Lee JC, Jenkins RG, et al. Adjustment for index event bias in genome-wide association studies of subsequent events. Nat Commun [Internet]. 2019 Apr 5;10(1):1561. Available from: https://pubmed.ncbi.nlm.nih.gov/30952951
330. Sandu, Meda R, Beynon, Rhona Richmond, Rebecca Santos Ferreira, Diana L, Hackshaw-McGeagh, Lucy, Davey Smith, George, Metcalfe Chis, Lane, J Athene, Martin R. Two-step Randomisation: Applying the Results of Small Feasibility Studies of Interventions to Large-scale Mendelian Randomisation Studies to Robustly Infer Causal Effects on Clinical Endpoints. Prerints. 2019;
331. Nichol AD, Bailey M, Cooper DJ. Challenging issues in randomised controlled trials. Injury [Internet]. 2010;41:S20-3. Available from: https://www.sciencedirect.com/science/article/pii/S0020138310002330
332. Hróbjartsson A, Boutron I. Blinding in Randomized Clinical Trials: Imposed Impartiality. Clin Pharmacol Ther [Internet]. 2011 Nov 1;90(5):732-6. Available from: https://doi.org/10.1038/clpt.2011.207
333. Hewitt CE, Torgerson DJ, Miles JN V. Is there another way to take account of noncompliance in randomized controlled trials? Can Med Assoc J [Internet]. 2006 Aug 15;175(4):347 LP - 347. Available from:
http://www.cmaj.ca/content/175/4/347.abstract
334. Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. Lancet [Internet]. 2002 Feb 16;359(9306):614. Available from: http://10.0.3.248/S0140-6736(02)07750-4
335. Clark L, Schmidt U, Tharmanathan P, Adamson J, Hewitt C, Torgerson D. Allocation concealment: a methodological review. J Eval Clin Pract [Internet]. 2013 Aug 1;19(4):708-12. Available from: https:/ / doi.org/10.1111/jep. 12032
336. Woolard RH, Carty K, Wirtz P, Longabaugh R, Nirenberg TD, Minugh PA, et al. Research Fundamentals: Follow-up of Subjects in Clinical Trials: Addressing Subject Attrition. Acad Emerg Med [Internet]. 2004 Aug 1;11(8):859-66. Available from: https://doi.org/10.1197/j.aem.2003.09.021
337. Robinson K, Allen F, Darby J, Fox C, Gordon AL, Horne JC, et al. Contamination in complex healthcare trials: the falls in care homes (FinCH) study experience. BMC Med Res Methodol [Internet]. 2020;20(1):46. Available from:
https://doi.org/10.1186/s12874-020-00925-z
338. Krishna R, Maithreyi R, Surapaneni KM. Research bias: a review for medical students. J Clin Diagn Res. 2010;4(2):2320-4.
339. Torgerson DJ, Torgerson C. Designing randomised trials in health, education and the social sciences : an introduction [Internet]. Basingstoke [England] ; Palgrave Macmillan; 2008. Available from:
http://www.dawsonera.com/depp/reader/protected/external/AbstractView/S978 0230583993
340. DeMets DL, Psaty BM, Fleming TR. When Can Intermediate Outcomes Be Used as Surrogate Outcomes? JAMA [Internet]. 2020 Mar 24;323(12):1184-5. Available from: https://doi.org/10.1001/jama.2020.1176
341. Prentice RL. Surrogate endpoints in clinical trials: Definition and operational criteria. Stat Med [Internet]. 1989 Apr 1;8(4):431-40. Available from: https://doi.org/10.1002/ sim. 4780080407
342. Denny JC, Ritchie MD, Basford MA, Pulley JM, Bastarache L, Brown-Gentry K, et al. PheWAS: demonstrating the feasibility of a phenome-wide scan to discover genedisease associations. Bioinformatics [Internet]. 2010 May;26(9):1205-10. Available from: https:/ / doi.org/10.1093/bioinformatics/btq126
343. Hebbring SJ. The challenges, advantages and future of phenome-wide association studies. Immunology [Internet]. 2014 Feb 1;141(2):157-65. Available from: https://doi.org/10.1111/imm. 12195
344. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. PLOS Med [Internet]. 2015 Mar 31;12(3):e1001779. Available from: https://doi.org/10.1371/journal.pmed. 1001779
345. Nguyen CD, Carlin JB, Lee KJ. Practical strategies for handling breakdown of multiple imputation procedures. Emerg Themes Epidemiol. 2021 Apr;18(1):5.
346. Statacorp. College Station, TX: StataCorp LLC.: Stata Statistical Software: Release 17; 2022.
347. Sterne JAC, Smith GD. Sifting the evidence - what's wrong with significance tests? Phys Ther [Internet]. 2001 Aug 1;81(8):1464-9. Available from: https://doi.org/10.1093/ptj/81.8.1464
348. Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. Nature Publishing Group; 2019.
349. Rosoff DB, Davey Smith G, Mehta N, Clarke T-K, Lohoff FW. Evaluating the relationship between alcohol consumption, tobacco use, and cardiovascular disease: A multivariable Mendelian randomization study. PLOS Med [Internet]. 2020 Dec 4;17(12):e1003410. Available from: https:// doi.org/10.1371/journal.pmed. 1003410
350. Adams CD, Richmond R, Ferreira DLS, Spiller W, Tan V, Zheng J, et al. Circulating Metabolic Biomarkers of Screen-Detected Prostate Cancer in the ProtecT Study. Cancer Epidemiol Biomarkers Prev. 2018/10/23. 2019 Jan;28(1):208-16.
351. Eeles RA, Olama AA Al, Benlloch S, Saunders EJ, Leongamornlert DA, Tymrakiewicz M , et al. Identification of 23 new prostate cancer susceptibility loci using the iCOGS
custom genotyping array. Nat Genet [Internet]. 2013;45(4):385-91. Available from: https://doi.org/10.1038/ng. 2560
352. Li J, Zhao Y, Zhang H, Hua W, Jiao W, Du X, et al. Contribution of Rs780094 and Rs1260326 Polymorphisms in GCKR Gene to Non-alcoholic Fatty Liver Disease: A Meta-Analysis Involving 26,552 Participants. Endocr Metab Immune Disord Drug Targets. 2021;21(9):1696-708.
353. Liu Z, Suo C, Shi O, Lin C, Zhao R, Yuan H, et al. The Health Impact of MAFLD, a Novel Disease Cluster of NAFLD, Is Amplified by the Integrated Effect of Fatty Liver Disease-Related Genetic Variants. Clin Gastroenterol Hepatol Off Clin Pract J Am Gastroenterol Assoc. 2022 Apr;20(4):e855-75.
354. Zahedi AS, Akbarzadeh M, Sedaghati-Khayat B, Seyedhamzehzadeh A, Daneshpour MS. GCKR common functional polymorphisms are associated with metabolic syndrome and its components: a 10-year retrospective cohort study in Iranian adults. Diabetol Metab Syndr. 2021 Feb;13(1):20.
355. Zhang C, Bao W, Rong Y, Yang H, Bowers K, Yeung E, et al. Genetic variants and the risk of gestational diabetes mellitus: a systematic review. Hum Reprod Update. 2013;19(4):376-90.
356. Lin Z, Wang Y, Zhang B, Jin Z. Association of type 2 diabetes susceptible genes GCKR, SLC30A8, and FTO polymorphisms with gestational diabetes mellitus risk: a meta-analysis. Endocrine [Internet]. 2018;62(1):34-45. Available from: https://doi.org/10.1007/s12020-018-1651-z
357. Wang Q, Tiffen J, Bailey CG, Lehman ML, Ritchie W, Fazli L, et al. Targeting Amino Acid Transport in Metastatic Castration-Resistant Prostate Cancer: Effects on Cell Cycle, Cell Growth, and Tumor Development. JNCI J Natl Cancer Inst [Internet]. 2013 Oct 2;105(19):1463-73. Available from: https:// doi.org/10.1093/jnci/djt241
358. Cacciatore S, Wium M, Licari C, Ajayi-Smith A, Masieri L, Anderson C, et al. Inflammatory metabolic profile of South African patients with prostate cancer. Cancer Metab [Internet]. 2021;9(1):29. Available from: https://doi.org/10.1186/s40170-021-00265-6
359. Hirasawa N. Expression of Histidine Decarboxylase and Its Roles in Inflammation. Vol. 20, International Journal of Molecular Sciences . 2019.
360. Niu Y-C, Feng R-N, Hou Y, Li K, Kang Z, Wang J, et al. Histidine and arginine are associated with inflammation and oxidative stress in obese women. Br J Nutr [Internet]. 2011/10/14. 2012;108(1):57-61. Available from: https://www.cambridge.org/core/article/histidine-and-arginine-are-associated-with-inflammation-and-oxidative-stress-in-obesewomen/21EC7FF249722367B3BA23C5F0B9D470
361. Putluri N, Shojaie A, Vasu VT, Nalluri S, Vareed SK, Putluri V, et al. Metabolomic Profiling Reveals a Role for Androgen in Activating Amino Acid Metabolism and Methylation in Prostate Cancer Cells. PLoS One [Internet]. 2011 Jul 18;6(7):e21417. Available from: https://doi.org/10.1371/journal.pone. 0021417
362. Kettunen J, Demirkan A, Würtz P, Draisma HHM, Haller T, Rawal R, et al. Genomewide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. Nat Commun [Internet]. 2016;7(1):11122. Available from: https://doi.org/10.1038/ncomms11122
363. Jia X, Chen J, Sun S, Yang W, Yang S, Shah P, et al. Detection of aggressive prostate cancer associated glycoproteins in urine using glycoproteomics and mass spectrometry. Proteomics. 2016 Dec;16(23):2989-96.
364. Liu Y, Chen J, Sethi A, Li QK, Chen L, Collins B, et al. Glycoproteomic analysis of prostate cancer tissues by SWATH mass spectrometry discovers N -acylethanolamine acid amidase and protein tyrosine kinase 7 as signatures for tumor aggressiveness. Mol Cell Proteomics. 2014 Jul;13(7):1753-68.
365. Cima I, Schiess R, Wild P, Kaelin M, Schüffler P, Lange V, et al. Cancer geneticsguided discovery of serum biomarker signatures for diagnosis and prognosis of prostate cancer. Proc Natl Acad Sci U S A. 2011 Feb;108(8):3342-7.
366. Kawahara R, Recuero S, Srougi M, Leite KRM, Thaysen-Andersen M, Palmisano G. The Complexity and Dynamics of the Tissue Glycoproteome Associated With Prostate Cancer Progression. Mol Cell Proteomics [Internet]. 2021 Jan 1;20. Available from: https:// doi.org/10.1074/mcp.RA120.002320
367. Williams N, Hughes LJ, Turner EL, Donovan JL, Hamdy FC, Neal DE, et al. Prostatespecific antigen testing rates remain low in UK general practice: a cross-sectional study in six English cities. BJU Int [Internet]. 2011 Nov 1;108(9):1402-8. Available from: https:/ / doi.org/10.1111/j.1464-410X.2011.10163.x
368. Lane JA, Er V, Avery KNL, Horwood J, Cantwell M, Caro GP, et al. ProDiet: A Phase II Randomized Placebo-controlled Trial of Green Tea Catechins and Lycopene in Men at Increased Risk of Prostate Cancer. Cancer Prev Res [Internet]. 2018;11(11):687. Available from:
http://cancerpreventionresearch.aacrjournals.org/content/11/11/687.abstract
369. Markley JL, Brüschweiler R, Edison AS, Eghbalnia HR, Powers R, Raftery D, et al. The future of NMR-based metabolomics. Curr Opin Biotechnol. 2017 Feb;43:34-40.
370. Sanderson E, Davey Smith G, Windmeijer F, Bowden J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. Int J Epidemiol [Internet]. 2019 Jun 1;48(3):713-27. Available from: https:/ / pubmed.ncbi.nlm.nih.gov/30535378
371. Holmes M V, Ala-Korpela M, Smith GD. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. Nat Rev Cardiol [Internet]. 2017/06/01. 2017;14(10):577-90. Available from: https://www.ncbi.nlm.nih.gov/pubmed/28569269
372. Richardson TG, Leyden GM, Wang Q, Bell JA, Elsworth B, Smith GD, et al. Metabolomic signatures of lipid-modifying therapies using drug target Mendelian randomization. medRxiv [Internet]. 2021 Jan 1;2021.08.06.21261699. Available from: http://medrxiv.org/content/early/2021/08/07/2021.08.06.21261699.abstract
373. Bull CJ, Bell JA, Murphy N, Sanderson E, Davey Smith G, Timpson NJ, et al. Adiposity, metabolites, and colorectal cancer risk: Mendelian randomization study. BMC Med [Internet]. 2020;18(1):396. Available from: https:// doi.org/10.1186/s12916-020-01855-9
374. Hamdy FC. The Prostate Testing for Cancer and Treatment (ProtecT) study: what have we learnt? Vol. 118, BJU international. England; 2016. p. 843.
375. Beynon RA, Richmond RC, Santos Ferreira DL, Ness AR, May M, Smith GD, et al. Investigating the effects of lycopene and green tea on the metabolome of men at risk of prostate cancer: The ProDiet randomised controlled trial. Int J Cancer [Internet]. 2019 Apr 15;144(8):1918-28. Available from: https://doi.org/10.1002/ijc. 31929
376. OLMEDILLA B, GRANADO F, SOUTHON S, WRIGHT AJA, BLANCO I, GILMARTINEZ E, et al. A European multicentre, placebo-controlled supplementation study with a-tocopherol, carotene-rich palm oil, lutein or lycopene: analysis of serum responses. Clin Sci [Internet]. 2002 Mar 18;102(4):447-56. Available from: https://doi.org/10.1042/cs1020447
377. Fenglei W, Jusheng Z, Bo Y, Jiajing J, Yuanqing F, Duo L. Effects of Vegetarian Diets on Blood Lipids: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. J Am Heart Assoc [Internet]. 2020 Apr 24;4(10):e002408. Available from: https://doi.org/10.1161/JAHA.115.002408
378. Yokoyama Y, Levin SM, Barnard ND. Association between plant-based diets and plasma lipids: a systematic review and meta-analysis. Nutr Rev [Internet]. 2017 Aug 21;75(9):683-98. Available from: https://doi.org/10.1093/nutrit/nux030
379. Huang B, Song B, Xu C. Cholesterol metabolism in cancer: mechanisms and therapeutic opportunities. Nat Metab [Internet]. 2020;2(2):132-41. Available from: https://doi.org/10.1038/ s42255-020-0174-0
380. Lee HJ, Li J, Vickman RE, Li J, Liu R, Durkes AC, et al. Cholesterol Esterification Inhibition Suppresses Prostate Cancer Metastasis by Impairing the Wnt/ $\beta$-catenin Pathway. Mol Cancer Res [Internet]. 2018 Jun 1;16(6):974 LP - 985. Available from: http://mcr.aacrjournals.org/content/16/6/974.abstract
381. Narayanan S. The Preanalytic Phase: An Important Component of Laboratory Medicine. Am J Clin Pathol [Internet]. 2000 Mar 1;113(3):429-52. Available from: https://doi.org/10.1309/C0NM-Q7R0-LL2E-B3UY
382. Turchiano M, Nguyen C, Fierman A, Lifshitz M, Convit A. Impact of Blood Sample Collection and Processing Methods on Glucose Levels in Community Outreach Studies. Gerber LM, editor. J Environ Public Health [Internet]. 2013;2013:256151. Available from: https://doi.org/10.1155/2013/256151
383. Ding M, Zeleznik OA, Guasch-Ferre M, Hu J, Lasky-Su J, Lee I-M, et al. MetabolomeWide Association Study of the Relationship Between Habitual Physical Activity and Plasma Metabolite Levels. Am J Epidemiol [Internet]. 2019 Jul 31;188(11):1932-43. Available from: https://doi.org/10.1093/aje/kwz171
384. Pang Y, Bennett DA, Holmes M V, Kartsonaki C, Du H, Millwood IY, et al. The
association of physical activity with plasma lipoproteins and inflammation measured by NMR-metabolomics: evidence from the China Kadoorie Biobank study. Eur Heart J [Internet]. 2017 Aug 29;38(suppl_1). Available from: https://doi.org/10.1093/eurheartj/ ehx504.4158
385. Ikonen E. Cellular cholesterol trafficking and compartmentalization. Nat Rev Mol Cell Biol [Internet]. 2008;9(2):125-38. Available from: https:// doi.org/10.1038/nrm2336
386. Wang S, Wu J, Suburu J, Gu Z, Cai J, Axanova LS, et al. Effect of dietary polyunsaturated fatty acids on castration-resistant Pten-null prostate cancer. Carcinogenesis [Internet]. 2012 Feb 1;33(2):404-12. Available from: https://doi.org/10.1093/carcin/bgr290
387. Bratton BA, Maly I V, Hofmann WA. Effect of polyunsaturated fatty acids on proliferation and survival of prostate cancer cells. PLoS One [Internet]. 2019 Jul 17;14(7):e0219822. Available from: https:// doi.org/10.1371/journal.pone. 0219822
388. Liu J, Hu S, Cui Y, Sun M-K, Xie F, Zhang Q, et al. Saturated fatty acids up-regulate COX-2 expression in prostate epithelial cells via toll-like receptor 4/NF-KB signaling. Inflammation. 2014;37(2):467-77.
389. Huang M, Koizumi A, Narita S, Inoue T, Tsuchiya N, Nakanishi H, et al. Diet-induced alteration of fatty acid synthase in prostate cancer progression. Oncogenesis [Internet]. 2016;5(2):e195-e195. Available from: https://doi.org/10.1038/oncsis.2015.42
390. Skvortsov S, Skvortsova I-I, Tang DG, Dubrovska A. Concise Review: Prostate Cancer Stem Cells: Current Understanding. Stem Cells [Internet]. 2018 Oct 1;36(10):1457-74. Available from: https://doi.org/10.1002/stem. 2859
391. Wang X, Sun B, Wei L, Jian X, Shan K, He Q, et al. Cholesterol and saturated fatty acids synergistically promote the malignant progression of prostate cancer. Neoplasia [Internet]. 2022;24(2):86-97. Available from: https://www.sciencedirect.com/science/article/pii/S1476558621000944
392. Kettunen J, Tukiainen T, Sarin A-PP, Ortega-Alonso A, Tikkanen E, Lyytikäinen L-PP, et al. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. Nat Genet [Internet]. 2012;44:269. Available from:
https://doi.org/10.1038/ng. 1073
393. Rao A V, Agarwal S. Bioavailability and in vivo antioxidant properties of lycopene from tomato products and their possible role in the prevention of cancer. Nutr Cancer [Internet]. 1998 Jan 1;31(3):199-203. Available from:
https://doi.org/10.1080/01635589809514703
394. Kumar NB, Besterman-Dahan K, Kang L, Pow-Sang J, Xu P, Allen K, et al. Results of a Randomized Clinical Trial of the Action of Several Doses of Lycopene in Localized Prostate Cancer: Administration Prior to Radical Prostatectomy. Clin Med Urol [Internet]. 2008 Jan 1;1:CMU.S718. Available from: https://doi.org/10.4137/CMU.S718
395. Johnson EJ, Qin J, Krinsky NI, Russell RM. Ingestion by Men of a Combined Dose of
$\beta$-Carotene and Lycopene Does Not Affect the Absorption of $\beta$-Carotene but Improves That of Lycopene. J Nutr [Internet]. 1997 Sep 1;127(9):1833-7. Available from: https:// doi.org/10.1093/jn/127.9.1833
396. Swelam BA, Verswijveren SJJM, Salmon J, Arundell L, Ridgers ND. Exploring activity compensation amongst youth and adults: a systematic review. Int J Behav Nutr Phys Act [Internet]. 2022;19(1):25. Available from: https:// doi.org/10.1186/s12966-022-01264-6
397. Felgueiras J, Silva JV, Fardilha M. Prostate cancer: the need for biomarkers and new therapeutic targets. J Zhejiang Univ Sci B [Internet]. 2014;15(1):16-42. Available from: https://doi.org/10.1631/jzus.B1300106
398. Sridhar SS, Freedland SJ, Gleave ME, Higano C, Mulders P, Parker C, et al. Castration-Resistant Prostate Cancer: From New Pathophysiology to New Treatment. Eur Urol [Internet]. 2014;65(2):289-99. Available from:
http://www.sciencedirect.com/science/article/pii/S0302283813008294
399. Bonkhoff H, Berges R. From pathogenesis to prevention of castration resistant prostate cancer. Prostate [Internet]. 2010 Jan 1;70(1):100-12. Available from: https://doi.org/10.1002/ pros. 21042
400. Gleason DF. Histologic grading and clinical staging of prostate carcinoma. Urol Pathol prostate. 1977;
401. Fischer K, Kettunen J, Würtz P, Haller T, Havulinna AS, Kangas AJ, et al. Biomarker Profiling by Nuclear Magnetic Resonance Spectroscopy for the Prediction of AllCause Mortality: An Observational Study of 17,345 Persons. PLOS Med [Internet]. 2014 Feb 25;11(2):e1001606. Available from:
https://doi.org/10.1371/journal.pmed. 1001606
402. Julkunen H, Cichońska A, Tiainen M, Koskela H, Nybo K, Mäkelä V, et al. Atlas of plasma nuclear magnetic resonance biomarkers for health and disease in 118,461 individuals from the UK Biobank. medRxiv [Internet]. 2022 Jan 1;2022.06.13.22276332. Available from:
http://medrxiv.org/content/early/2022/06/15/2022.06.13.22276332.abstract
403. Bonilla C, Lewis SJ, Martin RM, Donovan JL, Hamdy FC, Neal DE, et al. Pubertal development and prostate cancer risk: Mendelian randomization study in a population-based cohort. BMC Med [Internet]. 2016;14(1):66. Available from: https://doi.org/10.1186/s12916-016-0602-x
404. Turner EL, Metcalfe C, Donovan JL, Noble S, Sterne JAC, Lane JA, et al. Contemporary accuracy of death certificates for coding prostate cancer as a cause of death: Is reliance on death certification good enough? A comparison with blinded review by an independent cause of death evaluation committee. Br J Cancer. 2016;115(1):90-4.
405. Benesty J, Chen J, Huang Y, Cohen I. Pearson Correlation Coefficient BT - Noise Reduction in Speech Processing. In: Cohen I, Huang Y, Chen J, Benesty J, editors. Berlin, Heidelberg: Springer Berlin Heidelberg; 2009. p. 1-4. Available from:
https://doi.org/10.1007/978-3-642-00296-0_5
406. Merriel SWD, Ingle SM, May MT, Martin RM. Retrospective cohort study evaluating clinical, biochemical and pharmacological prognostic factors for prostate cancer progression using primary care data. BMJ Open [Internet]. 2021 Feb 1;11(2):e044420. Available from: http://bmjopen.bmj.com/content/11/2/e044420.abstract
407. Walz J, Gallina A, Saad F, Montorsi F, Perrotte P, Shariat SF, et al. A Nomogram Predicting 10-Year Life Expectancy in Candidates for Radical Prostatectomy or Radiotherapy for Prostate Cancer. J Clin Oncol [Internet]. 2007 Aug 20;25(24):3576-81. Available from: https:// doi.org/10.1200/JCO.2006.10.3820
408. Kerkmeijer LGW, Monninkhof EM, van Oort IM, van der Poel HG, de Meerleer G, van Vulpen M. PREDICT: model for prediction of survival in localized prostate cancer. World J Urol [Internet]. 2016;34(6):789-95. Available from:
https://doi.org/10.1007/s00345-015-1691-4
409. Stark JR, Perner S, Stampfer MJ, Sinnott JA, Finn S, Eisenstein AS, et al. Gleason score and lethal prostate cancer: does $3+4=4+3$ ? J Clin Oncol Off J Am Soc Clin Oncol. 2009 Jul;27(21):3459-64.
410. Meyer H-J, Wienke A, Surov A. Discrimination between clinical significant and insignificant prostate cancer with apparent diffusion coefficient - a systematic review and meta analysis. BMC Cancer. 2020 May;20(1):482.
411. Chuang C-K, Wang T-J, Yeung C-Y, Hsieh W-S, Lin D-S, Ho S-C, et al. Interference and blood sample preparation for a pyruvate enzymatic assay. Clin Biochem [Internet]. 2006;39(1):74-7. Available from: https://www.sciencedirect.com/science/ article/pii/S0009912005002924
412. Yin P, Lehmann R, Xu G. Effects of pre-analytical processes on blood samples used in metabolomics studies. Anal Bioanal Chem [Internet]. 2015;407(17):4879-92. Available from: https:/ / doi.org/10.1007/s00216-015-8565-x
413. Burgess S, Thompson SG, Collaboration CRPCHDG. Avoiding bias from weak instruments in Mendelian randomization studies. Int J Epidemiol [Internet]. 2011 Jun 1;40(3):755-64. Available from: https:/ / doi.org/10.1093/ije/ dyr036
414. Burgess S, Butterworth A, Thompson SG. Mendelian Randomization Analysis With Multiple Genetic Variants Using Summarized Data. Genet Epidemiol [Internet]. 2013 Nov 1;37(7):658-65. Available from: https://doi.org/10.1002/gepi.21758
415. Brumpton B, Sanderson E, Heilbron K, Hartwig FP, Harrison S, Vie GA, et al. Avoiding dynastic, assortative mating, and population stratification biases in Mendelian randomization through within-family analyses. Nat Commun [Internet]. 2020;11(1):3519. Available from: https://doi.org/10.1038/s41467-020-17117-4
416. Lawson DJ, Davies NM, Haworth S, Ashraf B, Howe L, Crawford A, et al. Is population structure in the genetic biobank era irrelevant, a challenge, or an opportunity? Hum Genet [Internet]. 2020;139(1):23-41. Available from: https://doi.org/10.1007/ s00439-019-02014-8
417. Hartwig FP, Davies NM, Hemani G, Davey Smith G. Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. Int J Epidemiol [Internet]. 2017 Mar 28;45(6):1717-26. Available from: https://doi.org/10.1093/ije/dyx028
418. Cortés J, González JA, Medina MN, Vogler M, Vilaró M, Elmore M, et al. Does evidence support the high expectations placed in precision medicine? A bibliographic review [version 5; peer review: 2 approved, 1 approved with reservations, 3 not approved]. 2019;7(30). Available from: http:/ /openr.es/gj9
419. Winkelbeiner S, Leucht S, Kane JM, Homan P. Evaluation of Differences in Individual Treatment Response in Schizophrenia Spectrum Disorders: A MetaanalysisEvaluation of Treatment Response Differences in SchizophreniaEvaluation of Treatment Response Differences in Schizophrenia. JAMA Psychiatry [Internet]. 2019; Available from: https://doi.org/10.1001/jamapsychiatry.2019.1530
420. Senn S. Mastering variation: variance components and personalised medicine. Stat Med [Internet]. 2016;35(7):966-77. Available from: https://doi.org/10.1002/sim. 6739

## Chapter 9. Appendices

### 9.1. Appendix A

Appendix Table A 1: Baseline assessment variable in UKBB.

| Questionnaire and interview |  |
| :---: | :---: |
| Sociodemographic | Social class; ethnicity; employment status; marital status; education; income; car ownership |
| Family history and early life exposures | Family history of major diseases; birth weight; breast feeding; maternal smoking; childhood body size; residence at birth |
| Psychosocial factors | Neurosis; depression (including bi-polar spectrum disorder); social support |
| Environmental factors | Current address; current (or last) occupation; domestic heating and cooking fuel; housing; means of travel; shift work; mobile phone use; sun exposure |
| Lifestyle | Smoking; alcohol consumption; physical activity; diet; sleep |
| Health status | Medical history; medications; disability; hearing; sight; sexual and reproductive history |
| Hearing threshold | Speech reception threshold* |
| Cognitive function | Pairs matching; reaction time; prospective memory*; fluid intelligence*; numeric memory ${ }^{\dagger}$ |
| Physical measures |  |
| Blood pressure and heart rate | two automated measures, one minute apart |
| Grip strength | Left- and right-hand grip strength |
| Anthropometrics | Standing and sitting height; weight and bio-impedance; hip and waist circumference |
| Spirometry | Up to three measures |
| Bone density ${ }^{\ddagger}$ | Calcaneal ultrasound |
| Arterial stiffness" | Pulse wave velocity |
| Eye examination ${ }^{\text {§ }}$ | Refractive index, intraocular pressure; acuity; retinal photograph; optical coherence tomography |
| Fitness test ${ }^{\text {¢ }}$ | Cycle ergometry with electrocardiogram (ECG) heart rate monitoring |
| * assessed in 170,000 participants; |  |
| ${ }^{\dagger}$ assessed in 50,000 participants; |  |
| ${ }^{\ddagger}$ measured in one heel for 170,000 participants and in both heels for 320,000 participants; |  |
| ${ }^{7}$ measured in 170,000 participants; |  |
| ${ }^{\text {§ }}$ measured in 100,000 participants |  |

doi:10.1371/jjournal.pmed. 1001779.1002
Adapted from Sudlow et al "UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age"

Appendix Table A 2: Current and future enhanced phenotyping variables in UKBB

| Data type | Number of participants | Details | Date of data acquisition | Date first available for research ${ }^{\ddagger}$ |
| :---: | :---: | :---: | :---: | :---: |
| Baseline assessment | Whole cohort | Questionnaire, physical measures, samples (see Table 2); haematological assays done on fresh blood samples | 2006-2010 | Q2 2012 |
| Repeat of baseline assessment | 20,000-25,000 | As above every few years, to allow correction for regression dilution due to measurement error and within person fluctuations in exposure levels [12]. | 2013- | Q3 2013 |
| Biochemical assays (of baseline samples) | Whole cohort | Biomarkers with known disease associations (e.g., lipids for vascular disease), diagnostic value (e.g., $\mathrm{HbA}_{1 \mathrm{c}}$ for diabetes), or ability to characterize phenotypes not otherwise well assessed (e.g., renal and liver function tests). | 2014-2015 | 2015 |
| Genotyping (of baseline samples) | Whole cohort | Dense genotyping chip with $>800,000$ markers including: approximately 250,000 SNPs in a whole-genome array; approximately 200,000 markers covering CNV, loss of function, insertions, deletions, and previously identified risk factor or disease associations; approximately 150,000 exome markers covering a high proportion of nonsynonymous coding variants with allele frequency $>0.02 \%$. | 2013-2015 | 2015 |
| Dietary Web questionnaire | 210,000 | Automatically coded dietary recall questionnaire, providing estimates of nutrient intake. 80,000 respondents completed it $\geq$ three times. | 2011-2012 | Q2 2013 |
| Other Web questionnaires | 350,000 to be approached | Participants invited by email to provide additional information via Web questionnaires about exposures (e.g., occupation) and health outcomes (cognitive function, depression) that are not readily identified from health record linkages. | 2014- | 2015 |
| Accelerometry | 100,000 | Wrist-worn tri-axial accelerometers record information on type, intensity, and duration of physical activity. | 2013-2015 | 2015 |
| Multimodal imaging | 100,000 | MRI brain, heart, and abdomen (for lipid distribution); ultrasound of carotid arteries; whole body DXA scan of bones and joints | Pilot phase: 20142015 Main phase: 2016-2019 | 2015 |
| Health record linkage | Whole cohort |  |  |  |
| Death registrations |  | ICD-coded cause specific mortality | 2006- | Q2 2013 |
| Cancer registrations |  | ICD-coded cancer diagnoses | 1971-* | Q2 2013 |
| Hospital inpatient episodes |  | ICD-coded diagnoses, OPCS-coded procedures | 1997-* | Q4 2013 |
| Hospital outpatient episodes |  | Limited ICD and OPCS coding | 2003-* | 2015 |
| Primary care |  | Read-coded information including diagnoses, measurements, referrals, prescriptions | Variable | 2015 |
| Other |  | UK Biobank will obtain data from national mental health care, residential history, laboratory and disease audit datasets and is considering the value of further linkages (e. g., imaging, cancer screening, dental). | Variable | Not yet determined |
| Adjudicated health outcomes | Whole cohort | Expert-led confirmation and subclassification of outcomes in a range of disease areas, including cancer, diabetes, heart disease, stroke, mental health, musculoskeletal, respiratory, neurodegenerative, and ocular disorders. |  | 2015 |

Hb: haemoglobin; SNPs: single nucleotide polymorphisms; CNV: copy number variations; MRI: magnetic resonance imaging; DXA: dual-energy X-ray absorptiometry; ICD: International Classification of Diseases; OPCS: Office of Population Censuses and Surveys Classification of Interventions and Procedures

* Future dates are estimated. Data available may be all or part of the relevant dataset.
* available from an earlier date from health record systems in Scotland
doi:10.1371/journal.pmed.1001779.t003

Adapted from Sudlow et al "UK Biobank: An Open Access Resource for Identifying the
Causes of a Wide Range of Complex Diseases of Middle and Old Age"

Appendix Table A 3: List of metabolites investigated in UKB


| M_HDL_P | MUFA_pct | S_LDL_CE | Total_FC |
| :---: | :---: | :---: | :---: |
| M_HDL_PL | non_HDL_C | S_LDL_CE_pct | Total_L |
| M_HDL_PL_pct | Omega_3 | S_LDL_FC | Total_P |
| M_HDL_TG | Omega_3_pct | S_LDL_FC_pct | Total_PL |
| M_HDL_TG_pct | Omega_6 | S_LDL_L | Total_TG |
| M_LDL_C | Omega_6_by_Omeg | S_LDL_P | Tyr |
| M_LDL_C_pct | a_3 | S_LDL_PL | Unsaturation |
| M_LDL_CE | Omega_6_pct | S_LDL_PL_pct | Val |
| M_LDL_CE_pct | Phe | S_LDL_TG | VLDL_C |
| M_LDL_FC | Phosphatidylc | S_LDL_TG_pct | VLDL_CE |
| M_LDL_FC_pct | Phosphoglyc | S_VLDL_C | VLDL_FC |
| M_LDL_L | PUFA | S_VLDL_C_pct | VLDL_L |
| M_LDL_P | PUFA_by_MUFA | S_VLDL_CE | VLDL_P |
| M_LDL_PL | PUFA_pct | S_VLDL_CE_pct | VLDL_PL |
| M_LDL_PL_pct | Pyruvate | S_VLDL_FC | VLDL_size |
| M_LDL_TG | Remnant_C | S_VLDL_FC_pct | VLDL_TG |
| M_LDL_TG_pct | S_HDL_C | S_VLDL_L | XL_HDL_C |
| M_VLDL_C | S_HDL_C_pct | S_VLDL_P | XL_HDL_C_pct |
| M_VLDL_C_pct | S_HDL_CE | S_VLDL_PL | XL_HDL_CE |
| M_VLDL_CE | S_HDL_CE_pct | S_VLDL_PL_pct | XL_HDL_CE_pct |
| M_VLDL_CE_pct | S_HDL_FC | S_VLDL_TG | XL_HDL_FC |
| M_VLDL_FC | S_HDL_FC_pct | S_VLDL_TG_pct | XL_HDL_FC_pct |
| M_VLDL_FC_pct | S_HDL_L | SFA | XL_HDL_L |
| M_VLDL_L | S_HDL_P | SFA_pct | XL_HDL_P |
| M_VLDL_P | S_HDL_PL | Sphingomyelins | XL_HDL_PL |
| M_VLDL_PL | S_HDL_PL_pct | TG_by_PG | XL_HDL_PL_pct |
| M_VLDL_PL_pct | S_HDL_TG | Total_BCAA | XL_HDL_TG |
| M_VLDL_TG | S_HDL_TG_pct | Total_C | XL_HDL_TG_pct |
| M_VLDL_TG_pct | S_LDL_C | Total_CE | XL_VLDL_C |
| MUFA | S_LDL_C_pct | Total_FA | XL_VLDL_C_pct |

XL_VLDL_CE
XL_VLDL_CE_pct
XL_VLDL_FC
XL_VLDL_FC_pct
XL_VLDL_L
XL_VLDL_P
XL_VLDL_PL
XL_VLDL_PL_pct
XL_VLDL_TG
XL_VLDL_TG_pct
XS_VLDL_C
XS_VLDL_C_pct
XS_VLDL_CE
XS_VLDL_CE_pct
XS_VLDL_FC
XS_VLDL_FC_pct
XS_VLDL_L
XS_VLDL_P
XS_VLDL_PL
XS_VLDL_PL_pct
XS_VLDL_TG
XS_VLDL_TG_pct
XXL_VLDL_C
XXL_VLDL_C_pct
XXL_VLDL_CE

XXL_VLDL_CE_pct
XXL_VLDL_FC
XXL_VLDL_FC_pct
XXL_VLDL_L
XXL_VLDL_P

XXL_VLDL_PL
XXL_VLDL_PL_pct
XXL_VLDL_TG
XXL_VLDL_TG_pct

Appendix Table A 4: List of studies included in the PRACTICAL consortium

|  |  | No PC |  |  | Inflation |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Study population | No. PC | deaths | Genotyping platform | No. Imputed SNPs | factor |
| ICOGS_CAPS | 354 | 44 | Custom Illumina Infinium (iCOGS) | 9786797 | 0.971 |
| ICOGS_CPCS | 4505 | 95 | Custom Illumina Infinium (iCOGS) | 9750186 | 1.003 |
| ICOGS_EPIC | 1170 | 35 | Custom Illumina Infinium (iCOGS) | 9799859 | 0.965 |
| ICOGS_FHCRC | 752 | 59 | Custom Illumina Infinium (iCOGS) | 9778209 | 0.979 |
| ICOGS_MAYO | 762 | 39 | Custom Illumina Infinium (iCOGS) | 9768079 | 0.961 |
| ICOGS_MCCS_PCFS | 1568 | 104 | Custom Illumina Infinium (iCOGS) | 9788606 | 1.066 |
| ICOGS_SEARCH | 1258 | 114 | Custom Illumina Infinium (iCOGS) | 9754581 | 1.003 |
| ICOGS_STHM1 | 2300 | 157 | Custom Illumina Infinium (iCOGS) | 9808465 | 0.986 |
| ICOGS_TAMPERE | 2385 | 241 | Custom Illumina Infinium (iCOGS) | 9855776 | 1.021 |
| ICOGS_UKGPCS | 156 | 26 | Custom Illumina Infinium (iCOGS) | 9763468 | 0.962 |
| ICOGS_ULM | 330 | 31 | Custom Illumina Infinium (iCOGS) | 9755747 | 0.947 |
| ONCO_AARHUS | 698 | 65 | Illumina OncoArray | 10001398 | 0.981 |
| ONCO_AHS | 491 | 23 | Illumina OncoArray | 10009305 | 0.798 |
| ONCO_ATBC | 1227 | 227 | Illumina OncoArray | 10085518 | 0.997 |
| ONCO_COSM | 2112 | 264 | Illumina OncoArray | 10022517 | 1.000 |
| ONCO_EPIC | 626 | 27 | Illumina OncoArray | 10051515 | 0.842 |
| ONCO_FHCRC | 421 | 31 | Illumina OncoArray | 10079015 | 0.880 |
| ONCO_HPFS | 1178 | 75 | Illumina OncoArray | 10093241 | 0.955 |
| ONCO_MCCS | 705 | 80 | Illumina OncoArray | 10037960 | 0.987 |
| ONCO_MEC | 637 | 29 | Illumina OncoArray | 10133224 | 0.864 |
| ONCO_OSLO | 1455 | 764 | Illumina OncoArray | 11323855 | 0.972 |
| ONCO_PHS | 640 | 118 | Illumina OncoArray | 10134011 | 1.008 |
| ONCO_PROCAP | 614 | 213 | Illumina OncoArray | 10004073 | 1.017 |
| ONCO_PROGRESS | 624 | 20 | Illumina OncoArray | 10213332 | 0.780 |
| ONCO_PROMPT | 57 | 41 | Illumina OncoArray | 9993073 | 1.028 |
| ONCO_QLD | 2992 | 54 | Illumina OncoArray | 10036463 | 0.935 |
| ONCO_SEARCH | 2564 | 150 | Illumina OncoArray | 10053660 | 1.021 |


| ONCO_SFPCS | 280 | 50 | Illumina OncoArray | 10054687 | 0.923 |
| :---: | :---: | :---: | :---: | :---: | :---: |
| ONCO_STHM2 | 3116 | 181 | Illumina OncoArray | 10052592 | 0.976 |
| ONCO_TAMPERE | 2108 | 146 | Illumina OncoArray | 10084392 | 1.028 |
| ONCO_UKGPCS1 | 6502 | 1476 | Illumina OncoArray | 10117802 | 1.029 |
| ONCO_UKGPCS2 | 5552 | 969 | Illumina OncoArray | 10400039 | 1.026 |
| GWAS_CAPS1 | 492 | 214 | Affymetrix GeneChip 500K | 10578874 | 1.019 |
| GWAS_CAPS2 | 1493 | 331 | Affymetrix GeneChip 5.0K | 10492976 | 1.019 |
| GWAS_BPC3_ATBC | 245 | 133 | Illumina Human610 Illumina 610K | 9268134 | 1.036 |
| GWAS_BPC3_CPSII | 636 | 79 | Illumina Human610 Illumina 610K | 8814571 | 1.031 |
| GWAS_BPC3_EPIC | 431 | 159 | Illumina Human610 Illumina 610K | 8887241 | 1.029 |
| GWAS_BPC3_HPFS | 214 | 37 | Illumina Human610 Illumina 610K | 8843698 | 1.071 |
| GWAS_BPC3_MEC | 244 | 23 | Illumina Human610 Illumina 610K | 8880354 | 1.155 |
| GWAS_BPC3_PHS | 298 | 97 | Illumina Human610 Illumina 610K | 8868315 | 1.030 |
| GWAS_MDC | 1093 | 263 | Illumina OmniExpress Exome BeadChip | 11154114 | 1.017 |
| GWAS_PEGASUS | 4643 | 234 | Human Omni 2.5 | 10878513 | 1.001 |
| GWAS_UKB | 7830 | 396 | Affymetrix UK Biobank Axiom array | 9857769 | 0.991 |
| Total | 67758 | 7914 |  |  |  |

Appendix Table A 5: Threats to robust inference when conducting two-sample MR

| Threats to |  |  |  |
| :---: | :---: | :---: | :---: |
| Mendelian randomisation | Definition | Methods to mitigate | References |
|  |  | Covariate adjustment |  |
| Weak <br> instrument bias | Bias generated by an instrument that is not strongly associated with the phenotype | Limited information maximum likelihood (LIML) | (285,293,302,413) |

Fuller (1) methods

The non-random distribution of genetic variants at different loci in the population

Population stratification implications

Collider bias which can overstate precision in MR studies that use methodologies of summarised data

When in the structure of the studied population there is a correlation between the distribution of genetic variants and that of the exposures/outcomes

Not applicable

Fixed family effects

Difference estimators

Negative control outcome
analysis

Adjusting genetic
associations
Collider bias occurs when conditioning on a variable that is a common effect of two variables. In MR collider bias results in violations of the instrumental variable assumption. This is particularly problematic
$(302,414)$ in disease progression studies, where risk factors are associated both with the development and progression of disease and can lead to under or over identification of progression genetic risk factors

Bias generated by overlapping of samples in the instrument-exposure and instrumentSample overlap outcome datasets, which can lead to biased causal estimates between the exposure and outcome.
(280,285,293,302,303,417)
(283,297,324-326,329)

Slope-Hunter

## Derive risk factor

associations in a nonoverlapping dataset

Apply equal or externally specified weights in the IVW method


### 9.2. Appendix B

Appendix B Table B 1: Cox regression results for clinical progression minimally adjusted model, fully adjusted model, censored minimally adjusted model

| alb | Minimally adjusted model |  |  |  |  | Fully adjusted model |  |  |  |  | Minimally adjusted model (Complete case) |  |  |  |  | Minimally adjusted model censored |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 1825 | 0.92 | 0.80 | 1.06 | 0.23 | 1229 | 0.90 | 0.76 | 1.05 | 0.19 | 1229 | 0.89 | 0.76 | 1.05 | 0.18 | 1825 | 0.91 | 0.79 | 1.05 | 0.21 |
| I_hdl_c | 1824 | 0.92 | 0.80 | 1.06 | 0.26 | 1228 | 0.89 | 0.75 | 1.06 | 0.20 | 1228 | 0.88 | 0.74 | 1.05 | 0.16 | 1824 | 0.92 | 0.79 | 1.06 | 0.25 |
| I_hdl_pl | 1825 | 0.92 | 0.80 | 1.06 | 0.26 | 1229 | 0.87 | 0.73 | 1.04 | 0.14 | 1229 | 0.87 | 0.73 | 1.03 | 0.10 | 1825 | 0.91 | 0.79 | 1.06 | 0.23 |
| gly | 1817 | 0.92 | 0.80 | 1.07 | 0.28 | 1221 | 1.00 | 0.83 | 1.19 | 0.96 | 1221 | 0.98 | 0.82 | 1.17 | 0.82 | 1817 | 0.94 | 0.81 | 1.09 | 0.41 |
| I_hdl_ce | 1824 | 0.92 | 0.80 | 1.07 | 0.28 | 1228 | 0.89 | 0.75 | 1.07 | 0.22 | 1228 | 0.89 | 0.75 | 1.05 | 0.17 | 1824 | 0.92 | 0.79 | 1.07 | 0.26 |
| I_hdi_I | 1824 | 0.92 | 0.80 | 1.07 | 0.28 | 1228 | 0.89 | 0.74 | 1.06 | 0.18 | 1228 | 0.88 | 0.74 | 1.04 | 0.14 | 1824 | 0.92 | 0.79 | 1.06 | 0.26 |
| 1_hdl_p | 1824 | 0.93 | 0.80 | 1.07 | 0.29 | 1228 | 0.89 | 0.74 | 1.06 | 0.18 | 1228 | 0.88 | 0.74 | 1.04 | 0.14 | 1824 | 0.92 | 0.79 | 1.07 | 0.27 |
| sm_ldi_fc | 1825 | 0.93 | 0.80 | 1.07 | 0.29 | 1229 | 0.83 | 0.70 | 0.99 | 0.04 | 1229 | 0.82 | 0.69 | 0.98 | 0.03 | 1825 | 0.91 | 0.78 | 1.05 | 0.20 |
| idl_fc | 1825 | 0.93 | 0.80 | 1.07 | 0.30 | 1229 | 0.84 | 0.70 | 1.00 | 0.05 | 1229 | 0.83 | 0.70 | 0.99 | 0.03 | 1825 | 0.91 | 0.79 | 1.06 | 0.23 |
| I_IdI_fc | 1825 | 0.93 | 0.80 | 1.07 | 0.30 | 1229 | 0.83 | 0.70 | 0.99 | 0.04 | 1229 | 0.83 | 0.69 | 0.99 | 0.03 | 1825 | 0.91 | 0.79 | 1.06 | 0.22 |
| s_ldi_c | 1825 | 0.93 | 0.81 | 1.07 | 0.32 | 1229 | 0.83 | 0.70 | 0.99 | 0.04 | 1229 | 0.82 | 0.69 | 0.98 | 0.03 | 1825 | 0.91 | 0.79 | 1.06 | 0.23 |
| s_ldi_ce | 1825 | 0.93 | 0.81 | 1.08 | 0.35 | 1229 | 0.83 | 0.70 | 0.99 | 0.04 | 1229 | 0.83 | 0.69 | 0.98 | 0.03 | 1825 | 0.92 | 0.79 | 1.06 | 0.26 |
| m_ldi_c | 1825 | 0.94 | 0.81 | 1.08 | 0.37 | 1229 | 0.84 | 0.70 | 1.00 | 0.05 | 1229 | 0.83 | 0.70 | 0.99 | 0.04 | 1825 | 0.92 | 0.79 | 1.07 | 0.28 |
| apoal | 1825 | 0.94 | 0.81 | 1.08 | 0.37 | 1229 | 0.83 | 0.70 | 0.99 | 0.04 | 1229 | 0.83 | 0.70 | 0.98 | 0.03 | 1825 | 0.92 | 0.79 | 1.06 | 0.24 |
| s_\|di_| | 1825 | 0.94 | 0.81 | 1.08 | 0.38 | 1229 | 0.84 | 0.70 | 1.00 | 0.04 | 1229 | 0.83 | 0.70 | 0.99 | 0.04 | 1825 | 0.92 | 0.79 | 1.07 | 0.28 |
| m_ldi_ce | 1825 | 0.94 | 0.81 | 1.08 | 0.40 | 1229 | 0.84 | 0.71 | 1.00 | 0.05 | 1229 | 0.83 | 0.70 | 0.99 | 0.04 | 1825 | 0.92 | 0.80 | 1.07 | 0.30 |
| IdI_c | 1825 | 0.94 | 0.81 | 1.09 | 0.40 | 1229 | 0.84 | 0.71 | 1.00 | 0.05 | 1229 | 0.83 | 0.70 | 0.99 | 0.04 | 1825 | 0.92 | 0.80 | 1.07 | 0.30 |
| s_ldi_p | 1825 | 0.94 | 0.82 | 1.09 | 0.41 | 1229 | 0.84 | 0.71 | 1.00 | 0.05 | 1229 | 0.83 | 0.70 | 0.99 | 0.04 | 1825 | 0.93 | 0.80 | 1.07 | 0.31 |
| m_ldi_\| | 1825 | 0.94 | 0.82 | 1.09 | 0.43 | 1229 | 0.85 | 0.71 | 1.01 | 0.06 | 1229 | 0.84 | 0.70 | 1.00 | 0.05 | 1825 | 0.93 | 0.80 | 1.08 | 0.33 |
| idl_pl | 1825 | 0.94 | 0.82 | 1.09 | 0.44 | 1229 | 0.85 | 0.71 | 1.01 | 0.07 | 1229 | 0.84 | 0.71 | 1.01 | 0.06 | 1825 | 0.93 | 0.80 | 1.08 | 0.36 |
| unsat | 1823 | 0.95 | 0.82 | 1.09 | 0.43 | 1227 | 0.94 | 0.79 | 1.12 | 0.50 | 1227 | 0.92 | 0.78 | 1.09 | 0.35 | 1823 | 0.94 | 0.81 | 1.09 | 0.42 |
| s_\|dil_pl | 1825 | 0.95 | 0.82 | 1.09 | 0.44 | 1229 | 0.84 | 0.71 | 1.00 | 0.05 | 1229 | 0.84 | 0.70 | 0.99 | 0.04 | 1825 | 0.93 | 0.80 | 1.07 | 0.31 |
| estc | 1824 | 0.95 | 0.82 | 1.09 | 0.45 | 1228 | 0.84 | 0.70 | 0.99 | 0.04 | 1228 | 0.83 | 0.70 | 0.98 | 0.03 | 1824 | 0.93 | 0.80 | 1.08 | 0.34 |
| m_ldi_p | 1825 | 0.95 | 0.82 | 1.09 | 0.46 | 1229 | 0.85 | 0.71 | 1.01 | 0.06 | 1229 | 0.84 | 0.71 | 1.00 | 0.05 | 1825 | 0.93 | 0.80 | 1.08 | 0.36 |
| I_ldi_c | 1825 | 0.95 | 0.82 | 1.09 | 0.46 | 1229 | 0.85 | 0.71 | 1.01 | 0.06 | 1229 | 0.84 | 0.70 | 1.00 | 0.05 | 1825 | 0.93 | 0.80 | 1.08 | 0.36 |
| serum_c | 1825 | 0.95 | 0.82 | 1.10 | 0.49 | 1229 | 0.84 | 0.71 | 1.00 | 0.05 | 1229 | 0.83 | 0.70 | 0.99 | 0.04 | 1825 | 0.94 | 0.81 | 1.08 | 0.38 |
| ala | 1825 | 0.95 | 0.83 | 1.10 | 0.49 | 1229 | 0.96 | 0.81 | 1.14 | 0.63 | 1229 | 0.96 | 0.81 | 1.14 | 0.63 | 1825 | 0.98 | 0.85 | 1.13 | 0.77 |


| I_IdI_I | Minimally adjusted model |  |  |  |  | Fully adjusted model |  |  |  |  | Minimally adjusted model (Complete case) |  |  |  |  | Minimally adjusted model censored |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 1825 | 0.95 | 0.82 | 1.10 | 0.51 | 1229 | 0.85 | 0.72 | 1.02 | 0.08 | 1229 | 0.85 | 0.71 | 1.01 | 0.07 | 1825 | 0.94 | 0.81 | 1.09 | 0.41 |
| I_\|di_pl | 1825 | 0.95 | 0.82 | 1.10 | 0.51 | 1229 | 0.85 | 0.71 | 1.01 | 0.07 | 1229 | 0.85 | 0.71 | 1.01 | 0.06 | 1825 | 0.94 | 0.81 | 1.09 | 0.39 |
| I_IdI_ce | 1825 | 0.95 | 0.83 | 1.10 | 0.51 | 1229 | 0.85 | 0.71 | 1.01 | 0.07 | 1229 | 0.85 | 0.71 | 1.01 | 0.06 | 1825 | 0.94 | 0.81 | 1.09 | 0.40 |
| idl_c | 1825 | 0.95 | 0.83 | 1.10 | 0.52 | 1229 | 0.85 | 0.72 | 1.02 | 0.08 | 1229 | 0.85 | 0.71 | 1.01 | 0.07 | 1825 | 0.94 | 0.81 | 1.09 | 0.42 |
| m_hdl_fc | 1825 | 0.96 | 0.83 | 1.10 | 0.54 | 1229 | 0.89 | 0.75 | 1.06 | 0.20 | 1229 | 0.88 | 0.74 | 1.05 | 0.16 | 1825 | 0.93 | 0.81 | 1.08 | 0.37 |
| I_IdI_p | 1825 | 0.96 | 0.83 | 1.10 | 0.54 | 1229 | 0.86 | 0.72 | 1.02 | 0.09 | 1229 | 0.85 | 0.71 | 1.02 | 0.07 | 1825 | 0.94 | 0.81 | 1.09 | 0.44 |
| pyr | 1823 | 0.96 | 0.82 | 1.11 | 0.56 | 1227 | 0.93 | 0.74 | 1.17 | 0.53 | 1227 | 0.93 | 0.74 | 1.17 | 0.54 | 1823 | 0.97 | 0.84 | 1.13 | 0.73 |
| freec | 1824 | 0.96 | 0.83 | 1.11 | 0.57 | 1228 | 0.85 | 0.71 | 1.01 | 0.07 | 1228 | 0.84 | 0.71 | 1.01 | 0.06 | 1824 | 0.95 | 0.82 | 1.10 | 0.47 |
| m_\|di_pl | 1825 | 0.96 | 0.83 | 1.11 | 0.58 | 1229 | 0.86 | 0.72 | 1.02 | 0.08 | 1229 | 0.85 | 0.72 | 1.02 | 0.08 | 1825 | 0.94 | 0.81 | 1.09 | 0.44 |
| idl_I | 1825 | 0.96 | 0.83 | 1.11 | 0.58 | 1229 | 0.86 | 0.73 | 1.03 | 0.10 | 1229 | 0.86 | 0.72 | 1.03 | 0.09 | 1825 | 0.95 | 0.82 | 1.10 | 0.49 |
| val | 1824 | 0.96 | 0.84 | 1.11 | 0.58 | 1229 | 1.00 | 0.85 | 1.18 | 0.99 | 1229 | 1.02 | 0.87 | 1.20 | 0.78 | 1824 | 0.94 | 0.81 | 1.09 | 0.40 |
| sm | 1824 | 0.96 | 0.83 | 1.11 | 0.59 | 1228 | 0.88 | 0.75 | 1.04 | 0.14 | 1228 | 0.88 | 0.74 | 1.04 | 0.12 | 1824 | 0.94 | 0.82 | 1.09 | 0.44 |
| idl_p | 1825 | 0.96 | 0.84 | 1.11 | 0.63 | 1229 | 0.87 | 0.73 | 1.04 | 0.12 | 1229 | 0.87 | 0.73 | 1.03 | 0.11 | 1825 | 0.96 | 0.82 | 1.11 | 0.55 |
| idl_ce | 1825 | 0.97 | 0.84 | 1.12 | 0.64 | 1229 | 0.86 | 0.72 | 1.03 | 0.10 | 1229 | 0.86 | 0.72 | 1.03 | 0.10 | 1825 | 0.95 | 0.82 | 1.10 | 0.52 |
| m_hdl_c | 1825 | 0.97 | 0.84 | 1.12 | 0.64 | 1229 | 0.89 | 0.75 | 1.07 | 0.22 | 1229 | 0.89 | 0.75 | 1.06 | 0.18 | 1825 | 0.94 | 0.81 | 1.09 | 0.42 |
| m_hdl_ce | 1825 | 0.97 | 0.84 | 1.12 | 0.69 | 1229 | 0.90 | 0.75 | 1.07 | 0.24 | 1229 | 0.89 | 0.75 | 1.06 | 0.20 | 1825 | 0.94 | 0.81 | 1.10 | 0.45 |
| bohbut | 1777 | 0.98 | 0.86 | 1.12 | 0.79 | 1195 | 1.01 | 0.88 | 1.16 | 0.86 | 1195 | 1.01 | 0.88 | 1.15 | 0.89 | 1777 | 0.97 | 0.84 | 1.11 | 0.64 |
| xs_vidl_pl | 1825 | 0.99 | 0.86 | 1.14 | 0.87 | 1229 | 0.89 | 0.75 | 1.06 | 0.20 | 1229 | 0.89 | 0.74 | 1.06 | 0.18 | 1825 | 0.98 | 0.85 | 1.14 | 0.82 |
| m_hdl_I | 1825 | 0.99 | 0.86 | 1.14 | 0.87 | 1229 | 0.91 | 0.76 | 1.09 | 0.31 | 1229 | 0.91 | 0.76 | 1.08 | 0.27 | 1825 | 0.97 | 0.83 | 1.12 | 0.63 |
| m_hdl_pl | 1824 | 0.99 | 0.86 | 1.14 | 0.88 | 1229 | 0.92 | 0.77 | 1.09 | 0.34 | 1229 | 0.91 | 0.77 | 1.08 | 0.30 | 1824 | 0.97 | 0.84 | 1.12 | 0.67 |
| s_hdl_l | 1825 | 0.99 | 0.86 | 1.14 | 0.90 | 1229 | 0.93 | 0.78 | 1.10 | 0.39 | 1229 | 0.92 | 0.78 | 1.10 | 0.36 | 1825 | 0.96 | 0.83 | 1.11 | 0.59 |
| I_hdl_tg | 1824 | 0.99 | 0.86 | 1.14 | 0.90 | 1228 | 0.97 | 0.83 | 1.14 | 0.74 | 1228 | 0.97 | 0.82 | 1.13 | 0.68 | 1824 | 1.00 | 0.87 | 1.15 | 0.99 |
| xs_vidl_fc | 1824 | 0.99 | 0.86 | 1.15 | 0.93 | 1228 | 0.91 | 0.77 | 1.09 | 0.32 | 1228 | 0.90 | 0.76 | 1.08 | 0.25 | 1824 | 0.99 | 0.86 | 1.15 | 0.94 |
| m_hdl_p | 1825 | 0.99 | 0.86 | 1.15 | 0.94 | 1229 | 0.92 | 0.77 | 1.10 | 0.35 | 1229 | 0.91 | 0.77 | 1.09 | 0.31 | 1825 | 0.97 | 0.84 | 1.12 | 0.70 |
| dha | 1823 | 0.99 | 0.87 | 1.14 | 0.94 | 1227 | 1.06 | 0.91 | 1.25 | 0.44 | 1227 | 1.05 | 0.90 | 1.23 | 0.52 | 1823 | 1.00 | 0.87 | 1.15 | 0.99 |
| leu | 1825 | 1.00 | 0.87 | 1.15 | 0.96 | 1229 | 1.02 | 0.86 | 1.21 | 0.80 | 1229 | 1.04 | 0.89 | 1.22 | 0.61 | 1825 | 0.98 | 0.85 | 1.14 | 0.82 |
| totcho | 1824 | 1.00 | 0.87 | 1.15 | 0.97 | 1228 | 0.90 | 0.76 | 1.06 | 0.20 | 1228 | 0.89 | 0.76 | 1.05 | 0.17 | 1824 | 0.98 | 0.85 | 1.14 | 0.83 |


|  | Minimally adjusted model |  |  |  |  | Fully adjusted model |  |  |  |  | Minimally adjusted model (Complete case) |  |  |  |  | Minimally adjusted model censored |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| gic | 1819 | 1.00 | 0.88 | 1.14 | 0.96 | 1225 | 0.95 | 0.80 | 1.13 | 0.54 | 1225 | 0.97 | 0.82 | 1.14 | 0.68 | 1819 | 1.01 | 0.88 | 1.16 | 0.87 |
| sfa | 1823 | 1.00 | 0.87 | 1.16 | 0.96 | 1227 | 1.02 | 0.86 | 1.21 | 0.81 | 1227 | 1.04 | 0.88 | 1.23 | 0.65 | 1823 | 1.00 | 0.86 | 1.16 | 0.98 |
| faw6_fa | 1823 | 1.00 | 0.87 | 1.16 | 0.96 | 1227 | 0.89 | 0.75 | 1.06 | 0.18 | 1227 | 0.89 | 0.75 | 1.05 | 0.16 | 1823 | 0.99 | 0.86 | 1.15 | 0.93 |
| la_fa | 1823 | 1.00 | 0.87 | 1.16 | 0.95 | 1227 | 0.90 | 0.76 | 1.06 | 0.21 | 1227 | 0.89 | 0.75 | 1.06 | 0.19 | 1823 | 1.00 | 0.86 | 1.15 | 0.97 |
| s_hdl_p | 1825 | 1.00 | 0.87 | 1.16 | 0.95 | 1229 | 0.94 | 0.79 | 1.12 | 0.48 | 1229 | 0.94 | 0.79 | 1.11 | 0.46 | 1825 | 0.98 | 0.84 | 1.13 | 0.74 |
| xs_vidl_c | 1825 | 1.01 | 0.87 | 1.16 | 0.92 | 1229 | 0.93 | 0.78 | 1.11 | 0.41 | 1229 | 0.92 | 0.78 | 1.10 | 0.38 | 1825 | 1.01 | 0.87 | 1.17 | 0.93 |
| pc | 1824 | 1.01 | 0.88 | 1.16 | 0.92 | 1228 | 0.89 | 0.76 | 1.06 | 0.19 | 1228 | 0.89 | 0.75 | 1.05 | 0.17 | 1824 | 1.00 | 0.86 | 1.15 | 0.95 |
| tyr | 1820 | 1.01 | 0.88 | 1.16 | 0.91 | 1225 | 1.00 | 0.84 | 1.18 | 0.97 | 1225 | 1.01 | 0.86 | 1.19 | 0.93 | 1820 | 0.99 | 0.86 | 1.15 | 0.94 |
| faw3 | 1823 | 1.01 | 0.88 | 1.16 | 0.90 | 1227 | 1.07 | 0.91 | 1.25 | 0.42 | 1227 | 1.06 | 0.90 | 1.24 | 0.51 | 1823 | 1.00 | 0.87 | 1.16 | 0.95 |
| s_hdl_fc | 1825 | 1.01 | 0.88 | 1.16 | 0.90 | 1229 | 0.98 | 0.82 | 1.16 | 0.77 | 1229 | 0.97 | 0.82 | 1.15 | 0.74 | 1825 | 0.99 | 0.86 | 1.14 | 0.86 |
| m_\|di_tg | 1824 | 1.01 | 0.88 | 1.16 | 0.88 | 1228 | 0.96 | 0.81 | 1.13 | 0.60 | 1228 | 0.95 | 0.80 | 1.13 | 0.58 | 1824 | 1.02 | 0.88 | 1.18 | 0.77 |
| pufa_fa | 1823 | 1.01 | 0.88 | 1.17 | 0.88 | 1227 | 0.91 | 0.77 | 1.08 | 0.28 | 1227 | 0.91 | 0.76 | 1.07 | 0.25 | 1823 | 1.00 | 0.86 | 1.16 | 0.99 |
| xs_vidl_ce | 1825 | 1.01 | 0.88 | 1.17 | 0.85 | 1229 | 0.94 | 0.79 | 1.12 | 0.48 | 1229 | 0.94 | 0.79 | 1.12 | 0.48 | 1825 | 1.01 | 0.87 | 1.18 | 0.86 |
| his | 1820 | 1.02 | 0.88 | 1.17 | 0.83 | 1225 | 1.07 | 0.91 | 1.26 | 0.43 | 1225 | 1.08 | 0.92 | 1.27 | 0.36 | 1820 | 0.99 | 0.86 | 1.15 | 0.90 |
| phe | 1825 | 1.02 | 0.89 | 1.16 | 0.80 | 1229 | 1.03 | 0.88 | 1.22 | 0.68 | 1229 | 1.05 | 0.89 | 1.23 | 0.58 | 1825 | 1.03 | 0.90 | 1.18 | 0.67 |
| totpg | 1824 | 1.02 | 0.89 | 1.17 | 0.80 | 1228 | 0.91 | 0.77 | 1.08 | 0.28 | 1228 | 0.91 | 0.77 | 1.07 | 0.25 | 1824 | 1.00 | 0.87 | 1.16 | 0.95 |
| xs_vldl_1 | 1824 | 1.03 | 0.89 | 1.19 | 0.68 | 1228 | 0.95 | 0.80 | 1.13 | 0.59 | 1228 | 0.95 | 0.80 | 1.13 | 0.56 | 1824 | 1.03 | 0.89 | 1.20 | 0.66 |
| ile | 1825 | 1.03 | 0.90 | 1.18 | 0.67 | 1229 | 1.06 | 0.90 | 1.25 | 0.46 | 1229 | 1.08 | 0.93 | 1.27 | 0.32 | 1825 | 1.01 | 0.88 | 1.17 | 0.87 |
| dha_fa | 1823 | 1.03 | 0.90 | 1.18 | 0.66 | 1227 | 1.03 | 0.88 | 1.20 | 0.74 | 1227 | 1.02 | 0.87 | 1.20 | 0.76 | 1823 | 1.03 | 0.90 | 1.18 | 0.68 |
| I_\|di_tg | 1824 | 1.03 | 0.90 | 1.19 | 0.67 | 1228 | 0.99 | 0.84 | 1.17 | 0.90 | 1228 | 0.99 | 0.84 | 1.17 | 0.89 | 1824 | 1.04 | 0.90 | 1.20 | 0.56 |
| IdI_tg | 1825 | 1.03 | 0.90 | 1.19 | 0.65 | 1229 | 0.98 | 0.83 | 1.16 | 0.83 | 1229 | 0.98 | 0.83 | 1.16 | 0.84 | 1825 | 1.04 | 0.90 | 1.21 | 0.55 |
| ace | 1825 | 1.03 | 0.91 | 1.17 | 0.59 | 1229 | 1.01 | 0.87 | 1.18 | 0.90 | 1229 | 1.02 | 0.88 | 1.19 | 0.81 | 1825 | 0.99 | 0.84 | 1.17 | 0.91 |
| apob_apoal | 1824 | 1.04 | 0.90 | 1.20 | 0.60 | 1228 | 0.94 | 0.80 | 1.12 | 0.50 | 1228 | 0.95 | 0.80 | 1.12 | 0.55 | 1824 | 1.03 | 0.89 | 1.19 | 0.67 |
| crea | 1820 | 1.04 | 0.91 | 1.19 | 0.57 | 1225 | 0.99 | 0.84 | 1.17 | 0.94 | 1225 | 0.98 | 0.83 | 1.16 | 0.84 | 1820 | 1.04 | 0.90 | 1.20 | 0.60 |
| xs_vldi_p | 1824 | 1.04 | 0.90 | 1.20 | 0.58 | 1228 | 0.97 | 0.81 | 1.15 | 0.69 | 1228 | 0.96 | 0.81 | 1.14 | 0.68 | 1824 | 1.04 | 0.90 | 1.21 | 0.56 |
| remnant_c | 1825 | 1.04 | 0.90 | 1.20 | 0.56 | 1229 | 0.95 | 0.80 | 1.13 | 0.58 | 1229 | 0.96 | 0.81 | 1.13 | 0.61 | 1825 | 1.04 | 0.90 | 1.20 | 0.60 |
| faw3_fa | 1823 | 1.04 | 0.91 | 1.19 | 0.54 | 1227 | 1.03 | 0.87 | 1.20 | 0.75 | 1227 | 1.02 | 0.87 | 1.20 | 0.77 | 1823 | 1.04 | 0.90 | 1.19 | 0.61 |


| glc | Minimally adjusted model |  |  |  |  | Fully adjusted model |  |  |  |  | Minimally adjusted model (Complete case) |  |  |  |  | Minimally adjusted model censored |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 1819 | 1.00 | 0.88 | 1.14 | 0.96 | 1225 | 0.95 | 0.80 | 1.13 | 0.54 | 1225 | 0.97 | 0.82 | 1.14 | 0.68 | 1819 | 1.01 | 0.88 | 1.16 | 0.87 |
| sfa | 1823 | 1.00 | 0.87 | 1.16 | 0.96 | 1227 | 1.02 | 0.86 | 1.21 | 0.81 | 1227 | 1.04 | 0.88 | 1.23 | 0.65 | 1823 | 1.00 | 0.86 | 1.16 | 0.98 |
| faw6_fa | 1823 | 1.00 | 0.87 | 1.16 | 0.96 | 1227 | 0.89 | 0.75 | 1.06 | 0.18 | 1227 | 0.89 | 0.75 | 1.05 | 0.16 | 1823 | 0.99 | 0.86 | 1.15 | 0.93 |
| la_fa | 1823 | 1.00 | 0.87 | 1.16 | 0.95 | 1227 | 0.90 | 0.76 | 1.06 | 0.21 | 1227 | 0.89 | 0.75 | 1.06 | 0.19 | 1823 | 1.00 | 0.86 | 1.15 | 0.97 |
| s_hdl_p | 1825 | 1.00 | 0.87 | 1.16 | 0.95 | 1229 | 0.94 | 0.79 | 1.12 | 0.48 | 1229 | 0.94 | 0.79 | 1.11 | 0.46 | 1825 | 0.98 | 0.84 | 1.13 | 0.74 |
| xs_vidl_c | 1825 | 1.01 | 0.87 | 1.16 | 0.92 | 1229 | 0.93 | 0.78 | 1.11 | 0.41 | 1229 | 0.92 | 0.78 | 1.10 | 0.38 | 1825 | 1.01 | 0.87 | 1.17 | 0.93 |
| pc | 1824 | 1.01 | 0.88 | 1.16 | 0.92 | 1228 | 0.89 | 0.76 | 1.06 | 0.19 | 1228 | 0.89 | 0.75 | 1.05 | 0.17 | 1824 | 1.00 | 0.86 | 1.15 | 0.95 |
| tyr | 1820 | 1.01 | 0.88 | 1.16 | 0.91 | 1225 | 1.00 | 0.84 | 1.18 | 0.97 | 1225 | 1.01 | 0.86 | 1.19 | 0.93 | 1820 | 0.99 | 0.86 | 1.15 | 0.94 |
| faw3 | 1823 | 1.01 | 0.88 | 1.16 | 0.90 | 1227 | 1.07 | 0.91 | 1.25 | 0.42 | 1227 | 1.06 | 0.90 | 1.24 | 0.51 | 1823 | 1.00 | 0.87 | 1.16 | 0.95 |
| s_hdl_fc | 1825 | 1.01 | 0.88 | 1.16 | 0.90 | 1229 | 0.98 | 0.82 | 1.16 | 0.77 | 1229 | 0.97 | 0.82 | 1.15 | 0.74 | 1825 | 0.99 | 0.86 | 1.14 | 0.86 |
| m_\|di_tg | 1824 | 1.01 | 0.88 | 1.16 | 0.88 | 1228 | 0.96 | 0.81 | 1.13 | 0.60 | 1228 | 0.95 | 0.80 | 1.13 | 0.58 | 1824 | 1.02 | 0.88 | 1.18 | 0.77 |
| pufa_fa | 1823 | 1.01 | 0.88 | 1.17 | 0.88 | 1227 | 0.91 | 0.77 | 1.08 | 0.28 | 1227 | 0.91 | 0.76 | 1.07 | 0.25 | 1823 | 1.00 | 0.86 | 1.16 | 0.99 |
| xs_vidl_ce | 1825 | 1.01 | 0.88 | 1.17 | 0.85 | 1229 | 0.94 | 0.79 | 1.12 | 0.48 | 1229 | 0.94 | 0.79 | 1.12 | 0.48 | 1825 | 1.01 | 0.87 | 1.18 | 0.86 |
| his | 1820 | 1.02 | 0.88 | 1.17 | 0.83 | 1225 | 1.07 | 0.91 | 1.26 | 0.43 | 1225 | 1.08 | 0.92 | 1.27 | 0.36 | 1820 | 0.99 | 0.86 | 1.15 | 0.90 |
| phe | 1825 | 1.02 | 0.89 | 1.16 | 0.80 | 1229 | 1.03 | 0.88 | 1.22 | 0.68 | 1229 | 1.05 | 0.89 | 1.23 | 0.58 | 1825 | 1.03 | 0.90 | 1.18 | 0.67 |
| totpg | 1824 | 1.02 | 0.89 | 1.17 | 0.80 | 1228 | 0.91 | 0.77 | 1.08 | 0.28 | 1228 | 0.91 | 0.77 | 1.07 | 0.25 | 1824 | 1.00 | 0.87 | 1.16 | 0.95 |
| xs_vldı ${ }^{\text {l }}$ | 1824 | 1.03 | 0.89 | 1.19 | 0.68 | 1228 | 0.95 | 0.80 | 1.13 | 0.59 | 1228 | 0.95 | 0.80 | 1.13 | 0.56 | 1824 | 1.03 | 0.89 | 1.20 | 0.66 |
| ile | 1825 | 1.03 | 0.90 | 1.18 | 0.67 | 1229 | 1.06 | 0.90 | 1.25 | 0.46 | 1229 | 1.08 | 0.93 | 1.27 | 0.32 | 1825 | 1.01 | 0.88 | 1.17 | 0.87 |
| dha_fa | 1823 | 1.03 | 0.90 | 1.18 | 0.66 | 1227 | 1.03 | 0.88 | 1.20 | 0.74 | 1227 | 1.02 | 0.87 | 1.20 | 0.76 | 1823 | 1.03 | 0.90 | 1.18 | 0.68 |
| I_IdI_tg | 1824 | 1.03 | 0.90 | 1.19 | 0.67 | 1228 | 0.99 | 0.84 | 1.17 | 0.90 | 1228 | 0.99 | 0.84 | 1.17 | 0.89 | 1824 | 1.04 | 0.90 | 1.20 | 0.56 |
| IdI_tg | 1825 | 1.03 | 0.90 | 1.19 | 0.65 | 1229 | 0.98 | 0.83 | 1.16 | 0.83 | 1229 | 0.98 | 0.83 | 1.16 | 0.84 | 1825 | 1.04 | 0.90 | 1.21 | 0.55 |
| ace | 1825 | 1.03 | 0.91 | 1.17 | 0.59 | 1229 | 1.01 | 0.87 | 1.18 | 0.90 | 1229 | 1.02 | 0.88 | 1.19 | 0.81 | 1825 | 0.99 | 0.84 | 1.17 | 0.91 |
| apob_apoal | 1824 | 1.04 | 0.90 | 1.20 | 0.60 | 1228 | 0.94 | 0.80 | 1.12 | 0.50 | 1228 | 0.95 | 0.80 | 1.12 | 0.55 | 1824 | 1.03 | 0.89 | 1.19 | 0.67 |
| crea | 1820 | 1.04 | 0.91 | 1.19 | 0.57 | 1225 | 0.99 | 0.84 | 1.17 | 0.94 | 1225 | 0.98 | 0.83 | 1.16 | 0.84 | 1820 | 1.04 | 0.90 | 1.20 | 0.60 |
| xs_vldi_p | 1824 | 1.04 | 0.90 | 1.20 | 0.58 | 1228 | 0.97 | 0.81 | 1.15 | 0.69 | 1228 | 0.96 | 0.81 | 1.14 | 0.68 | 1824 | 1.04 | 0.90 | 1.21 | 0.56 |
| remnant_c | 1825 | 1.04 | 0.90 | 1.20 | 0.56 | 1229 | 0.95 | 0.80 | 1.13 | 0.58 | 1229 | 0.96 | 0.81 | 1.13 | 0.61 | 1825 | 1.04 | 0.90 | 1.20 | 0.60 |
| faw3_fa | 1823 | 1.04 | 0.91 | 1.19 | 0.54 | 1227 | 1.03 | 0.87 | 1.20 | 0.75 | 1227 | 1.02 | 0.87 | 1.20 | 0.77 | 1823 | 1.04 | 0.90 | 1.19 | 0.61 |


|  | Minimally adjusted model |  |  |  |  | Fully adjusted model |  |  |  |  | Minimally adjusted model (Complete case) |  |  |  |  | Minimally adjusted model censored |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_ldl_tg | 1824 | 1.04 | 0.91 | 1.20 | 0.53 | 1228 | 0.98 | 0.83 | 1.15 | 0.77 | 1228 | 0.98 | 0.83 | 1.15 | 0.80 | 1824 | 1.05 | 0.91 | 1.21 | 0.48 |
| xl_hdl_tg | 1822 | 1.05 | 0.92 | 1.20 | 0.49 | 1228 | 0.99 | 0.84 | 1.17 | 0.89 | 1228 | 1.00 | 0.84 | 1.18 | 0.97 | 1822 | 1.06 | 0.92 | 1.22 | 0.44 |
| acace | 1825 | 1.06 | 0.95 | 1.17 | 0.31 | 1229 | 1.07 | 0.96 | 1.19 | 0.23 | 1229 | 1.07 | 0.97 | 1.19 | 0.16 | 1825 | 1.04 | 0.93 | 1.16 | 0.48 |
| s_vidi_ce | 1824 | 1.06 | 0.92 | 1.22 | 0.41 | 1228 | 0.99 | 0.83 | 1.17 | 0.91 | 1228 | 0.99 | 0.83 | 1.17 | 0.88 | 1824 | 1.07 | 0.92 | 1.23 | 0.40 |
| totfa | 1823 | 1.06 | 0.93 | 1.22 | 0.39 | 1227 | 0.97 | 0.82 | 1.15 | 0.74 | 1227 | 0.98 | 0.83 | 1.16 | 0.80 | 1823 | 1.05 | 0.92 | 1.22 | 0.46 |
| idl_tg | 1824 | 1.07 | 0.93 | 1.22 | 0.37 | 1228 | 1.03 | 0.88 | 1.22 | 0.69 | 1228 | 1.03 | 0.88 | 1.22 | 0.69 | 1824 | 1.08 | 0.94 | 1.25 | 0.27 |
| s_hdl_pl | 1825 | 1.07 | 0.93 | 1.22 | 0.36 | 1229 | 1.03 | 0.87 | 1.23 | 0.72 | 1229 | 1.04 | 0.88 | 1.23 | 0.67 | 1825 | 1.05 | 0.91 | 1.21 | 0.51 |
| sfa_fa | 1823 | 1.07 | 0.93 | 1.22 | 0.35 | 1227 | 0.98 | 0.83 | 1.16 | 0.84 | 1227 | 0.99 | 0.84 | 1.17 | 0.93 | 1823 | 1.06 | 0.92 | 1.22 | 0.42 |
| apob | 1824 | 1.07 | 0.93 | 1.24 | 0.32 | 1228 | 1.03 | 0.86 | 1.22 | 0.77 | 1228 | 1.04 | 0.88 | 1.23 | 0.66 | 1824 | 1.08 | 0.93 | 1.25 | 0.31 |
| s_vidl_c | 1824 | 1.08 | 0.94 | 1.24 | 0.28 | 1228 | 1.02 | 0.86 | 1.20 | 0.84 | 1228 | 1.02 | 0.86 | 1.20 | 0.84 | 1824 | 1.09 | 0.94 | 1.26 | 0.26 |
| xxl_vidl_tg | 1820 | 1.09 | 0.95 | 1.24 | 0.23 | 1228 | 1.04 | 0.88 | 1.22 | 0.65 | 1228 | 1.06 | 0.90 | 1.24 | 0.50 | 1820 | 1.09 | 0.95 | 1.25 | 0.24 |
| mufa_fa | 1823 | 1.09 | 0.96 | 1.25 | 0.20 | 1227 | 1.02 | 0.87 | 1.20 | 0.79 | 1227 | 1.03 | 0.88 | 1.22 | 0.68 | 1823 | 1.09 | 0.95 | 1.25 | 0.23 |
| xxl_vidl_l | 1819 | 1.09 | 0.95 | 1.25 | 0.20 | 1227 | 1.05 | 0.89 | 1.23 | 0.59 | 1227 | 1.06 | 0.91 | 1.25 | 0.45 | 1819 | 1.09 | 0.95 | 1.25 | 0.21 |
| vidl_d | 1825 | 1.09 | 0.95 | 1.26 | 0.20 | 1229 | 1.08 | 0.91 | 1.28 | 0.37 | 1229 | 1.09 | 0.93 | 1.29 | 0.29 | 1825 | 1.10 | 0.95 | 1.27 | 0.21 |
| m_vidl_tg | 1823 | 1.09 | 0.96 | 1.25 | 0.19 | 1228 | 1.05 | 0.89 | 1.23 | 0.59 | 1228 | 1.06 | 0.90 | 1.24 | 0.47 | 1823 | 1.10 | 0.96 | 1.26 | 0.19 |
| hdi_tg | 1825 | 1.10 | 0.96 | 1.25 | 0.18 | 1229 | 1.06 | 0.91 | 1.25 | 0.44 | 1229 | 1.07 | 0.92 | 1.26 | 0.38 | 1825 | 1.11 | 0.97 | 1.27 | 0.15 |
| xxl_vidl_ce | 1822 | 1.10 | 0.96 | 1.25 | 0.18 | 1227 | 1.04 | 0.88 | 1.23 | 0.65 | 1227 | 1.05 | 0.90 | 1.24 | 0.54 | 1822 | 1.10 | 0.96 | 1.27 | 0.16 |
| m_vidl_। | 1822 | 1.10 | 0.96 | 1.25 | 0.18 | 1227 | 1.05 | 0.89 | 1.23 | 0.60 | 1227 | 1.06 | 0.90 | 1.24 | 0.49 | 1822 | 1.10 | 0.96 | 1.26 | 0.18 |
| I_vidl_ce | 1822 | 1.10 | 0.96 | 1.25 | 0.18 | 1227 | 1.05 | 0.89 | 1.24 | 0.54 | 1227 | 1.06 | 0.91 | 1.25 | 0.45 | 1822 | 1.10 | 0.96 | 1.27 | 0.16 |
| xl_vidl_ce | 1822 | 1.10 | 0.96 | 1.25 | 0.18 | 1227 | 1.05 | 0.89 | 1.23 | 0.59 | 1227 | 1.06 | 0.90 | 1.24 | 0.48 | 1822 | 1.10 | 0.96 | 1.27 | 0.15 |
| m_vidl_fc | 1822 | 1.10 | 0.96 | 1.25 | 0.18 | 1227 | 1.05 | 0.89 | 1.23 | 0.58 | 1227 | 1.06 | 0.90 | 1.24 | 0.48 | 1822 | 1.10 | 0.96 | 1.26 | 0.17 |
| I_vidl_tg | 1823 | 1.10 | 0.96 | 1.25 | 0.17 | 1228 | 1.04 | 0.88 | 1.23 | 0.62 | 1228 | 1.06 | 0.90 | 1.24 | 0.48 | 1823 | 1.10 | 0.96 | 1.26 | 0.18 |
| m_vldi_p | 1822 | 1.10 | 0.96 | 1.25 | 0.18 | 1227 | 1.05 | 0.89 | 1.23 | 0.58 | 1227 | 1.06 | 0.90 | 1.24 | 0.48 | 1822 | 1.10 | 0.96 | 1.26 | 0.18 |
| xxl_vidl_pl | 1819 | 1.10 | 0.96 | 1.25 | 0.17 | 1227 | 1.05 | 0.89 | 1.24 | 0.54 | 1227 | 1.07 | 0.91 | 1.25 | 0.41 | 1819 | 1.10 | 0.96 | 1.26 | 0.18 |
| I_vidl_I | 1823 | 1.10 | 0.96 | 1.25 | 0.17 | 1228 | 1.04 | 0.88 | 1.23 | 0.62 | 1228 | 1.06 | 0.90 | 1.24 | 0.48 | 1823 | 1.10 | 0.96 | 1.26 | 0.17 |
| I_vidi_p | 1823 | 1.10 | 0.96 | 1.25 | 0.17 | 1228 | 1.04 | 0.88 | 1.23 | 0.62 | 1228 | 1.06 | 0.90 | 1.24 | 0.48 | 1823 | 1.10 | 0.96 | 1.26 | 0.17 |
| vidl_c | 1825 | 1.10 | 0.96 | 1.26 | 0.18 | 1229 | 1.03 | 0.88 | 1.22 | 0.70 | 1229 | 1.04 | 0.89 | 1.23 | 0.62 | 1825 | 1.10 | 0.96 | 1.27 | 0.17 |


|  | Minimally adjusted model |  |  |  |  | Fully adjusted model |  |  |  |  | Minimally adjusted model (Complete case) |  |  |  |  | Minimally adjusted model censored |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| I_vidl_pl | 1823 | 1.10 | 0.96 | 1.25 | 0.17 | 1228 | 1.04 | 0.88 | 1.23 | 0.63 | 1228 | 1.06 | 0.90 | 1.24 | 0.49 | 1823 | 1.10 | 0.96 | 1.26 | 0.17 |
| I_vidl_c | 1822 | 1.10 | 0.96 | 1.25 | 0.16 | 1227 | 1.05 | 0.89 | 1.24 | 0.57 | 1227 | 1.06 | 0.91 | 1.25 | 0.45 | 1822 | 1.10 | 0.96 | 1.27 | 0.15 |
| I_vidl_fc | 1823 | 1.10 | 0.96 | 1.25 | 0.16 | 1228 | 1.04 | 0.88 | 1.23 | 0.62 | 1228 | 1.06 | 0.90 | 1.24 | 0.48 | 1823 | 1.10 | 0.96 | 1.26 | 0.16 |
| xl_vidl_c | 1823 | 1.10 | 0.96 | 1.25 | 0.16 | 1228 | 1.04 | 0.88 | 1.23 | 0.62 | 1228 | 1.06 | 0.90 | 1.24 | 0.49 | 1823 | 1.11 | 0.96 | 1.27 | 0.15 |
| s_vidl_fc | 1824 | 1.10 | 0.96 | 1.26 | 0.16 | 1228 | 1.06 | 0.90 | 1.25 | 0.48 | 1228 | 1.06 | 0.90 | 1.25 | 0.46 | 1824 | 1.11 | 0.97 | 1.28 | 0.14 |
| m_vldi_ce | 1824 | 1.10 | 0.96 | 1.26 | 0.16 | 1228 | 1.05 | 0.90 | 1.24 | 0.53 | 1228 | 1.06 | 0.90 | 1.25 | 0.47 | 1824 | 1.11 | 0.97 | 1.28 | 0.14 |
| xl_vidl_p | 1823 | 1.10 | 0.97 | 1.26 | 0.15 | 1228 | 1.04 | 0.89 | 1.23 | 0.61 | 1228 | 1.06 | 0.90 | 1.25 | 0.46 | 1823 | 1.11 | 0.97 | 1.27 | 0.15 |
| xs_vldi_tg | 1824 | 1.10 | 0.97 | 1.26 | 0.15 | 1228 | 1.08 | 0.92 | 1.27 | 0.34 | 1228 | 1.08 | 0.92 | 1.27 | 0.32 | 1824 | 1.12 | 0.97 | 1.28 | 0.11 |
| xl_vidl_tg | 1822 | 1.11 | 0.97 | 1.26 | 0.14 | 1228 | 1.04 | 0.88 | 1.23 | 0.61 | 1228 | 1.06 | 0.90 | 1.24 | 0.47 | 1822 | 1.11 | 0.97 | 1.27 | 0.14 |
| xl_vidl_I | 1822 | 1.11 | 0.97 | 1.26 | 0.13 | 1228 | 1.04 | 0.89 | 1.23 | 0.61 | 1228 | 1.06 | 0.91 | 1.24 | 0.47 | 1822 | 1.11 | 0.97 | 1.27 | 0.13 |
| s_vidl_I | 1824 | 1.11 | 0.97 | 1.27 | 0.14 | 1228 | 1.07 | 0.91 | 1.25 | 0.44 | 1228 | 1.07 | 0.91 | 1.26 | 0.40 | 1824 | 1.12 | 0.97 | 1.28 | 0.12 |
| s_vidl_p | 1824 | 1.11 | 0.97 | 1.27 | 0.13 | 1228 | 1.07 | 0.91 | 1.26 | 0.40 | 1228 | 1.08 | 0.92 | 1.26 | 0.36 | 1824 | 1.12 | 0.97 | 1.28 | 0.12 |
| xl_vidl_fc | 1822 | 1.11 | 0.97 | 1.26 | 0.12 | 1228 | 1.04 | 0.89 | 1.23 | 0.61 | 1228 | 1.06 | 0.90 | 1.24 | 0.48 | 1822 | 1.11 | 0.97 | 1.27 | 0.12 |
| serum_tg | 1825 | 1.11 | 0.97 | 1.26 | 0.12 | 1229 | 1.07 | 0.92 | 1.26 | 0.38 | 1229 | 1.09 | 0.93 | 1.27 | 0.30 | 1825 | 1.11 | 0.98 | 1.27 | 0.11 |
| xxl_vidl_c | 1821 | 1.11 | 0.97 | 1.27 | 0.12 | 1228 | 1.04 | 0.89 | 1.23 | 0.62 | 1228 | 1.06 | 0.90 | 1.24 | 0.49 | 1821 | 1.12 | 0.97 | 1.28 | 0.12 |
| s_vidl_pl | 1824 | 1.11 | 0.97 | 1.27 | 0.13 | 1228 | 1.07 | 0.91 | 1.26 | 0.41 | 1228 | 1.07 | 0.92 | 1.26 | 0.38 | 1824 | 1.12 | 0.98 | 1.29 | 0.11 |
| mufa | 1823 | 1.11 | 0.97 | 1.27 | 0.12 | 1227 | 1.10 | 0.94 | 1.29 | 0.23 | 1227 | 1.12 | 0.96 | 1.31 | 0.16 | 1823 | 1.12 | 0.97 | 1.28 | 0.11 |
| vidl_tg | 1825 | 1.11 | 0.98 | 1.27 | 0.11 | 1229 | 1.08 | 0.92 | 1.27 | 0.34 | 1229 | 1.09 | 0.94 | 1.28 | 0.26 | 1825 | 1.12 | 0.98 | 1.27 | 0.10 |
| xl_vidl_pl | 1821 | 1.11 | 0.98 | 1.27 | 0.11 | 1227 | 1.05 | 0.89 | 1.24 | 0.58 | 1227 | 1.06 | 0.91 | 1.25 | 0.45 | 1821 | 1.12 | 0.97 | 1.28 | 0.11 |
| m_vldi_c | 1823 | 1.11 | 0.97 | 1.27 | 0.11 | 1228 | 1.07 | 0.91 | 1.25 | 0.43 | 1228 | 1.08 | 0.92 | 1.26 | 0.36 | 1823 | 1.12 | 0.98 | 1.29 | 0.10 |
| xxl_vldi_p | 1825 | 1.11 | 0.98 | 1.26 | 0.09 | 1229 | 1.08 | 0.92 | 1.27 | 0.36 | 1229 | 1.10 | 0.93 | 1.29 | 0.26 | 1825 | 1.12 | 0.98 | 1.27 | 0.09 |
| s_vidl_tg | 1823 | 1.12 | 0.98 | 1.27 | 0.11 | 1228 | 1.09 | 0.93 | 1.28 | 0.29 | 1228 | 1.10 | 0.94 | 1.28 | 0.24 | 1823 | 1.12 | 0.98 | 1.29 | 0.09 |
| IdI_d | 1825 | 1.12 | 0.98 | 1.28 | 0.11 | 1229 | 1.18 | 1.00 | 1.38 | 0.05 | 1229 | 1.19 | 1.01 | 1.40 | 0.03 | 1825 | 1.15 | 1.00 | 1.32 | 0.05 |
| tg_pg | 1824 | 1.12 | 0.98 | 1.28 | 0.10 | 1228 | 1.11 | 0.94 | 1.31 | 0.20 | 1228 | 1.13 | 0.96 | 1.32 | 0.14 | 1824 | 1.12 | 0.98 | 1.29 | 0.10 |
| m_vldi_pl | 1823 | 1.12 | 0.98 | 1.28 | 0.10 | 1228 | 1.07 | 0.92 | 1.26 | 0.38 | 1228 | 1.09 | 0.93 | 1.27 | 0.30 | 1823 | 1.12 | 0.98 | 1.29 | 0.09 |
| xxl_vldi_fc | 1820 | 1.12 | 0.98 | 1.28 | 0.09 | 1227 | 1.06 | 0.90 | 1.25 | 0.50 | 1227 | 1.07 | 0.92 | 1.26 | 0.38 | 1820 | 1.12 | 0.98 | 1.29 | 0.09 |
| s_hdl_tg | 1824 | 1.12 | 0.98 | 1.28 | 0.10 | 1228 | 1.12 | 0.95 | 1.32 | 0.17 | 1228 | 1.13 | 0.96 | 1.32 | 0.13 | 1824 | 1.13 | 0.99 | 1.30 | 0.08 |
| m_hdl_tg | 1824 | 1.14 | 1.00 | 1.30 | 0.05 | 1228 | 1.10 | 0.93 | 1.29 | 0.26 | 1228 | 1.11 | 0.94 | 1.29 | 0.21 | 1824 | 1.15 | 1.00 | 1.31 | 0.05 |
| gp | 1825 | 1.18 | 1.03 | 1.35 | 0.01 | 1229 | 1.15 | 0.97 | 1.36 | 0.10 | 1229 | 1.15 | 0.98 | 1.36 | 0.08 | 1825 | 1.18 | 1.03 | 1.35 | 0.02 |

Appendix B Table B 2: Cox regression results for metastases or PCa death- minimally adjusted model, fully adjusted model, censored minimally adjusted model

|  | Minimally adjusted model |  |  |  |  | Fully adjusted model |  |  |  |  | Minimally adjusted(Complete case) |  |  |  |  | Minimally adjusted model censored |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| metabolite | N | Hazard Ratio | $\begin{aligned} & \text { Lower } \\ & 95 \% \mathrm{Cl} \\ & \hline \end{aligned}$ | $\begin{aligned} & \hline \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p -value | N | $\begin{aligned} & \hline \begin{array}{l} \text { Hazard } \\ \text { Ratio } \end{array} \\ & \hline \end{aligned}$ | $\begin{aligned} & \hline \text { Lower } \\ & 95 \% \mathrm{Cl} \\ & \hline \end{aligned}$ | $\begin{aligned} & \hline \text { Upper } \\ & 95 \% \mathrm{Cl} \\ & \hline \end{aligned}$ | p-value | N | Hazard Ratio | $\begin{aligned} & \hline \text { Lower } \\ & 95 \% \mathrm{Cl} \\ & \hline \end{aligned}$ | $\begin{aligned} & \hline \text { Upper } \\ & 95 \% \mathrm{Cl} \\ & \hline \end{aligned}$ | $p$-value | N | Hazard Ratio | $\begin{aligned} & \hline \text { Lower } \\ & 95 \% \mathrm{Cl} \\ & \hline \end{aligned}$ | $\begin{aligned} & \hline \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | pvalue |
| s_hdl_ce | 1822 | 0.78 | 0.61 | 0.99 | 0.04 | 1228 | 0.74 | 0.56 | 0.99 | 0.04 | 1228 | 0.76 | 0.58 | 1.01 | 0.06 | 1822 | 0.79 | 0.61 | 1.03 | 0.09 |
| s_hdl_c | 1822 | 0.78 | 0.61 | 0.99 | 0.04 | 1228 | 0.75 | 0.56 | 1.01 | 0.05 | 1228 | 0.77 | 0.58 | 1.03 | 0.08 | 1822 | 0.81 | 0.62 | 1.06 | 0.12 |
| s_ldi_ce | 1825 | 0.79 | 0.61 | 1.02 | 0.07 | 1229 | 0.70 | 0.51 | 0.96 | 0.03 | 1229 | 0.69 | 0.51 | 0.95 | 0.02 | 1825 | 0.77 | 0.59 | 1.02 | 0.07 |
| s_ldi_c | 1825 | 0.79 | 0.61 | 1.02 | 0.07 | 1229 | 0.70 | 0.51 | 0.97 | 0.03 | 1229 | 0.70 | 0.51 | 0.95 | 0.02 | 1825 | 0.78 | 0.59 | 1.03 | 0.08 |
| idl_fc | 1825 | 0.79 | 0.61 | 1.03 | 0.08 | 1229 | 0.72 | 0.52 | 0.98 | 0.04 | 1229 | 0.73 | 0.54 | 1.00 | 0.05 | 1825 | 0.80 | 0.61 | 1.06 | 0.12 |
| m_ldi_ce | 1825 | 0.79 | 0.61 | 1.03 | 0.08 | 1229 | 0.71 | 0.52 | 0.98 | 0.04 | 1229 | 0.70 | 0.51 | 0.96 | 0.03 | 1825 | 0.78 | 0.59 | 1.03 | 0.08 |
| s_ldi_fc | 1825 | 0.80 | 0.61 | 1.03 | 0.09 | 1229 | 0.71 | 0.52 | 0.97 | 0.03 | 1229 | 0.71 | 0.52 | 0.97 | 0.03 | 1825 | 0.79 | 0.59 | 1.04 | 0.09 |
| I_IdI_fc | 1825 | 0.80 | 0.61 | 1.03 | 0.09 | 1229 | 0.71 | 0.52 | 0.98 | 0.04 | 1229 | 0.72 | 0.53 | 0.99 | 0.04 | 1825 | 0.79 | 0.60 | 1.05 | 0.11 |
| m_Idl_c | 1825 | 0.80 | 0.61 | 1.04 | 0.09 | 1229 | 0.71 | 0.52 | 0.98 | 0.04 | 1229 | 0.70 | 0.51 | 0.97 | 0.03 | 1825 | 0.78 | 0.59 | 1.04 | 0.09 |
| s_ldi_\| | 1825 | 0.80 | 0.62 | 1.05 | 0.11 | 1229 | 0.72 | 0.52 | 0.99 | 0.04 | 1229 | 0.71 | 0.52 | 0.97 | 0.03 | 1825 | 0.79 | 0.60 | 1.05 | 0.11 |
| s_ldi_p | 1825 | 0.81 | 0.62 | 1.05 | 0.11 | 1229 | 0.72 | 0.53 | 1.00 | 0.05 | 1229 | 0.71 | 0.52 | 0.98 | 0.03 | 1825 | 0.80 | 0.60 | 1.06 | 0.11 |
| m_ldi_fc | 1825 | 0.81 | 0.62 | 1.05 | 0.11 | 1229 | 0.73 | 0.53 | 1.00 | 0.05 | 1229 | 0.72 | 0.52 | 0.98 | 0.04 | 1825 | 0.79 | 0.60 | 1.05 | 0.10 |
| IdI_c | 1825 | 0.81 | 0.62 | 1.05 | 0.11 | 1229 | 0.72 | 0.52 | 1.00 | 0.05 | 1229 | 0.72 | 0.52 | 0.98 | 0.04 | 1825 | 0.80 | 0.60 | 1.06 | 0.11 |
| m_ldi_\| | 1825 | 0.81 | 0.62 | 1.05 | 0.12 | 1229 | 0.73 | 0.53 | 1.01 | 0.05 | 1229 | 0.72 | 0.52 | 0.99 | 0.04 | 1825 | 0.80 | 0.60 | 1.06 | 0.12 |
| m_ldi_p | 1825 | 0.81 | 0.62 | 1.06 | 0.12 | 1229 | 0.74 | 0.53 | 1.01 | 0.06 | 1229 | 0.72 | 0.52 | 0.99 | 0.04 | 1825 | 0.80 | 0.61 | 1.07 | 0.13 |
| hdl3_c | 1825 | 0.82 | 0.63 | 1.05 | 0.11 | 1229 | 0.74 | 0.55 | 1.01 | 0.05 | 1229 | 0.79 | 0.59 | 1.06 | 0.12 | 1825 | 0.85 | 0.65 | 1.11 | 0.24 |
| idl_pl | 1825 | 0.82 | 0.63 | 1.06 | 0.13 | 1229 | 0.74 | 0.54 | 1.03 | 0.07 | 1229 | 0.75 | 0.54 | 1.03 | 0.08 | 1825 | 0.83 | 0.62 | 1.10 | 0.18 |
| unsat | 1823 | 0.82 | 0.64 | 1.06 | 0.13 | 1227 | 0.77 | 0.57 | 1.05 | 0.10 | 1227 | 0.82 | 0.61 | 1.10 | 0.18 | 1823 | 0.84 | 0.64 | 1.10 | 0.21 |
| I_Id_c | 1825 | 0.83 | 0.64 | 1.07 | 0.15 | 1229 | 0.74 | 0.54 | 1.03 | 0.07 | 1229 | 0.74 | 0.54 | 1.02 | 0.06 | 1825 | 0.82 | 0.62 | 1.08 | 0.16 |
| I_\|di_| | 1825 | 0.83 | 0.64 | 1.08 | 0.17 | 1229 | 0.76 | 0.55 | 1.05 | 0.09 | 1229 | 0.75 | 0.55 | 1.04 | 0.08 | 1825 | 0.83 | 0.63 | 1.10 | 0.20 |
| I_\|di_pl | 1825 | 0.83 | 0.64 | 1.09 | 0.18 | 1229 | 0.76 | 0.55 | 1.04 | 0.09 | 1229 | 0.75 | 0.55 | 1.04 | 0.08 | 1825 | 0.83 | 0.63 | 1.11 | 0.21 |
| I_ldi_ce | 1825 | 0.84 | 0.64 | 1.09 | 0.18 | 1229 | 0.76 | 0.55 | 1.04 | 0.09 | 1229 | 0.75 | 0.54 | 1.03 | 0.07 | 1825 | 0.83 | 0.62 | 1.09 | 0.18 |
| I_IdI_p | 1825 | 0.84 | 0.65 | 1.09 | 0.20 | 1229 | 0.77 | 0.56 | 1.06 | 0.11 | 1229 | 0.76 | 0.55 | 1.05 | 0.09 | 1825 | 0.84 | 0.63 | 1.11 | 0.23 |
| faw6_fa | 1823 | 0.84 | 0.67 | 1.06 | 0.15 | 1227 | 0.77 | 0.59 | 1.01 | 0.06 | 1227 | 0.81 | 0.62 | 1.06 | 0.13 | 1823 | 0.85 | 0.66 | 1.09 | 0.21 |


|  | Minimally adjusted model |  |  |  |  | Fully adjusted model |  |  |  |  | Minimally adjusted(Complete case) |  |  |  |  | Minimally adjusted model censored |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| pufa_fa | 1823 | 0.85 | 0.67 | 1.07 | 0.17 | 1227 | 0.79 | 0.60 | 1.05 | 0.10 | 1227 | 0.83 | 0.63 | 1.10 | 0.19 | 1823 | 0.85 | 0.66 | 1.10 | 0.21 |
| s_ldi_pl | 1825 | 0.85 | 0.66 | 1.10 | 0.21 | 1229 | 0.76 | 0.56 | 1.04 | 0.08 | 1229 | 0.75 | 0.55 | 1.02 | 0.06 | 1825 | 0.83 | 0.63 | 1.10 | 0.20 |
| estc | 1824 | 0.85 | 0.66 | 1.10 | 0.23 | 1228 | 0.76 | 0.55 | 1.04 | 0.09 | 1228 | 0.77 | 0.56 | 1.05 | 0.10 | 1824 | 0.86 | 0.65 | 1.14 | 0.31 |
| m_ldi_pl | 1825 | 0.85 | 0.66 | 1.11 | 0.24 | 1229 | 0.78 | 0.57 | 1.06 | 0.12 | 1229 | 0.76 | 0.55 | 1.04 | 0.08 | 1825 | 0.84 | 0.63 | 1.11 | 0.22 |
| serum_c | 1825 | 0.85 | 0.66 | 1.11 | 0.24 | 1229 | 0.76 | 0.55 | 1.05 | 0.10 | 1229 | 0.77 | 0.57 | 1.06 | 0.11 | 1825 | 0.87 | 0.66 | 1.15 | 0.32 |
| idl_1 | 1825 | 0.86 | 0.66 | 1.11 | 0.24 | 1229 | 0.79 | 0.57 | 1.09 | 0.15 | 1229 | 0.79 | 0.57 | 1.09 | 0.15 | 1825 | 0.86 | 0.65 | 1.14 | 0.31 |
| idl_c | 1825 | 0.86 | 0.66 | 1.11 | 0.24 | 1229 | 0.78 | 0.56 | 1.08 | 0.13 | 1229 | 0.78 | 0.57 | 1.08 | 0.13 | 1825 | 0.86 | 0.65 | 1.13 | 0.28 |
| hdl_c | 1825 | 0.86 | 0.66 | 1.11 | 0.25 | 1229 | 0.77 | 0.57 | 1.05 | 0.10 | 1229 | 0.86 | 0.63 | 1.16 | 0.33 | 1825 | 0.92 | 0.70 | 1.21 | 0.55 |
| freec | 1824 | 0.86 | 0.66 | 1.12 | 0.26 | 1228 | 0.77 | 0.56 | 1.06 | 0.11 | 1228 | 0.78 | 0.57 | 1.07 | 0.12 | 1824 | 0.88 | 0.66 | 1.16 | 0.35 |
| xs_vidl_pl | 1825 | 0.86 | 0.66 | 1.12 | 0.27 | 1229 | 0.81 | 0.59 | 1.12 | 0.20 | 1229 | 0.80 | 0.59 | 1.11 | 0.18 | 1825 | 0.90 | 0.68 | 1.19 | 0.45 |
| idl_p | 1825 | 0.86 | 0.67 | 1.12 | 0.28 | 1229 | 0.81 | 0.59 | 1.11 | 0.18 | 1229 | 0.80 | 0.58 | 1.10 | 0.18 | 1825 | 0.88 | 0.66 | 1.16 | 0.36 |
| hdl2_c | 1825 | 0.87 | 0.67 | 1.12 | 0.28 | 1229 | 0.78 | 0.57 | 1.06 | 0.12 | 1229 | 0.87 | 0.64 | 1.18 | 0.37 | 1825 | 0.93 | 0.71 | 1.22 | 0.59 |
| gly | 1817 | 0.87 | 0.67 | 1.14 | 0.32 | 1221 | 0.83 | 0.59 | 1.15 | 0.27 | 1221 | 0.86 | 0.62 | 1.19 | 0.36 | 1817 | 0.92 | 0.70 | 1.22 | 0.57 |
| m_hdl_fc | 1825 | 0.87 | 0.68 | 1.12 | 0.30 | 1229 | 0.83 | 0.61 | 1.14 | 0.25 | 1229 | 0.91 | 0.68 | 1.24 | 0.56 | 1825 | 0.98 | 0.75 | 1.28 | 0.87 |
| cit | 1824 | 0.88 | 0.68 | 1.14 | 0.34 | 1228 | 0.94 | 0.70 | 1.26 | 0.67 | 1228 | 0.92 | 0.69 | 1.23 | 0.57 | 1824 | 1.00 | 0.76 | 1.30 | 0.98 |
| 1_hdl_pl | 1825 | 0.89 | 0.69 | 1.15 | 0.36 | 1229 | 0.81 | 0.60 | 1.11 | 0.19 | 1229 | 0.91 | 0.68 | 1.23 | 0.54 | 1825 | 0.95 | 0.73 | 1.25 | 0.72 |
| la_fa | 1823 | 0.89 | 0.70 | 1.13 | 0.32 | 1227 | 0.82 | 0.63 | 1.07 | 0.14 | 1227 | 0.83 | 0.64 | 1.09 | 0.19 | 1823 | 0.89 | 0.69 | 1.15 | 0.39 |
| idl_ce | 1825 | 0.89 | 0.68 | 1.15 | 0.37 | 1229 | 0.81 | 0.59 | 1.12 | 0.21 | 1229 | 0.81 | 0.59 | 1.12 | 0.20 | 1825 | 0.88 | 0.67 | 1.17 | 0.38 |
| 1_hdl_fc | 1824 | 0.89 | 0.69 | 1.16 | 0.39 | 1228 | 0.82 | 0.60 | 1.11 | 0.20 | 1228 | 0.91 | 0.68 | 1.23 | 0.55 | 1824 | 0.94 | 0.72 | 1.24 | 0.67 |
| xs_vidl_fc | 1824 | 0.89 | 0.69 | 1.16 | 0.40 | 1228 | 0.86 | 0.63 | 1.18 | 0.35 | 1228 | 0.85 | 0.62 | 1.17 | 0.32 | 1824 | 0.94 | 0.71 | 1.25 | 0.67 |
| pyr | 1823 | 0.90 | 0.66 | 1.23 | 0.50 | 1227 | 0.75 | 0.49 | 1.16 | 0.20 | 1227 | 0.71 | 0.46 | 1.10 | 0.12 | 1823 | 0.95 | 0.70 | 1.28 | 0.74 |
| s_hdl_l | 1825 | 0.90 | 0.70 | 1.16 | 0.42 | 1229 | 0.90 | 0.66 | 1.22 | 0.50 | 1229 | 0.91 | 0.67 | 1.23 | 0.52 | 1825 | 0.98 | 0.74 | 1.28 | 0.86 |
| lac | 1825 | 0.90 | 0.68 | 1.20 | 0.48 | 1229 | 1.04 | 0.70 | 1.55 | 0.83 | 1229 | 1.01 | 0.67 | 1.51 | 0.96 | 1825 | 0.86 | 0.61 | 1.20 | 0.37 |
| m_hdl_c | 1825 | 0.91 | 0.70 | 1.17 | 0.44 | 1229 | 0.86 | 0.63 | 1.17 | 0.33 | 1229 | 0.93 | 0.69 | 1.26 | 0.66 | 1825 | 0.99 | 0.76 | 1.29 | 0.94 |
| I_hdl_I | 1824 | 0.91 | 0.70 | 1.17 | 0.45 | 1228 | 0.84 | 0.62 | 1.14 | 0.26 | 1228 | 0.93 | 0.70 | 1.26 | 0.66 | 1824 | 0.97 | 0.74 | 1.27 | 0.81 |
| hdl_d | 1825 | 0.91 | 0.70 | 1.17 | 0.46 | 1229 | 0.86 | 0.64 | 1.16 | 0.33 | 1229 | 0.95 | 0.71 | 1.27 | 0.74 | 1825 | 0.96 | 0.74 | 1.26 | 0.79 |
| I_hdl_p | 1824 | 0.91 | 0.70 | 1.17 | 0.47 | 1228 | 0.84 | 0.62 | 1.14 | 0.27 | 1228 | 0.94 | 0.70 | 1.26 | 0.67 | 1824 | 0.97 | 0.74 | 1.27 | 0.84 |


|  | Minimally adjusted model |  |  |  |  | Fully adjusted model |  |  |  |  | Minimally adjusted(Complete case) |  |  |  |  | Minimally adjusted model censored |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| apoa1 | 1825 | 0.91 | 0.71 | 1.17 | 0.47 | 1229 | 0.81 | 0.60 | 1.09 | 0.16 | 1229 | 0.88 | 0.65 | 1.18 | 0.38 | 1825 | 0.96 | 0.74 | 1.26 | 0.78 |
| ala | 1825 | 0.92 | 0.71 | 1.18 | 0.49 | 1229 | 0.82 | 0.61 | 1.11 | 0.20 | 1229 | 0.77 | 0.57 | 1.04 | 0.09 | 1825 | 0.99 | 0.76 | 1.30 | 0.96 |
| m_hdl_ce | 1825 | 0.92 | 0.71 | 1.18 | 0.50 | 1229 | 0.86 | 0.63 | 1.18 | 0.36 | 1229 | 0.94 | 0.69 | 1.27 | 0.69 | 1825 | 0.99 | 0.76 | 1.30 | 0.96 |
| I_hdl_c | 1824 | 0.92 | 0.71 | 1.18 | 0.51 | 1228 | 0.85 | 0.63 | 1.15 | 0.29 | 1228 | 0.94 | 0.70 | 1.27 | 0.71 | 1824 | 0.97 | 0.74 | 1.27 | 0.84 |
| m_\|di_tg | 1824 | 0.92 | 0.71 | 1.19 | 0.52 | 1228 | 0.92 | 0.68 | 1.25 | 0.60 | 1228 | 0.91 | 0.67 | 1.23 | 0.53 | 1824 | 0.98 | 0.75 | 1.29 | 0.90 |
| dha_fa | 1823 | 0.92 | 0.71 | 1.19 | 0.53 | 1227 | 0.98 | 0.73 | 1.32 | 0.90 | 1227 | 1.01 | 0.75 | 1.34 | 0.97 | 1823 | 0.94 | 0.71 | 1.24 | 0.64 |
| s_hdl_p | 1825 | 0.92 | 0.71 | 1.19 | 0.53 | 1229 | 0.93 | 0.69 | 1.26 | 0.64 | 1229 | 0.93 | 0.69 | 1.26 | 0.66 | 1825 | 1.00 | 0.76 | 1.31 | 0.99 |
| I_hdl_ce | 1824 | 0.92 | 0.72 | 1.19 | 0.55 | 1228 | 0.86 | 0.63 | 1.17 | 0.33 | 1228 | 0.96 | 0.71 | 1.28 | 0.76 | 1824 | 0.98 | 0.75 | 1.28 | 0.89 |
| xl_hdl_pl | 1823 | 0.92 | 0.72 | 1.19 | 0.55 | 1228 | 0.85 | 0.63 | 1.14 | 0.28 | 1228 | 0.93 | 0.70 | 1.25 | 0.64 | 1823 | 0.97 | 0.74 | 1.27 | 0.82 |
| m_hdl_I | 1825 | 0.93 | 0.73 | 1.20 | 0.58 | 1229 | 0.90 | 0.66 | 1.22 | 0.50 | 1229 | 0.97 | 0.72 | 1.31 | 0.86 | 1825 | 1.03 | 0.79 | 1.34 | 0.81 |
| m_hdl_pl | 1824 | 0.94 | 0.73 | 1.20 | 0.61 | 1229 | 0.92 | 0.68 | 1.24 | 0.58 | 1229 | 0.99 | 0.74 | 1.33 | 0.95 | 1824 | 1.05 | 0.81 | 1.36 | 0.72 |
| m_hdl_p | 1825 | 0.94 | 0.74 | 1.21 | 0.64 | 1229 | 0.92 | 0.68 | 1.24 | 0.58 | 1229 | 0.99 | 0.73 | 1.33 | 0.94 | 1825 | 1.05 | 0.81 | 1.36 | 0.73 |
| sm | 1824 | 0.94 | 0.73 | 1.21 | 0.65 | 1228 | 0.87 | 0.65 | 1.17 | 0.36 | 1228 | 0.89 | 0.66 | 1.20 | 0.46 | 1824 | 0.96 | 0.73 | 1.25 | 0.74 |
| glc | 1819 | 0.94 | 0.74 | 1.21 | 0.65 | 1225 | 0.88 | 0.63 | 1.23 | 0.45 | 1225 | 0.85 | 0.61 | 1.20 | 0.36 | 1819 | 0.97 | 0.75 | 1.24 | 0.78 |
| tyr | 1820 | 0.94 | 0.73 | 1.22 | 0.66 | 1225 | 0.94 | 0.69 | 1.28 | 0.71 | 1225 | 0.88 | 0.65 | 1.19 | 0.40 | 1820 | 0.92 | 0.70 | 1.22 | 0.57 |
| gln | 1824 | 0.96 | 0.74 | 1.23 | 0.74 | 1228 | 0.96 | 0.72 | 1.27 | 0.76 | 1228 | 0.98 | 0.73 | 1.31 | 0.88 | 1824 | 1.07 | 0.82 | 1.41 | 0.61 |
| xl_hdl_p | 1823 | 0.97 | 0.75 | 1.24 | 0.79 | 1228 | 0.87 | 0.65 | 1.17 | 0.35 | 1228 | 0.94 | 0.71 | 1.26 | 0.70 | 1823 | 1.00 | 0.76 | 1.30 | 0.97 |
| xl_hdl_1 | 1823 | 0.97 | 0.76 | 1.24 | 0.80 | 1228 | 0.87 | 0.65 | 1.17 | 0.35 | 1228 | 0.94 | 0.70 | 1.26 | 0.69 | 1823 | 0.99 | 0.76 | 1.30 | 0.97 |
| faw3_fa | 1823 | 0.97 | 0.75 | 1.25 | 0.80 | 1227 | 1.03 | 0.78 | 1.38 | 0.82 | 1227 | 1.03 | 0.77 | 1.37 | 0.84 | 1823 | 0.96 | 0.73 | 1.26 | 0.77 |
| faw6 | 1823 | 0.97 | 0.75 | 1.25 | 0.82 | 1227 | 0.87 | 0.64 | 1.18 | 0.37 | 1227 | 0.85 | 0.63 | 1.16 | 0.30 | 1823 | 0.97 | 0.74 | 1.27 | 0.83 |
| xs_vidl_1 | 1824 | 0.97 | 0.75 | 1.26 | 0.82 | 1228 | 0.97 | 0.71 | 1.31 | 0.82 | 1228 | 0.94 | 0.69 | 1.27 | 0.68 | 1824 | 1.01 | 0.77 | 1.33 | 0.92 |
| xs_vild_c | 1825 | 0.98 | 0.76 | 1.27 | 0.86 | 1229 | 0.96 | 0.70 | 1.30 | 0.78 | 1229 | 0.95 | 0.69 | 1.29 | 0.73 | 1825 | 1.02 | 0.77 | 1.34 | 0.91 |
| pufa | 1823 | 0.98 | 0.76 | 1.27 | 0.89 | 1227 | 0.90 | 0.67 | 1.22 | 0.51 | 1227 | 0.88 | 0.65 | 1.19 | 0.42 | 1823 | 0.98 | 0.75 | 1.29 | 0.89 |
| totcho | 1824 | 0.98 | 0.77 | 1.26 | 0.89 | 1228 | 0.90 | 0.68 | 1.20 | 0.48 | 1228 | 0.92 | 0.69 | 1.24 | 0.60 | 1824 | 1.02 | 0.78 | 1.33 | 0.90 |
| xs_vidl_p | 1824 | 0.98 | 0.76 | 1.27 | 0.90 | 1228 | 0.99 | 0.73 | 1.33 | 0.92 | 1228 | 0.95 | 0.70 | 1.29 | 0.75 | 1824 | 1.03 | 0.78 | 1.35 | 0.84 |
| dha | 1823 | 0.99 | 0.77 | 1.27 | 0.91 | 1227 | 1.03 | 0.78 | 1.37 | 0.82 | 1227 | 1.02 | 0.77 | 1.36 | 0.87 | 1823 | 1.01 | 0.77 | 1.32 | 0.95 |
| I_\|dI_tg | 1824 | 0.99 | 0.77 | 1.27 | 0.91 | 1228 | 1.02 | 0.76 | 1.37 | 0.89 | 1228 | 1.00 | 0.75 | 1.35 | 0.97 | 1824 | 1.06 | 0.81 | 1.38 | 0.67 |


|  | Minimally adjusted model |  |  |  |  | Fully adjusted model |  |  |  |  | Minimally adjusted(Complete case) |  |  |  |  | Minimally adjusted model censored |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| crea | 1820 | 0.99 | 0.77 | 1.26 | 0.92 | 1225 | 1.11 | 0.84 | 1.45 | 0.46 | 1225 | 1.07 | 0.82 | 1.41 | 0.61 | 1820 | 1.03 | 0.80 | 1.33 | 0.82 |
| s_ldi_tg | 1824 | 0.99 | 0.77 | 1.26 | 0.92 | 1228 | 0.97 | 0.72 | 1.30 | 0.84 | 1228 | 0.93 | 0.69 | 1.24 | 0.60 | 1824 | 1.01 | 0.77 | 1.31 | 0.95 |
| la | 1823 | 0.99 | 0.77 | 1.27 | 0.93 | 1227 | 0.89 | 0.66 | 1.20 | 0.43 | 1227 | 0.86 | 0.64 | 1.16 | 0.33 | 1823 | 0.98 | 0.75 | 1.29 | 0.91 |
| pc | 1824 | 0.99 | 0.77 | 1.27 | 0.94 | 1228 | 0.90 | 0.67 | 1.21 | 0.47 | 1228 | 0.92 | 0.68 | 1.24 | 0.58 | 1824 | 1.03 | 0.79 | 1.34 | 0.83 |
| Idl_tg | 1825 | 0.99 | 0.77 | 1.27 | 0.94 | 1229 | 1.01 | 0.76 | 1.35 | 0.95 | 1229 | 0.99 | 0.74 | 1.33 | 0.97 | 1825 | 1.06 | 0.81 | 1.38 | 0.69 |
| xl_hdl_ce | 1823 | 1.00 | 0.78 | 1.27 | 0.98 | 1228 | 0.87 | 0.65 | 1.17 | 0.35 | 1228 | 0.92 | 0.69 | 1.24 | 0.59 | 1823 | 1.00 | 0.77 | 1.30 | 0.98 |
| xl_hdl_c | 1823 | 1.00 | 0.78 | 1.28 | 0.99 | 1228 | 0.88 | 0.66 | 1.18 | 0.39 | 1228 | 0.94 | 0.70 | 1.26 | 0.67 | 1823 | 1.01 | 0.77 | 1.31 | 0.96 |
| 1_hdl_tg | 1824 | 1.00 | 0.78 | 1.28 | 0.99 | 1228 | 1.03 | 0.79 | 1.35 | 0.82 | 1228 | 1.08 | 0.83 | 1.42 | 0.56 | 1824 | 1.10 | 0.85 | 1.41 | 0.47 |
| apob | 1824 | 1.01 | 0.78 | 1.29 | 0.96 | 1228 | 0.96 | 0.72 | 1.30 | 0.81 | 1228 | 0.92 | 0.68 | 1.24 | 0.58 | 1824 | 0.99 | 0.76 | 1.30 | 0.95 |
| xl_hdl_fc | 1823 | 1.01 | 0.79 | 1.29 | 0.95 | 1228 | 0.90 | 0.68 | 1.21 | 0.50 | 1228 | 0.98 | 0.73 | 1.30 | 0.87 | 1823 | 1.02 | 0.79 | 1.33 | 0.86 |
| s_hdi_fc | 1825 | 1.02 | 0.80 | 1.30 | 0.88 | 1229 | 1.05 | 0.78 | 1.41 | 0.77 | 1229 | 1.06 | 0.79 | 1.42 | 0.68 | 1825 | 1.10 | 0.85 | 1.42 | 0.47 |
| totpg | 1824 | 1.02 | 0.80 | 1.31 | 0.88 | 1228 | 0.94 | 0.70 | 1.26 | 0.67 | 1228 | 0.96 | 0.71 | 1.28 | 0.76 | 1824 | 1.06 | 0.81 | 1.37 | 0.69 |
| xs_vidl_ce | 1825 | 1.02 | 0.79 | 1.32 | 0.88 | 1229 | 1.01 | 0.74 | 1.37 | 0.96 | 1229 | 1.00 | 0.73 | 1.36 | 0.99 | 1825 | 1.05 | 0.80 | 1.38 | 0.73 |
| glol | 409 | 1.02 | 0.49 | 2.12 | 0.96 | 288 | 1.17 | 0.42 | 3.24 | 0.76 | 288 | 1.09 | 0.41 | 2.88 | 0.86 | 409 | 1.27 | 0.56 | 2.89 | 0.56 |
| s_vidl_ce | 1824 | 1.02 | 0.79 | 1.32 | 0.86 | 1228 | 1.01 | 0.75 | 1.37 | 0.93 | 1228 | 0.96 | 0.71 | 1.30 | 0.81 | 1824 | 1.04 | 0.80 | 1.37 | 0.76 |
| bohbut | 1777 | 1.03 | 0.84 | 1.25 | 0.78 | 1195 | 1.03 | 0.84 | 1.27 | 0.76 | 1195 | 1.04 | 0.85 | 1.27 | 0.70 | 1777 | 1.01 | 0.81 | 1.27 | 0.90 |
| idl_tg | 1824 | 1.03 | 0.80 | 1.32 | 0.82 | 1228 | 1.09 | 0.82 | 1.44 | 0.57 | 1228 | 1.05 | 0.79 | 1.40 | 0.72 | 1824 | 1.10 | 0.85 | 1.43 | 0.46 |
| phe | 1825 | 1.03 | 0.81 | 1.31 | 0.80 | 1229 | 1.08 | 0.82 | 1.43 | 0.59 | 1229 | 1.03 | 0.78 | 1.36 | 0.84 | 1825 | 1.06 | 0.83 | 1.37 | 0.63 |
| remnant_c | 1825 | 1.03 | 0.80 | 1.33 | 0.80 | 1229 | 1.00 | 0.74 | 1.35 | 0.98 | 1229 | 0.96 | 0.71 | 1.30 | 0.81 | 1825 | 1.03 | 0.79 | 1.35 | 0.81 |
| faw3 | 1823 | 1.04 | 0.81 | 1.32 | 0.77 | 1227 | 1.07 | 0.82 | 1.41 | 0.62 | 1227 | 1.04 | 0.79 | 1.38 | 0.77 | 1823 | 1.04 | 0.80 | 1.35 | 0.78 |
| ace | 1825 | 1.05 | 0.85 | 1.31 | 0.64 | 1229 | 1.04 | 0.83 | 1.29 | 0.75 | 1229 | 1.07 | 0.87 | 1.32 | 0.52 | 1825 | 1.07 | 0.88 | 1.31 | 0.49 |
| val | 1824 | 1.06 | 0.83 | 1.35 | 0.67 | 1229 | 1.15 | 0.86 | 1.55 | 0.35 | 1229 | 1.06 | 0.80 | 1.40 | 0.71 | 1824 | 1.02 | 0.78 | 1.32 | 0.90 |
| alb | 1825 | 1.06 | 0.82 | 1.37 | 0.67 | 1229 | 1.14 | 0.84 | 1.54 | 0.39 | 1229 | 1.15 | 0.85 | 1.56 | 0.37 | 1825 | 1.07 | 0.81 | 1.41 | 0.65 |
| S_vidl_c | 1824 | 1.06 | 0.83 | 1.37 | 0.63 | 1228 | 1.07 | 0.80 | 1.43 | 0.65 | 1228 | 1.01 | 0.75 | 1.35 | 0.95 | 1824 | 1.09 | 0.83 | 1.42 | 0.54 |
| apob_apoa1 | 1824 | 1.07 | 0.84 | 1.38 | 0.58 | 1228 | 1.07 | 0.80 | 1.44 | 0.66 | 1228 | 0.99 | 0.74 | 1.33 | 0.96 | 1824 | 1.02 | 0.78 | 1.34 | 0.88 |
| his | 1820 | 1.09 | 0.86 | 1.40 | 0.48 | 1225 | 1.22 | 0.92 | 1.63 | 0.17 | 1225 | 1.22 | 0.92 | 1.63 | 0.17 | 1820 | 1.15 | 0.88 | 1.49 | 0.31 |
| s_hdl_pl | 1825 | 1.10 | 0.87 | 1.39 | 0.44 | 1229 | 1.15 | 0.86 | 1.54 | 0.33 | 1229 | 1.16 | 0.87 | 1.54 | 0.30 | 1825 | 1.18 | 0.92 | 1.51 | 0.20 |


|  | Minimally adjusted model |  |  |  |  | Fully adjusted model |  |  |  |  | Minimally adjusted(Complete case) |  |  |  |  | Minimally adjusted model censored |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xs_vidl_tg | 1824 | 1.10 | 0.87 | 1.39 | 0.43 | 1228 | 1.15 | 0.88 | 1.51 | 0.30 | 1228 | 1.09 | 0.83 | 1.44 | 0.53 | 1824 | 1.14 | 0.89 | 1.47 | 0.30 |
| totfa | 1823 | 1.10 | 0.87 | 1.40 | 0.42 | 1227 | 1.06 | 0.80 | 1.41 | 0.68 | 1227 | 1.02 | 0.77 | 1.37 | 0.87 | 1823 | 1.10 | 0.85 | 1.42 | 0.48 |
| mufa_fa | 1823 | 1.11 | 0.88 | 1.41 | 0.37 | 1227 | 1.13 | 0.87 | 1.47 | 0.36 | 1227 | 1.10 | 0.83 | 1.45 | 0.51 | 1823 | 1.13 | 0.88 | 1.45 | 0.35 |
| S_vill_fc | 1824 | 1.12 | 0.88 | 1.42 | 0.36 | 1228 | 1.15 | 0.87 | 1.51 | 0.33 | 1228 | 1.08 | 0.82 | 1.43 | 0.59 | 1824 | 1.15 | 0.89 | 1.48 | 0.30 |
| m_vldi_fc | 1822 | 1.13 | 0.89 | 1.43 | 0.32 | 1227 | 1.12 | 0.85 | 1.49 | 0.42 | 1227 | 1.04 | 0.79 | 1.38 | 0.78 | 1822 | 1.09 | 0.85 | 1.41 | 0.49 |
| acace | 1825 | 1.13 | 0.98 | 1.30 | 0.10 | 1229 | 1.13 | 0.99 | 1.30 | 0.08 | 1229 | 1.14 | 1.00 | 1.30 | 0.05 | 1825 | 1.13 | 0.97 | 1.31 | 0.13 |
| m_vidi_l | 1822 | 1.13 | 0.89 | 1.43 | 0.32 | 1227 | 1.12 | 0.84 | 1.49 | 0.43 | 1227 | 1.04 | 0.78 | 1.38 | 0.80 | 1822 | 1.09 | 0.84 | 1.41 | 0.52 |
| s_vidl_। | 1824 | 1.13 | 0.89 | 1.43 | 0.32 | 1228 | 1.15 | 0.88 | 1.52 | 0.31 | 1228 | 1.08 | 0.82 | 1.43 | 0.59 | 1824 | 1.14 | 0.88 | 1.47 | 0.31 |
| m_vidl_p | 1822 | 1.13 | 0.89 | 1.43 | 0.31 | 1227 | 1.12 | 0.85 | 1.49 | 0.42 | 1227 | 1.04 | 0.78 | 1.38 | 0.80 | 1822 | 1.09 | 0.84 | 1.41 | 0.51 |
| s_vidl_pl | 1824 | 1.13 | 0.89 | 1.44 | 0.31 | 1228 | 1.16 | 0.89 | 1.53 | 0.27 | 1228 | 1.10 | 0.83 | 1.45 | 0.51 | 1824 | 1.16 | 0.90 | 1.50 | 0.25 |
| m_vldi_tg | 1823 | 1.13 | 0.90 | 1.43 | 0.30 | 1228 | 1.12 | 0.84 | 1.48 | 0.43 | 1228 | 1.04 | 0.78 | 1.38 | 0.80 | 1823 | 1.09 | 0.84 | 1.41 | 0.53 |
| s_vidl_p | 1824 | 1.13 | 0.89 | 1.44 | 0.30 | 1228 | 1.16 | 0.88 | 1.53 | 0.28 | 1228 | 1.09 | 0.82 | 1.43 | 0.56 | 1824 | 1.14 | 0.89 | 1.47 | 0.30 |
| mufa | 1823 | 1.14 | 0.90 | 1.43 | 0.27 | 1227 | 1.11 | 0.85 | 1.46 | 0.44 | 1227 | 1.08 | 0.81 | 1.43 | 0.60 | 1823 | 1.14 | 0.89 | 1.46 | 0.30 |
| xl_hdl_tg | 1822 | 1.14 | 0.90 | 1.45 | 0.27 | 1228 | 1.09 | 0.82 | 1.46 | 0.55 | 1228 | 1.07 | 0.81 | 1.43 | 0.63 | 1822 | 1.11 | 0.85 | 1.43 | 0.45 |
| sfa | 1823 | 1.15 | 0.91 | 1.45 | 0.25 | 1227 | 1.13 | 0.85 | 1.50 | 0.40 | 1227 | 1.09 | 0.82 | 1.44 | 0.57 | 1823 | 1.13 | 0.88 | 1.45 | 0.34 |
| I_vidi_ce | 1822 | 1.15 | 0.91 | 1.45 | 0.24 | 1227 | 1.15 | 0.87 | 1.53 | 0.33 | 1227 | 1.07 | 0.81 | 1.41 | 0.65 | 1822 | 1.12 | 0.87 | 1.44 | 0.39 |
| m_vldi_ce | 1824 | 1.15 | 0.91 | 1.46 | 0.23 | 1228 | 1.17 | 0.89 | 1.54 | 0.26 | 1228 | 1.10 | 0.83 | 1.45 | 0.50 | 1824 | 1.15 | 0.89 | 1.48 | 0.28 |
| vidl_d | 1825 | 1.15 | 0.90 | 1.47 | 0.25 | 1229 | 1.20 | 0.90 | 1.61 | 0.22 | 1229 | 1.09 | 0.82 | 1.46 | 0.54 | 1825 | 1.12 | 0.86 | 1.46 | 0.40 |
| vidl_c | 1825 | 1.15 | 0.91 | 1.47 | 0.24 | 1229 | 1.16 | 0.88 | 1.52 | 0.30 | 1229 | 1.10 | 0.83 | 1.45 | 0.52 | 1825 | 1.16 | 0.89 | 1.49 | 0.27 |
| xxi_vldi_tg | 1820 | 1.16 | 0.92 | 1.46 | 0.23 | 1228 | 1.21 | 0.91 | 1.60 | 0.19 | 1228 | 1.12 | 0.85 | 1.48 | 0.41 | 1820 | 1.17 | 0.91 | 1.49 | 0.22 |
| xxi_vldil_ | 1819 | 1.16 | 0.92 | 1.46 | 0.22 | 1227 | 1.21 | 0.91 | 1.59 | 0.19 | 1227 | 1.13 | 0.86 | 1.48 | 0.40 | 1819 | 1.17 | 0.91 | 1.50 | 0.21 |
| s_vidl_tg | 1823 | 1.16 | 0.92 | 1.46 | 0.22 | 1228 | 1.19 | 0.91 | 1.55 | 0.20 | 1228 | 1.11 | 0.85 | 1.46 | 0.45 | 1823 | 1.16 | 0.90 | 1.48 | 0.26 |
| xxl_vldi_pl | 1819 | 1.16 | 0.92 | 1.46 | 0.20 | 1227 | 1.21 | 0.92 | 1.60 | 0.18 | 1227 | 1.13 | 0.86 | 1.49 | 0.37 | 1819 | 1.18 | 0.92 | 1.50 | 0.19 |
| I_vidl_c | 1822 | 1.16 | 0.93 | 1.46 | 0.20 | 1227 | 1.16 | 0.88 | 1.54 | 0.29 | 1227 | 1.08 | 0.81 | 1.43 | 0.60 | 1822 | 1.12 | 0.87 | 1.45 | 0.37 |
| ile | 1825 | 1.17 | 0.92 | 1.48 | 0.20 | 1229 | 1.29 | 0.97 | 1.71 | 0.08 | 1229 | 1.16 | 0.89 | 1.52 | 0.28 | 1825 | 1.12 | 0.87 | 1.44 | 0.39 |
| I_vidi_pl | 1823 | 1.17 | 0.93 | 1.47 | 0.19 | 1228 | 1.16 | 0.87 | 1.53 | 0.31 | 1228 | 1.08 | 0.81 | 1.42 | 0.61 | 1823 | 1.12 | 0.87 | 1.44 | 0.39 |
| I_vidi_p | 1823 | 1.17 | 0.93 | 1.47 | 0.19 | 1228 | 1.16 | 0.87 | 1.54 | 0.31 | 1228 | 1.08 | 0.81 | 1.42 | 0.61 | 1823 | 1.12 | 0.87 | 1.44 | 0.40 |


| I_vidl_I | Minimally adjusted model |  |  |  |  | Fully adjusted model |  |  |  |  | Minimally adjusted(Complete case) |  |  |  |  | Minimally adjusted model censored |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 1823 | 1.17 | 0.93 | 1.47 | 0.18 | 1228 | 1.16 | 0.87 | 1.54 | 0.31 | 1228 | 1.08 | 0.81 | 1.42 | 0.61 | 1823 | 1.12 | 0.87 | 1.44 | 0.39 |
| I_vidl_tg | 1823 | 1.17 | 0.93 | 1.47 | 0.18 | 1228 | 1.16 | 0.87 | 1.54 | 0.31 | 1228 | 1.08 | 0.81 | 1.42 | 0.61 | 1823 | 1.11 | 0.86 | 1.43 | 0.41 |
| hdil_tg | 1825 | 1.17 | 0.93 | 1.47 | 0.17 | 1229 | 1.22 | 0.95 | 1.57 | 0.13 | 1229 | 1.20 | 0.92 | 1.56 | 0.17 | 1825 | 1.22 | 0.96 | 1.54 | 0.11 |
| m_hdl_tg | 1824 | 1.17 | 0.93 | 1.48 | 0.18 | 1228 | 1.25 | 0.96 | 1.64 | 0.10 | 1228 | 1.21 | 0.92 | 1.58 | 0.17 | 1824 | 1.25 | 0.97 | 1.59 | 0.08 |
| I_vidi_fc | 1823 | 1.17 | 0.94 | 1.47 | 0.16 | 1228 | 1.16 | 0.88 | 1.54 | 0.29 | 1228 | 1.09 | 0.82 | 1.44 | 0.57 | 1823 | 1.13 | 0.88 | 1.45 | 0.36 |
| xxl_vidl_ce | 1822 | 1.18 | 0.94 | 1.48 | 0.16 | 1227 | 1.17 | 0.88 | 1.56 | 0.29 | 1227 | 1.09 | 0.82 | 1.44 | 0.55 | 1822 | 1.14 | 0.88 | 1.46 | 0.32 |
| xl_vidi_ce | 1822 | 1.18 | 0.94 | 1.48 | 0.15 | 1227 | 1.18 | 0.89 | 1.56 | 0.26 | 1227 | 1.09 | 0.82 | 1.44 | 0.54 | 1822 | 1.13 | 0.88 | 1.46 | 0.33 |
| m_vidl_c | 1823 | 1.18 | 0.94 | 1.49 | 0.15 | 1228 | 1.20 | 0.92 | 1.56 | 0.18 | 1228 | 1.12 | 0.86 | 1.47 | 0.40 | 1823 | 1.17 | 0.91 | 1.50 | 0.21 |
| leu | 1825 | 1.19 | 0.94 | 1.50 | 0.14 | 1229 | 1.29 | 0.98 | 1.70 | 0.07 | 1229 | 1.19 | 0.91 | 1.55 | 0.20 | 1825 | 1.14 | 0.89 | 1.47 | 0.29 |
| xl_vidl_c | 1823 | 1.19 | 0.95 | 1.49 | 0.13 | 1228 | 1.18 | 0.89 | 1.56 | 0.26 | 1228 | 1.10 | 0.83 | 1.45 | 0.51 | 1823 | 1.14 | 0.89 | 1.46 | 0.31 |
| serum_tg | 1825 | 1.19 | 0.96 | 1.49 | 0.12 | 1229 | 1.22 | 0.94 | 1.57 | 0.14 | 1229 | 1.15 | 0.88 | 1.51 | 0.30 | 1825 | 1.18 | 0.93 | 1.49 | 0.18 |
| xl_vidl_p | 1823 | 1.19 | 0.95 | 1.49 | 0.12 | 1228 | 1.18 | 0.89 | 1.57 | 0.25 | 1228 | 1.10 | 0.83 | 1.46 | 0.50 | 1823 | 1.13 | 0.88 | 1.46 | 0.32 |
| xl_vill_tg | 1822 | 1.20 | 0.96 | 1.50 | 0.11 | 1228 | 1.18 | 0.89 | 1.57 | 0.25 | 1228 | 1.10 | 0.83 | 1.45 | 0.51 | 1822 | 1.13 | 0.88 | 1.46 | 0.32 |
| m_vidi_pl | 1823 | 1.20 | 0.96 | 1.50 | 0.11 | 1228 | 1.21 | 0.93 | 1.57 | 0.15 | 1228 | 1.14 | 0.87 | 1.48 | 0.35 | 1823 | 1.18 | 0.92 | 1.50 | 0.19 |
| xl_vidi_1 | 1822 | 1.20 | 0.96 | 1.50 | 0.11 | 1228 | 1.18 | 0.89 | 1.57 | 0.25 | 1228 | 1.10 | 0.83 | 1.45 | 0.50 | 1822 | 1.14 | 0.89 | 1.46 | 0.30 |
| s_hdl_tg | 1824 | 1.20 | 0.95 | 1.52 | 0.12 | 1228 | 1.29 | 0.99 | 1.68 | 0.06 | 1228 | 1.21 | 0.92 | 1.58 | 0.16 | 1824 | 1.22 | 0.95 | 1.56 | 0.12 |
| sfa_fa | 1823 | 1.20 | 0.93 | 1.55 | 0.16 | 1227 | 1.34 | 0.99 | 1.82 | 0.06 | 1227 | 1.29 | 0.96 | 1.72 | 0.09 | 1823 | 1.15 | 0.87 | 1.52 | 0.32 |
| vidl_tg | 1825 | 1.21 | 0.97 | 1.50 | 0.09 | 1229 | 1.23 | 0.95 | 1.59 | 0.11 | 1229 | 1.16 | 0.89 | 1.52 | 0.27 | 1825 | 1.18 | 0.93 | 1.49 | 0.17 |
| xl_vidl_pl | 1821 | 1.21 | 0.97 | 1.51 | 0.10 | 1227 | 1.19 | 0.90 | 1.57 | 0.23 | 1227 | 1.11 | 0.84 | 1.46 | 0.47 | 1821 | 1.15 | 0.90 | 1.47 | 0.27 |
| xl_vidl_fc | 1822 | 1.21 | 0.97 | 1.51 | 0.10 | 1228 | 1.19 | 0.90 | 1.57 | 0.23 | 1228 | 1.11 | 0.84 | 1.46 | 0.46 | 1822 | 1.15 | 0.90 | 1.48 | 0.26 |
| xxi_vidl_c | 1821 | 1.21 | 0.96 | 1.52 | 0.10 | 1228 | 1.18 | 0.89 | 1.56 | 0.24 | 1228 | 1.11 | 0.84 | 1.46 | 0.47 | 1821 | 1.16 | 0.90 | 1.48 | 0.25 |
| tg_pg | 1824 | 1.23 | 0.98 | 1.54 | 0.08 | 1228 | 1.28 | 0.98 | 1.68 | 0.07 | 1228 | 1.18 | 0.90 | 1.56 | 0.23 | 1824 | 1.17 | 0.91 | 1.50 | 0.22 |
| xxi_vidi_fc | 1820 | 1.23 | 0.99 | 1.53 | 0.07 | 1227 | 1.21 | 0.92 | 1.60 | 0.17 | 1227 | 1.14 | 0.87 | 1.50 | 0.35 | 1820 | 1.17 | 0.92 | 1.50 | 0.20 |
| ldi_d | 1825 | 1.23 | 0.98 | 1.55 | 0.08 | 1229 | 1.39 | 1.07 | 1.82 | 0.01 | 1229 | 1.45 | 1.11 | 1.88 | 0.01 | 1825 | 1.31 | 1.03 | 1.67 | 0.03 |
| xxl_vldil_p | 1825 | 1.26 | 1.04 | 1.53 | 0.02 | 1229 | 1.31 | 1.02 | 1.67 | 0.03 | 1229 | 1.24 | 0.96 | 1.60 | 0.10 | 1825 | 1.22 | 0.99 | 1.51 | 0.07 |
| gp | 1825 | 1.31 | 1.04 | 1.64 | 0.02 | 1229 | 1.29 | 0.97 | 1.71 | 0.08 | 1229 | 1.19 | 0.90 | 1.57 | 0.23 | 1825 | 1.26 | 0.98 | 1.61 | 0.07 |

Appendix B Table B 3: Proportional hazard assumption test in the minimally adjusted cox
model for clinical progression, metastases or death and any progression model for clinical progression, metastases or death and any progression

|  | Clinical Progression |  |  |  |  |  |  | Metastases |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| metabolite | metabolite $p$-value | gleason $p$ - <br> value | Stage p- <br> value | PS p-value | age $p$-value | centre p - <br> value | global p- <br> value | metabolite p-value | gleason $p$ - <br> value | Stage p- <br> value | PS p-value | age p -value | centre p - <br> value | global p- <br> value |
| gp | 0.31 | 0.23 | 0.77 | 0.11 | 0.05 | 0.89 | 0.10 | 0.47 | 0.11 | 0.57 | 0.73 | 0.28 | 0.81 | 0.51 |
| hdl_c | 0.40 | 0.23 | 0.78 | 0.10 | 0.04 | 0.97 | 0.11 | 0.97 | 0.10 | 0.57 | 0.68 | 0.27 | 0.78 | 0.55 |
| hdl_d | 0.82 | 0.22 | 0.77 | 0.10 | 0.05 | 0.93 | 0.13 | 0.92 | 0.09 | 0.57 | 0.69 | 0.27 | 0.79 | 0.53 |
| hdil_tg | 0.60 | 0.23 | 0.79 | 0.11 | 0.05 | 0.93 | 0.13 | 0.40 | 0.09 | 0.52 | 0.73 | 0.29 | 0.71 | 0.45 |
| hdi2_c | 0.31 | 0.23 | 0.79 | 0.09 | 0.04 | 0.98 | 0.10 | 0.96 | 0.10 | 0.57 | 0.68 | 0.27 | 0.78 | 0.55 |
| hdl3_c | 0.37 | 0.24 | 0.77 | 0.10 | 0.04 | 0.93 | 0.10 | 0.54 | 0.11 | 0.59 | 0.67 | 0.26 | 0.77 | 0.50 |
| his | 0.87 | 0.22 | 0.79 | 0.11 | 0.05 | 0.96 | 0.14 | 0.57 | 0.08 | 0.56 | 0.76 | 0.29 | 0.74 | 0.48 |
| idl_c | 0.14 | 0.21 | 0.84 | 0.11 | 0.05 | 0.99 | 0.06 | 0.52 | 0.08 | 0.55 | 0.72 | 0.27 | 0.74 | 0.46 |
| idl_ce | 0.19 | 0.21 | 0.83 | 0.11 | 0.05 | 0.99 | 0.07 | 0.54 | 0.08 | 0.55 | 0.72 | 0.27 | 0.74 | 0.46 |
| idl_fc | 0.07 | 0.21 | 0.83 | 0.11 | 0.05 | 0.97 | 0.04 | 0.48 | 0.09 | 0.56 | 0.73 | 0.27 | 0.74 | 0.46 |
| idl_1 | 0.11 | 0.21 | 0.84 | 0.11 | 0.05 | 0.99 | 0.06 | 0.43 | 0.08 | 0.55 | 0.73 | 0.27 | 0.74 | 0.43 |
| idl_p | 0.11 | 0.21 | 0.85 | 0.11 | 0.05 | 0.99 | 0.06 | 0.40 | 0.08 | 0.55 | 0.73 | 0.27 | 0.74 | 0.43 |
| idl_pl | 0.10 | 0.21 | 0.84 | 0.11 | 0.05 | 0.99 | 0.05 | 0.44 | 0.08 | 0.55 | 0.73 | 0.27 | 0.74 | 0.44 |
| idl_tg | 0.11 | 0.20 | 0.88 | 0.10 | 0.04 | 0.94 | 0.06 | 0.14 | 0.07 | 0.50 | 0.70 | 0.26 | 0.79 | 0.29 |
| ile | 0.20 | 0.24 | 0.79 | 0.09 | 0.04 | 0.94 | 0.08 | 0.84 | 0.08 | 0.53 | 0.73 | 0.29 | 0.76 | 0.52 |
| I_hdl_c | 0.48 | 0.23 | 0.78 | 0.10 | 0.05 | 0.96 | 0.12 | 0.83 | 0.09 | 0.57 | 0.68 | 0.27 | 0.78 | 0.53 |
| I_hdl_ce | 0.44 | 0.23 | 0.78 | 0.10 | 0.05 | 0.96 | 0.11 | 0.83 | 0.09 | 0.56 | 0.68 | 0.27 | 0.78 | 0.53 |
| I_hdl_fc | 0.64 | 0.23 | 0.78 | 0.10 | 0.05 | 0.95 | 0.13 | 0.83 | 0.10 | 0.57 | 0.69 | 0.27 | 0.78 | 0.54 |
| I_hdl_I | 0.46 | 0.23 | 0.78 | 0.10 | 0.04 | 0.96 | 0.11 | 0.92 | 0.10 | 0.57 | 0.69 | 0.27 | 0.78 | 0.54 |
| I_hdl_p | 0.45 | 0.23 | 0.78 | 0.10 | 0.04 | 0.96 | 0.11 | 0.93 | 0.09 | 0.57 | 0.69 | 0.27 | 0.78 | 0.54 |
| 1_hdl_pl | 0.42 | 0.23 | 0.78 | 0.10 | 0.04 | 0.96 | 0.11 | 0.96 | 0.10 | 0.57 | 0.69 | 0.27 | 0.79 | 0.54 |
| 1_hdl_tg | 0.84 | 0.22 | 0.79 | 0.11 | 0.04 | 0.94 | 0.13 | 0.22 | 0.08 | 0.56 | 0.75 | 0.25 | 0.77 | 0.35 |
| I_\|dI_c | 0.11 | 0.21 | 0.84 | 0.11 | 0.05 | 0.99 | 0.05 | 0.53 | 0.08 | 0.55 | 0.72 | 0.27 | 0.74 | 0.46 |
| I_IdI_ce | 0.12 | 0.21 | 0.84 | 0.11 | 0.05 | 0.99 | 0.06 | 0.54 | 0.08 | 0.55 | 0.72 | 0.27 | 0.74 | 0.46 |
| I_ldi_fc | 0.08 | 0.21 | 0.84 | 0.11 | 0.05 | 0.98 | 0.04 | 0.55 | 0.09 | 0.55 | 0.72 | 0.27 | 0.74 | 0.47 |


|  | Clinical Progression |  |  |  |  |  |  | Metastases |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| metabolite | metabolite $p$-value | gleason p value | Stage pvalue | PS p-value | age $p$-value | centre p- <br> value | global pvalue | metabolite p-value | gleason pvalue | Stage pvalue | PS p-value | age $p$-value | centre p - <br> value | global pvalue |
| I_IdI_\| | 0.11 | 0.21 | 0.84 | 0.11 | 0.05 | 0.99 | 0.05 | 0.48 | 0.08 | 0.55 | 0.72 | 0.27 | 0.74 | 0.45 |
| I_\|di_p | 0.11 | 0.21 | 0.84 | 0.11 | 0.05 | 0.99 | 0.06 | 0.46 | 0.08 | 0.55 | 0.72 | 0.27 | 0.74 | 0.44 |
| I_ldi_pl | 0.16 | 0.21 | 0.83 | 0.11 | 0.05 | 0.99 | 0.07 | 0.50 | 0.08 | 0.55 | 0.72 | 0.27 | 0.74 | 0.45 |
| I_IdI_tg | 0.11 | 0.20 | 0.87 | 0.11 | 0.04 | 0.94 | 0.06 | 0.16 | 0.07 | 0.51 | 0.72 | 0.26 | 0.78 | 0.30 |
| I_vidl_c | 0.35 | 0.23 | 0.80 | 0.11 | 0.04 | 0.90 | 0.11 | 0.92 | 0.10 | 0.53 | 0.73 | 0.28 | 0.76 | 0.54 |
| I_vidi_ce | 0.51 | 0.23 | 0.80 | 0.11 | 0.05 | 0.92 | 0.13 | 0.76 | 0.09 | 0.53 | 0.73 | 0.28 | 0.75 | 0.53 |
| I_vidl_fc | 0.23 | 0.23 | 0.79 | 0.11 | 0.04 | 0.89 | 0.09 | 0.93 | 0.10 | 0.53 | 0.73 | 0.27 | 0.77 | 0.54 |
| I_vidl_I | 0.21 | 0.23 | 0.79 | 0.10 | 0.04 | 0.89 | 0.09 | 0.92 | 0.10 | 0.53 | 0.73 | 0.27 | 0.77 | 0.54 |
| I_vidl_p | 0.20 | 0.23 | 0.79 | 0.10 | 0.04 | 0.89 | 0.08 | 0.90 | 0.10 | 0.53 | 0.73 | 0.27 | 0.77 | 0.54 |
| I_vidi_pl | 0.23 | 0.23 | 0.79 | 0.11 | 0.04 | 0.89 | 0.09 | 0.93 | 0.10 | 0.53 | 0.73 | 0.27 | 0.77 | 0.54 |
| I_vidl_tg | 0.16 | 0.23 | 0.79 | 0.10 | 0.04 | 0.88 | 0.07 | 0.85 | 0.10 | 0.54 | 0.73 | 0.27 | 0.78 | 0.54 |
| la | 0.54 | 0.22 | 0.80 | 0.11 | 0.05 | 0.96 | 0.12 | 0.50 | 0.08 | 0.56 | 0.73 | 0.28 | 0.76 | 0.46 |
| la_fa | 0.09 | 0.22 | 0.79 | 0.11 | 0.05 | 0.93 | 0.05 | 0.60 | 0.09 | 0.55 | 0.68 | 0.27 | 0.78 | 0.49 |
| lac | 0.45 | 0.22 | 0.77 | 0.10 | 0.05 | 0.96 | 0.11 | 0.86 | 0.09 | 0.56 | 0.69 | 0.29 | 0.82 | 0.52 |
| ldi_c | 0.09 | 0.22 | 0.83 | 0.11 | 0.05 | 0.99 | 0.05 | 0.54 | 0.09 | 0.55 | 0.72 | 0.27 | 0.74 | 0.46 |
| \|dlıd | 0.64 | 0.22 | 0.81 | 0.11 | 0.05 | 0.95 | 0.13 | 0.56 | 0.09 | 0.53 | 0.68 | 0.30 | 0.76 | 0.49 |
| ldil_tg | 0.19 | 0.21 | 0.85 | 0.11 | 0.04 | 0.95 | 0.08 | 0.17 | 0.07 | 0.52 | 0.72 | 0.26 | 0.77 | 0.31 |
| leu | 0.54 | 0.23 | 0.78 | 0.10 | 0.04 | 0.95 | 0.12 | 0.65 | 0.08 | 0.52 | 0.75 | 0.29 | 0.79 | 0.51 |
| m_hdl_c | 0.02 | 0.23 | 0.83 | 0.09 | 0.03 | 0.99 | 0.02 | 0.96 | 0.09 | 0.56 | 0.68 | 0.27 | 0.78 | 0.53 |
| m_hdl_ce | 0.02 | 0.23 | 0.83 | 0.09 | 0.03 | 0.99 | 0.02 | 0.94 | 0.09 | 0.56 | 0.68 | 0.27 | 0.78 | 0.53 |
| m_hdl_fc | 0.08 | 0.24 | 0.80 | 0.10 | 0.03 | 1.00 | 0.05 | 0.60 | 0.10 | 0.58 | 0.69 | 0.29 | 0.80 | 0.51 |
| m_hdl_1 | 0.02 | 0.24 | 0.81 | 0.09 | 0.03 | 1.00 | 0.02 | 0.73 | 0.09 | 0.57 | 0.69 | 0.29 | 0.79 | 0.51 |
| m_hdl_p | 0.01 | 0.24 | 0.81 | 0.09 | 0.03 | 1.00 | 0.02 | 0.69 | 0.09 | 0.57 | 0.69 | 0.29 | 0.79 | 0.50 |
| m_hdl_pl | 0.02 | 0.25 | 0.80 | 0.09 | 0.03 | 1.00 | 0.02 | 0.60 | 0.09 | 0.57 | 0.69 | 0.29 | 0.79 | 0.49 |
| m_hdl_tg | 0.26 | 0.24 | 0.76 | 0.10 | 0.05 | 0.90 | 0.10 | 0.28 | 0.09 | 0.53 | 0.73 | 0.32 | 0.70 | 0.41 |


|  | Clinical Progression |  |  |  |  |  |  | Metastases |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\underline{\text { metabolite }}$ | metabolite p-value | $\begin{aligned} & \hline \text { gleason p- } \\ & \text { value } \\ & \hline \end{aligned}$ | Stage pvalue | PS p-value | age p-value | centre pvalue | global pvalue | metabolite $p$-value | $\begin{aligned} & \hline \text { gleason } \mathrm{p} \text { - } \\ & \text { value } \\ & \hline \end{aligned}$ | Stage pvalue | PS p-value | age $p$-value | centre p- <br> value | global pvalue |
| m_ldi_c | 0.08 | 0.22 | 0.83 | 0.11 | 0.05 | 0.99 | 0.05 | 0.53 | 0.09 | 0.55 | 0.71 | 0.27 | 0.74 | 0.46 |
| m_ldi_ce | 0.08 | 0.22 | 0.83 | 0.11 | 0.05 | 0.99 | 0.04 | 0.53 | 0.09 | 0.55 | 0.72 | 0.27 | 0.74 | 0.46 |
| m_ld_ff | 0.13 | 0.22 | 0.83 | 0.11 | 0.05 | 0.98 | 0.06 | 0.58 | 0.09 | 0.55 | 0.70 | 0.27 | 0.74 | 0.47 |
| m_ldi_l | 0.11 | 0.22 | 0.83 | 0.11 | 0.05 | 0.99 | 0.05 | 0.50 | 0.09 | 0.55 | 0.72 | 0.27 | 0.74 | 0.45 |
| m_ldi_p | 0.11 | 0.22 | 0.83 | 0.11 | 0.05 | 0.99 | 0.05 | 0.48 | 0.09 | 0.55 | 0.72 | 0.27 | 0.74 | 0.45 |
| m_ldi_pl | 0.35 | 0.22 | 0.82 | 0.11 | 0.05 | 0.98 | 0.10 | 0.57 | 0.08 | 0.55 | 0.71 | 0.27 | 0.74 | 0.46 |
| m_ldı_tg | 0.16 | 0.21 | 0.85 | 0.11 | 0.04 | 0.94 | 0.07 | 0.18 | 0.08 | 0.54 | 0.72 | 0.26 | 0.79 | 0.32 |
| m_vidl_c | 0.80 | 0.23 | 0.80 | 0.11 | 0.05 | 0.93 | 0.15 | 0.57 | 0.10 | 0.53 | 0.73 | 0.29 | 0.75 | 0.52 |
| m_vidl_ce | 0.84 | 0.24 | 0.80 | 0.11 | 0.05 | 0.96 | 0.15 | 0.45 | 0.10 | 0.53 | 0.73 | 0.29 | 0.73 | 0.48 |
| m_vldi_fc | 0.42 | 0.23 | 0.80 | 0.11 | 0.04 | 0.91 | 0.12 | 0.81 | 0.09 | 0.53 | 0.72 | 0.28 | 0.76 | 0.53 |
| m_vidi_l | 0.40 | 0.23 | 0.80 | 0.11 | 0.04 | 0.91 | 0.11 | 0.84 | 0.09 | 0.53 | 0.72 | 0.28 | 0.76 | 0.53 |
| m_vidl_p | 0.37 | 0.23 | 0.80 | 0.11 | 0.04 | 0.91 | 0.11 | 0.87 | 0.09 | 0.53 | 0.72 | 0.28 | 0.76 | 0.53 |
| m_vldi_pl | 0.50 | 0.23 | 0.79 | 0.11 | 0.05 | 0.91 | 0.13 | 0.75 | 0.10 | 0.53 | 0.73 | 0.28 | 0.75 | 0.54 |
| m_vidi_tg | 0.26 | 0.23 | 0.80 | 0.10 | 0.04 | 0.90 | 0.09 | 0.98 | 0.09 | 0.54 | 0.72 | 0.27 | 0.76 | 0.54 |
| mufa | 0.44 | 0.23 | 0.78 | 0.11 | 0.05 | 0.91 | 0.12 | 0.78 | 0.09 | 0.54 | 0.71 | 0.29 | 0.76 | 0.54 |
| mufa_fa | 0.09 | 0.22 | 0.77 | 0.11 | 0.05 | 0.92 | 0.05 | 0.88 | 0.10 | 0.56 | 0.69 | 0.28 | 0.78 | 0.55 |
| pc | 0.95 | 0.23 | 0.79 | 0.11 | 0.04 | 0.95 | 0.14 | 0.62 | 0.08 | 0.56 | 0.71 | 0.28 | 0.77 | 0.49 |
| phe | 0.32 | 0.20 | 0.82 | 0.11 | 0.04 | 0.96 | 0.10 | 0.33 | 0.07 | 0.52 | 0.72 | 0.24 | 0.76 | 0.40 |
| pufa | 0.46 | 0.22 | 0.80 | 0.11 | 0.05 | 0.96 | 0.11 | 0.54 | 0.08 | 0.56 | 0.72 | 0.28 | 0.77 | 0.47 |
| pufa_fa | 0.03 | 0.22 | 0.76 | 0.11 | 0.04 | 0.86 | 0.03 | 0.82 | 0.10 | 0.55 | 0.69 | 0.27 | 0.78 | 0.55 |
| pyr | 0.92 | 0.23 | 0.78 | 0.10 | 0.05 | 0.92 | 0.14 | 0.07 | 0.07 | 0.58 | 0.66 | 0.25 | 0.95 | 0.20 |
| remnant_c | 0.30 | 0.22 | 0.83 | 0.11 | 0.05 | 0.99 | 0.10 | 0.36 | 0.08 | 0.53 | 0.73 | 0.28 | 0.74 | 0.42 |
| s_hdl_c | 0.41 | 0.24 | 0.81 | 0.10 | 0.04 | 0.96 | 0.11 | 0.59 | 0.10 | 0.52 | 0.69 | 0.26 | 0.73 | 0.49 |
| s_hdl_ce | 0.20 | 0.24 | 0.81 | 0.10 | 0.04 | 0.97 | 0.08 | 0.64 | 0.10 | 0.51 | 0.70 | 0.26 | 0.73 | 0.50 |
| s_hdi_fc | 0.02 | 0.23 | 0.80 | 0.11 | 0.03 | 0.99 | 0.02 | 0.68 | 0.08 | 0.56 | 0.68 | 0.29 | 0.79 | 0.49 |


|  | Clinical Progression |  |  |  |  |  |  | Metastases |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| metabolite | metabolite p -value | gleason p value | Stage pvalue | PS p-value | age $p$-value | centre $p$ - <br> value | global p- <br> value | metabolite $p$-value | gleason $p$ - <br> value | Stage pvalue | PS p-value | age $p$-value | centre $p$ - <br> value | global p- <br> value |
| s_hdl_1 | 0.06 | 0.23 | 0.79 | 0.11 | 0.03 | 0.95 | 0.04 | 0.57 | 0.09 | 0.56 | 0.66 | 0.29 | 0.78 | 0.48 |
| s_hdl_p | 0.05 | 0.23 | 0.79 | 0.11 | 0.03 | 0.95 | 0.03 | 0.55 | 0.09 | 0.56 | 0.67 | 0.30 | 0.78 | 0.48 |
| s_hdi_pl | 0.00 | 0.25 | 0.78 | 0.10 | 0.03 | 0.96 | 0.00 | 0.85 | 0.08 | 0.55 | 0.70 | 0.29 | 0.78 | 0.50 |
| s_hdl_tg | 0.74 | 0.23 | 0.79 | 0.11 | 0.05 | 0.92 | 0.14 | 0.49 | 0.09 | 0.52 | 0.71 | 0.28 | 0.74 | 0.49 |
| s_ldi_c | 0.08 | 0.22 | 0.83 | 0.11 | 0.05 | 0.98 | 0.05 | 0.53 | 0.09 | 0.55 | 0.71 | 0.28 | 0.74 | 0.46 |
| s_ldi_ce | 0.07 | 0.22 | 0.83 | 0.11 | 0.05 | 0.98 | 0.04 | 0.54 | 0.09 | 0.55 | 0.72 | 0.28 | 0.74 | 0.47 |
| s_IdI_fc | 0.13 | 0.22 | 0.82 | 0.10 | 0.05 | 0.97 | 0.06 | 0.50 | 0.09 | 0.56 | 0.70 | 0.27 | 0.73 | 0.46 |
| s_ldi_\| | 0.14 | 0.22 | 0.83 | 0.11 | 0.05 | 0.98 | 0.06 | 0.49 | 0.09 | 0.55 | 0.71 | 0.28 | 0.74 | 0.45 |
| s_ldi_p | 0.15 | 0.22 | 0.83 | 0.11 | 0.05 | 0.98 | 0.06 | 0.49 | 0.09 | 0.55 | 0.71 | 0.28 | 0.74 | 0.45 |
| s_\|di_pl | 0.58 | 0.22 | 0.80 | 0.11 | 0.05 | 0.96 | 0.12 | 0.60 | 0.09 | 0.56 | 0.70 | 0.28 | 0.75 | 0.48 |
| s_ldi_tg | 0.79 | 0.22 | 0.81 | 0.11 | 0.05 | 0.95 | 0.14 | 0.33 | 0.08 | 0.54 | 0.71 | 0.28 | 0.76 | 0.40 |
| s_vidil_c | 0.12 | 0.22 | 0.85 | 0.11 | 0.05 | 0.99 | 0.06 | 0.14 | 0.08 | 0.52 | 0.73 | 0.28 | 0.74 | 0.29 |
| s_vidl_ce | 0.05 | 0.21 | 0.86 | 0.11 | 0.05 | 0.99 | 0.03 | 0.14 | 0.07 | 0.52 | 0.73 | 0.28 | 0.74 | 0.29 |
| s_vidl_fc | 0.47 | 0.23 | 0.82 | 0.11 | 0.05 | 0.96 | 0.12 | 0.19 | 0.09 | 0.52 | 0.72 | 0.29 | 0.74 | 0.34 |
| s_vidl_I | 0.69 | 0.23 | 0.81 | 0.11 | 0.05 | 0.96 | 0.14 | 0.28 | 0.09 | 0.53 | 0.72 | 0.29 | 0.74 | 0.41 |
| s_vidl_p | 0.80 | 0.23 | 0.80 | 0.11 | 0.05 | 0.95 | 0.15 | 0.33 | 0.09 | 0.53 | 0.72 | 0.29 | 0.74 | 0.43 |
| s_vidl_pl | 0.69 | 0.23 | 0.81 | 0.11 | 0.05 | 0.95 | 0.14 | 0.21 | 0.09 | 0.52 | 0.72 | 0.29 | 0.74 | 0.36 |
| s_vidl_tg | 0.73 | 0.23 | 0.79 | 0.11 | 0.05 | 0.93 | 0.14 | 0.51 | 0.10 | 0.53 | 0.72 | 0.28 | 0.75 | 0.49 |
| serum_c | 0.26 | 0.21 | 0.82 | 0.11 | 0.05 | 0.96 | 0.09 | 0.43 | 0.09 | 0.56 | 0.72 | 0.28 | 0.75 | 0.44 |
| serum_tg | 0.45 | 0.23 | 0.79 | 0.11 | 0.05 | 0.91 | 0.12 | 0.80 | 0.10 | 0.54 | 0.72 | 0.29 | 0.74 | 0.54 |
| sfa | 0.58 | 0.23 | 0.78 | 0.11 | 0.05 | 0.92 | 0.13 | 0.79 | 0.09 | 0.54 | 0.73 | 0.28 | 0.76 | 0.53 |
| sfa_fa | 0.12 | 0.23 | 0.77 | 0.10 | 0.05 | 0.86 | 0.06 | 0.76 | 0.09 | 0.55 | 0.70 | 0.25 | 0.79 | 0.49 |
| sm | 0.28 | 0.21 | 0.82 | 0.11 | 0.05 | 0.97 | 0.09 | 0.63 | 0.09 | 0.56 | 0.71 | 0.28 | 0.76 | 0.50 |
| tg_pg | 0.28 | 0.23 | 0.78 | 0.10 | 0.05 | 0.91 | 0.10 | 0.86 | 0.11 | 0.55 | 0.71 | 0.28 | 0.76 | 0.56 |
| totcho | 0.81 | 0.22 | 0.79 | 0.11 | 0.04 | 0.95 | 0.13 | 0.72 | 0.09 | 0.56 | 0.71 | 0.27 | 0.77 | 0.50 |


| metabolite | Clinical Progression |  |  |  |  |  |  | Metastases |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | metabolite p-value | gleason p value | Stage pvalue | PS p-value | age p-value | centre pvalue | global p- <br> value | metabolite $p$-value | $\begin{aligned} & \text { gleason } \mathrm{p} \text { - } \\ & \text { value } \\ & \hline \end{aligned}$ | Stage pvalue | PS p-value | age $p$-value | centre pvalue | global pvalue |
| totfa | 0.77 | 0.23 | 0.78 | 0.11 | 0.05 | 0.94 | 0.14 | 0.68 | 0.09 | 0.54 | 0.72 | 0.29 | 0.76 | 0.51 |
| totpg | 0.82 | 0.23 | 0.78 | 0.11 | 0.04 | 0.94 | 0.14 | 0.69 | 0.08 | 0.56 | 0.71 | 0.28 | 0.77 | 0.50 |
| tyr | 0.42 | 0.23 | 0.80 | 0.11 | 0.05 | 0.94 | 0.11 | 0.79 | 0.09 | 0.56 | 0.70 | 0.28 | 0.78 | 0.50 |
| unsat | 0.04 | 0.22 | 0.76 | 0.10 | 0.04 | 0.88 | 0.03 | 0.89 | 0.10 | 0.57 | 0.72 | 0.27 | 0.78 | 0.55 |
| val | 0.93 | 0.22 | 0.78 | 0.11 | 0.04 | 0.94 | 0.13 | 0.57 | 0.08 | 0.55 | 0.72 | 0.28 | 0.79 | 0.47 |
| vidl_c | 0.60 | 0.23 | 0.81 | 0.11 | 0.05 | 0.97 | 0.14 | 0.38 | 0.09 | 0.53 | 0.73 | 0.29 | 0.73 | 0.45 |
| vidl_d | 0.05 | 0.21 | 0.79 | 0.10 | 0.04 | 0.86 | 0.04 | 0.88 | 0.10 | 0.55 | 0.72 | 0.27 | 0.77 | 0.55 |
| vidl_tg | 0.33 | 0.23 | 0.78 | 0.11 | 0.05 | 0.89 | 0.11 | 0.95 | 0.10 | 0.54 | 0.72 | 0.28 | 0.75 | 0.56 |
| xl_hdl_c | 0.93 | 0.23 | 0.76 | 0.10 | 0.05 | 0.93 | 0.13 | 0.87 | 0.09 | 0.55 | 0.69 | 0.28 | 0.78 | 0.51 |
| xl_hdl_ce | 0.87 | 0.22 | 0.76 | 0.10 | 0.05 | 0.93 | 0.13 | 0.90 | 0.09 | 0.56 | 0.69 | 0.28 | 0.78 | 0.51 |
| xl_hdl_fc | 0.93 | 0.23 | 0.76 | 0.10 | 0.05 | 0.93 | 0.14 | 0.80 | 0.09 | 0.55 | 0.68 | 0.28 | 0.78 | 0.51 |
| xl_hdl_1 | 0.95 | 0.23 | 0.76 | 0.10 | 0.04 | 0.93 | 0.14 | 0.92 | 0.09 | 0.56 | 0.68 | 0.27 | 0.78 | 0.52 |
| xl_hdl_p | 0.95 | 0.22 | 0.76 | 0.10 | 0.04 | 0.93 | 0.14 | 0.93 | 0.09 | 0.56 | 0.69 | 0.27 | 0.78 | 0.52 |
| xl_hdl_pl | 0.94 | 0.23 | 0.77 | 0.11 | 0.04 | 0.92 | 0.14 | 0.95 | 0.09 | 0.57 | 0.69 | 0.27 | 0.78 | 0.53 |
| xl_hdl_tg | 0.71 | 0.23 | 0.80 | 0.11 | 0.04 | 0.94 | 0.13 | 0.99 | 0.09 | 0.52 | 0.72 | 0.28 | 0.76 | 0.53 |
| xl_vidi_c | 0.36 | 0.23 | 0.80 | 0.11 | 0.04 | 0.90 | 0.11 | 0.98 | 0.10 | 0.53 | 0.74 | 0.28 | 0.77 | 0.55 |
| xl_vill_ce | 0.43 | 0.23 | 0.80 | 0.11 | 0.04 | 0.91 | 0.12 | 0.96 | 0.10 | 0.53 | 0.74 | 0.28 | 0.77 | 0.55 |
| xl_vidl_fc | 0.27 | 0.23 | 0.79 | 0.11 | 0.04 | 0.90 | 0.10 | 0.87 | 0.10 | 0.54 | 0.74 | 0.27 | 0.77 | 0.55 |
| xl_vidl_\| | 0.18 | 0.23 | 0.79 | 0.10 | 0.04 | 0.89 | 0.08 | 0.80 | 0.10 | 0.54 | 0.74 | 0.27 | 0.77 | 0.54 |
| xl_vidi_p | 0.18 | 0.23 | 0.79 | 0.11 | 0.04 | 0.88 | 0.08 | 0.82 | 0.10 | 0.53 | 0.74 | 0.27 | 0.78 | 0.54 |
| xl_vidl_pl | 0.19 | 0.23 | 0.78 | 0.10 | 0.04 | 0.91 | 0.08 | 0.82 | 0.10 | 0.54 | 0.74 | 0.27 | 0.77 | 0.54 |
| xl_vill_tg | 0.13 | 0.23 | 0.79 | 0.10 | 0.04 | 0.89 | 0.07 | 0.74 | 0.10 | 0.54 | 0.73 | 0.26 | 0.78 | 0.53 |
| xs_vidl_c | 0.03 | 0.18 | 0.90 | 0.11 | 0.05 | 0.99 | 0.02 | 0.17 | 0.07 | 0.52 | 0.74 | 0.27 | 0.74 | 0.31 |
| xs_vldi_ce | 0.04 | 0.18 | 0.89 | 0.11 | 0.05 | 0.99 | 0.03 | 0.24 | 0.07 | 0.52 | 0.74 | 0.28 | 0.74 | 0.36 |
| xs_vidl_fc | 0.01 | 0.19 | 0.89 | 0.11 | 0.04 | 0.95 | 0.01 | 0.10 | 0.07 | 0.52 | 0.74 | 0.26 | 0.78 | 0.24 |


|  | Clinical Progression |  |  |  |  |  |  | Metastases |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| metabolite | metabolite p-value | $\begin{aligned} & \text { gleason p- } \\ & \text { value } \\ & \hline \end{aligned}$ | Stage pvalue | PS p-value | age p-value | centre pvalue | global p- <br> value | metabolite p-value | ```gleason p- value``` | Stage pvalue | PS p-value | age $p$-value | centre p value | global pvalue |
| xs_vldi_1 | 0.05 | 0.19 | 0.88 | 0.11 | 0.05 | 0.99 | 0.03 | 0.14 | 0.07 | 0.51 | 0.74 | 0.27 | 0.75 | 0.28 |
| xs_vidl_p | 0.06 | 0.20 | 0.88 | 0.11 | 0.05 | 0.99 | 0.04 | 0.14 | 0.07 | 0.51 | 0.74 | 0.27 | 0.75 | 0.28 |
| xs_vldi_pl | 0.06 | 0.20 | 0.86 | 0.11 | 0.05 | 0.98 | 0.04 | 0.21 | 0.08 | 0.54 | 0.75 | 0.27 | 0.75 | 0.33 |
| xs_vidl_tg | 0.47 | 0.22 | 0.83 | 0.11 | 0.05 | 0.96 | 0.12 | 0.20 | 0.08 | 0.51 | 0.70 | 0.27 | 0.76 | 0.34 |
| xxi_vidl_c | 0.34 | 0.24 | 0.78 | 0.11 | 0.04 | 0.92 | 0.11 | 0.94 | 0.10 | 0.55 | 0.74 | 0.27 | 0.76 | 0.55 |
| xxi_vidl_ce | 0.65 | 0.24 | 0.80 | 0.11 | 0.05 | 0.92 | 0.14 | 0.87 | 0.10 | 0.53 | 0.74 | 0.28 | 0.76 | 0.55 |
| xxi_vidl_fc | 0.16 | 0.23 | 0.77 | 0.10 | 0.04 | 0.91 | 0.07 | 0.78 | 0.10 | 0.55 | 0.74 | 0.26 | 0.77 | 0.54 |
| xxi_vidil_ | 0.18 | 0.24 | 0.73 | 0.11 | 0.03 | 0.89 | 0.08 | 0.53 | 0.11 | 0.76 | 0.89 | 0.20 | 0.73 | 0.50 |
| xxl_vidl_p | 0.19 | 0.23 | 0.79 | 0.11 | 0.04 | 0.86 | 0.08 | 0.72 | 0.10 | 0.54 | 0.75 | 0.28 | 0.79 | 0.55 |
| xxi_vidl_pl | 0.18 | 0.24 | 0.73 | 0.11 | 0.03 | 0.90 | 0.08 | 0.51 | 0.11 | 0.76 | 0.89 | 0.20 | 0.73 | 0.49 |
| xxi_vidl_tg | 0.15 | 0.24 | 0.73 | 0.11 | 0.03 | 0.88 | 0.07 | 0.56 | 0.11 | 0.76 | 0.88 | 0.20 | 0.73 | 0.51 |

Appendix B Table B 4: F-statistic and r-square measures for the instruments of individual metabolites in UKBB

| Metabolite | Total | Number of <br> SNPs | Sample <br> size |
| :--- | :--- | :--- | :--- | :--- |
| r2 | statistic |  |  |


| Metabolite | Total r2 | Number of SNPS | Sample Size | F statistic |
| :---: | :---: | :---: | :---: | :---: |
| Triglycerides in IDL | 0.076 | 44 | 115078 | 214.7 |
| Triglycerides to total lipids ratio in IDL | 0.064 | 57 | 115078 | 137.9 |
| Isoleucine | 0.007 | 7 | 115075 | 114.0 |
| Cholesterol in large HDL | 0.075 | 73 | 115078 | 128.5 |
| Cholesterol to total lipids ratio in large HDL | 0.045 | 58 | 115078 | 92.6 |
| Cholesteryl esters in large HDL | 0.074 | 73 | 115078 | 126.4 |
| Cholesteryl esters to total lipids ratio in large HDL | 0.035 | 49 | 115078 | 86.0 |
| Free cholesterol in large HDL | 0.076 | 69 | 115078 | 136.9 |
| Free cholesterol to total lipids ratio in large HDL | 0.075 | 53 | 115078 | 176.2 |
| Total lipids in large HDL | 0.081 | 76 | 115078 | 132.9 |
| Concentration of large HDL particles | 0.074 | 68 | 115078 | 135.7 |
| Phospholipids in large HDL | 0.075 | 63 | 115078 | 147.9 |
| Phospholipids to total lipids ratio in large HDL | 0.048 | 56 | 115078 | 103.0 |
| Triglycerides in large HDL | 0.090 | 38 | 115078 | 299.2 |
| Triglycerides to total lipids ratio in large HDL | 0.047 | 57 | 115078 | 99.1 |
| Cholesterol in large LDL | 0.039 | 37 | 115078 | 127.0 |
| Cholesterol to total lipids ratio in large LDL | 0.033 | 43 | 115078 | 91.1 |
| Cholesteryl esters in large LDL | 0.038 | 35 | 115078 | 131.2 |
| Cholesteryl esters to total lipids ratio in large LDL | 0.019 | 26 | 115078 | 85.0 |
| Free cholesterol in large LDL | 0.042 | 43 | 115078 | 117.4 |
| Free cholesterol to total lipids ratio in large LDL | 0.047 | 59 | 115078 | 96.8 |
| Total lipids in large LDL | 0.039 | 35 | 115078 | 135.2 |
| Concentration of large LDL particles | 0.045 | 42 | 115078 | 129.5 |
| Phospholipids in large LDL | 0.038 | 35 | 115078 | 128.2 |
| Phospholipids to total lipids ratio in large LDL | 0.057 | 33 | 115078 | 211.9 |
| Triglycerides in large LDL | 0.066 | 48 | 115078 | 169.8 |
| Triglycerides to total lipids ratio in large LDL | 0.048 | 48 | 115078 | 120.3 |
| Cholesterol in large VLDL | 0.043 | 39 | 115078 | 133.1 |
| Cholesterol to total lipids ratio in large VLDL | 0.056 | 42 | 115067 | 162.2 |
| Cholesteryl esters in large VLDL | 0.039 | 36 | 115078 | 129.1 |
| Cholesteryl esters to total lipids ratio in large VLDL | 0.056 | 46 | 115067 | 147.8 |
| Free cholesterol in large VLDL | 0.042 | 40 | 115078 | 124.6 |
| Free cholesterol to total lipids ratio in large VLDL | 0.033 | 37 | 115067 | 105.9 |
| Total lipids in large VLDL | 0.040 | 41 | 115078 | 118.0 |
| Concentration of large VLDL particles | 0.042 | 41 | 115078 | 122.6 |
| Phospholipids in large VLDL | 0.043 | 39 | 115078 | 131.1 |
| Phospholipids to total lipids ratio in large VLDL | 0.040 | 45 | 115067 | 107.7 |
| Triglycerides in large VLDL | 0.040 | 43 | 115078 | 111.5 |
| Triglycerides to total lipids ratio in large VLDL | 0.059 | 42 | 115067 | 170.5 |
| Linoleic acid | 0.039 | 36 | 114999 | 128.6 |
| Ratio of linoleic acid to total fatty acids | 0.013 | 24 | 114999 | 63.8 |
| Lactate | 0.003 | 6 | 114802 | 51.9 |


| Metabolite | Total r2 | Number of SNPS | Sample Size | $F$ <br> statistic |
| :---: | :---: | :---: | :---: | :---: |
| LDL cholesterol | 0.038 | 36 | 115078 | 126.1 |
| Cholesteryl esters in LDL | 0.036 | 33 | 115078 | 129.6 |
| Free cholesterol in LDL | 0.041 | 39 | 115078 | 124.7 |
| Total lipids in LDL | 0.038 | 35 | 115078 | 131.4 |
| Concentration of LDL particles | 0.042 | 40 | 115078 | 127.6 |
| Phospholipids in LDL | 0.040 | 36 | 115078 | 132.0 |
| Average diameter for LDL particles | 0.022 | 27 | 115078 | 95.8 |
| Triglycerides in LDL | 0.067 | 53 | 115078 | 156.3 |
| Leucine | 0.010 | 11 | 115074 | 101.3 |
| Cholesterol in medium HDL | 0.049 | 65 | 115078 | 91.1 |
| Cholesterol to total lipids ratio in medium HDL | 0.037 | 58 | 115078 | 75.5 |
| Cholesteryl esters in medium HDL | 0.048 | 66 | 115078 | 87.6 |
| Cholesteryl esters to total lipids ratio in medium HDL | 0.040 | 50 | 115078 | 95.5 |
| Free cholesterol in medium HDL | 0.053 | 58 | 115078 | 111.1 |
| Free cholesterol to total lipids ratio in medium HDL | 0.061 | 61 | 115078 | 122.3 |
| Total lipids in medium HDL | 0.043 | 49 | 115078 | 106.6 |
| Concentration of medium HDL particles | 0.045 | 51 | 115078 | 105.0 |
| Phospholipids in medium HDL | 0.045 | 47 | 115078 | 114.4 |
| Phospholipids to total lipids ratio in medium HDL | 0.033 | 55 | 115078 | 71.1 |
| Triglycerides in medium HDL | 0.049 | 41 | 115078 | 145.7 |
| Triglycerides to total lipids ratio in medium HDL | 0.048 | 50 | 115078 | 115.1 |
| Cholesterol in medium LDL | 0.036 | 35 | 115078 | 122.4 |
| Cholesterol to total lipids ratio in medium LDL | 0.026 | 28 | 115078 | 109.5 |
| Cholesteryl esters in medium LDL | 0.036 | 37 | 115078 | 115.2 |
| Cholesteryl esters to total lipids ratio in medium LDL | 0.036 | 32 | 115078 | 135.4 |
| Free cholesterol in medium LDL | 0.039 | 38 | 115078 | 123.1 |
| Free cholesterol to total lipids ratio in medium LDL | 0.052 | 59 | 115078 | 107.8 |
| Total lipids in medium LDL | 0.037 | 34 | 115078 | 128.2 |
| Concentration of medium LDL particles | 0.035 | 35 | 115078 | 120.3 |
| Phospholipids in medium LDL | 0.038 | 34 | 115078 | 133.5 |
| Phospholipids to total lipids ratio in medium LDL | 0.047 | 30 | 115078 | 188.6 |
| Triglycerides in medium LDL | 0.062 | 51 | 115078 | 148.7 |
| Triglycerides to total lipids ratio in medium LDL | 0.066 | 49 | 115078 | 165.4 |
| Cholesterol in medium VLDL | 0.040 | 38 | 115078 | 126.2 |
| Cholesterol to total lipids ratio in medium VLDL | 0.045 | 56 | 115078 | 97.5 |
| Cholesteryl esters in medium VLDL | 0.039 | 38 | 115078 | 124.2 |
| Cholesteryl esters to total lipids ratio in medium VLDL | 0.043 | 54 | 115078 | 95.2 |
| Free cholesterol in medium VLDL | 0.040 | 41 | 115078 | 118.4 |
| Free cholesterol to total lipids ratio in medium VLDL | 0.045 | 45 | 115078 | 120.9 |
| Total lipids in medium VLDL | 0.035 | 36 | 115078 | 116.8 |
| Concentration of medium VLDL particles | 0.040 | 43 | 115078 | 110.9 |
| Phospholipids in medium VLDL | 0.043 | 45 | 115078 | 113.5 |


| Metabolite | Total <br> r2 | Number of SNPS | Sample Size | F statistic |
| :---: | :---: | :---: | :---: | :---: |
| Phospholipids to total lipids ratio in medium VLDL | 0.053 | 50 | 115078 | 129.4 |
| Triglycerides in medium VLDL | 0.046 | 46 | 115078 | 121.2 |
| Triglycerides to total lipids ratio in medium VLDL | 0.043 | 49 | 115078 | 104.7 |
| Monounsaturated fatty acids | 0.043 | 47 | 114999 | 109.0 |
| Ratio of monounsaturated fatty acids to total fatty acids | 0.052 | 44 | 114999 | 143.1 |
| Total cholesterol minus HDL-C | 0.038 | 35 | 115078 | 129.9 |
| Omega-3 fatty acids | 0.091 | 37 | 114999 | 309.4 |
| Ratio of omega-3 fatty acids to total fatty acids | 0.087 | 26 | 114999 | 423.0 |
| Omega-6 fatty acids | 0.044 | 42 | 114999 | 125.2 |
| Ratio of omega-6 fatty acids to omega-3 fatty acids | 0.084 | 27 | 114999 | 391.8 |
| Ratio of omega-6 fatty acids to total fatty acids | 0.028 | 36 | 114999 | 91.8 |
| Phenylalanine | 0.009 | 5 | 115025 | 211.8 |
| Phosphatidylcholines | 0.057 | 43 | 114999 | 160.7 |
| Phosphoglycerides | 0.053 | 43 | 114999 | 150.6 |
| Polyunsaturated fatty acids | 0.055 | 44 | 114999 | 151.6 |
| Ratio of polyunsaturated fatty acids to monounsaturated fatty acids | 0.044 | 38 | 114999 | 138.1 |
| Ratio of polyunsaturated fatty acids to total fatty acids | 0.031 | 35 | 114999 | 103.9 |
| Pyruvate | 0.014 | 15 | 114748 | 109.0 |
| Remnant cholesterol (non-HDL non-LDL -cholesterol) | 0.042 | 40 | 115078 | 125.5 |
| Cholesterol in small HDL | 0.038 | 36 | 115078 | 126.5 |
| Cholesterol to total lipids ratio in small HDL | 0.046 | 42 | 115078 | 130.8 |
| Cholesteryl esters in small HDL | 0.045 | 34 | 115078 | 157.8 |
| Cholesteryl esters to total lipids ratio in small HDL | 0.062 | 43 | 115078 | 177.1 |
| Free cholesterol in small HDL | 0.038 | 37 | 115078 | 122.2 |
| Free cholesterol to total lipids ratio in small HDL | 0.084 | 55 | 115078 | 192.0 |
| Total lipids in small HDL | 0.042 | 37 | 115078 | 136.9 |
| Concentration of small HDL particles | 0.036 | 31 | 115078 | 138.3 |
| Phospholipids in small HDL | 0.042 | 43 | 115078 | 117.8 |
| Phospholipids to total lipids ratio in small HDL | 0.050 | 44 | 115078 | 136.7 |
| Triglycerides in small HDL | 0.045 | 48 | 115078 | 113.8 |
| Triglycerides to total lipids ratio in small HDL | 0.052 | 61 | 115078 | 104.4 |
| Cholesterol in small LDL | 0.039 | 38 | 115078 | 124.4 |
| Cholesterol to total lipids ratio in small LDL | 0.025 | 27 | 115078 | 108.1 |
| Cholesteryl esters in small LDL | 0.038 | 36 | 115078 | 126.2 |
| Cholesteryl esters to total lipids ratio in small LDL | 0.041 | 34 | 115078 | 145.1 |
| Free cholesterol in small LDL | 0.045 | 41 | 115078 | 132.0 |
| Free cholesterol to total lipids ratio in small LDL | 0.050 | 53 | 115078 | 113.6 |
| Total lipids in small LDL | 0.041 | 37 | 115078 | 134.1 |
| Concentration of small LDL particles | 0.042 | 40 | 115078 | 125.1 |
| Phospholipids in small LDL | 0.047 | 36 | 115078 | 156.5 |
| Phospholipids to total lipids ratio in small LDL | 0.040 | 36 | 115078 | 133.2 |
| Triglycerides in small LDL | 0.055 | 48 | 115078 | 139.9 |


| Metabolite | Total <br> r2 | Number of SNPs | Sample <br> Size | F statistic |
| :---: | :---: | :---: | :---: | :---: |
| Free cholesterol to total lipids ratio in very large HDL | 0.058 | 52 | 115053 | 135.7 |
| Total lipids in very large HDL | 0.071 | 55 | 115078 | 158.7 |
| Concentration of very large HDL particles | 0.081 | 67 | 115078 | 150.5 |
| Phospholipids in very large HDL | 0.072 | 53 | 115078 | 168.0 |
| Phospholipids to total lipids ratio in very large HDL | 0.057 | 41 | 115053 | 168.4 |
| Triglycerides in very large HDL | 0.079 | 46 | 115078 | 214.8 |
| Triglycerides to total lipids ratio in very large HDL | 0.054 | 52 | 115053 | 126.4 |
| Cholesterol in very large VLDL | 0.040 | 37 | 115078 | 128.8 |
| Cholesterol to total lipids ratio in very large VLDL | 0.047 | 51 | 114155 | 110.3 |
| Cholesteryl esters in very large VLDL | 0.042 | 42 | 115078 | 119.6 |
| Cholesteryl esters to total lipids ratio in very large VLDL | 0.049 | 54 | 114155 | 108.7 |
| Free cholesterol in very large VLDL | 0.039 | 34 | 115078 | 136.7 |
| Free cholesterol to total lipids ratio in very large VLDL | 0.045 | 44 | 114155 | 121.8 |
| Total lipids in very large VLDL | 0.042 | 41 | 115078 | 122.2 |
| Concentration of very large VLDL particles | 0.046 | 43 | 115078 | 130.1 |
| Phospholipids in very large VLDL | 0.042 | 39 | 115078 | 128.9 |
| Phospholipids to total lipids ratio in very large VLDL | 0.027 | 36 | 114155 | 88.0 |
| Triglycerides in very large VLDL | 0.044 | 49 | 115078 | 108.4 |
| Triglycerides to total lipids ratio in very large VLDL | 0.048 | 42 | 114155 | 136.1 |
| Cholesterol in very small VLDL | 0.062 | 44 | 115078 | 173.7 |
| Cholesterol to total lipids ratio in very small VLDL | 0.048 | 51 | 115078 | 114.6 |
| Cholesteryl esters in very small VLDL | 0.060 | 44 | 115078 | 167.3 |
| Cholesteryl esters to total lipids ratio in very small VLDL | 0.049 | 53 | 115078 | 111.4 |
| Free cholesterol in very small VLDL | 0.068 | 45 | 115078 | 185.6 |
| Free cholesterol to total lipids ratio in very small VLDL | 0.039 | 36 | 115078 | 128.6 |
| Total lipids in very small VLDL | 0.072 | 50 | 115078 | 177.9 |
| Concentration of very small VLDL particles | 0.066 | 46 | 115078 | 176.0 |
| Phospholipids in very small VLDL | 0.075 | 51 | 115078 | 182.0 |
| Phospholipids to total lipids ratio in very small VLDL | 0.032 | 47 | 115078 | 81.4 |
| Triglycerides in very small VLDL | 0.069 | 58 | 115078 | 147.4 |
| Triglycerides to total lipids ratio in very small VLDL | 0.049 | 53 | 115078 | 112.7 |
| Cholesterol in chylomicrons and extremely large VLDL | 0.047 | 44 | 115078 | 128.5 |
| Cholesterol to total lipids ratio in chylomicrons and extremely large VLDL | 0.029 | 19 | 111631 | 178.5 |
| Cholesteryl esters in chylomicrons and extremely large VLDL | 0.049 | 44 | 115078 | 136.0 |
| Cholesteryl esters to total lipids ratio in chylomicrons and extremely large |  |  |  |  |
| VLDL | 0.025 | 17 | 111631 | 169.9 |
| Free cholesterol in chylomicrons and extremely large VLDL | 0.045 | 45 | 115078 | 119.4 |
| Free cholesterol to total lipids ratio in chylomicrons and extremely large VLDL | 0.030 | 20 | 111631 | 170.8 |
| Total lipids in chylomicrons and extremely large VLDL | 0.046 | 45 | 115078 | 122.7 |
| Concentration of chylomicrons and extremely large VLDL particles | 0.043 | 42 | 115078 | 123.2 |
| Phospholipids in chylomicrons and extremely large VLDL | 0.045 | 45 | 115078 | 121.4 |
| Triglycerides in chylomicrons and extremely large VLDL | 0.047 | 42 | 115078 | 133.9 |
| Triglycerides to total lipids ratio in chylomicrons and extremely large VLDL | 0.029 | 13 | 111631 | 257.5 |


| Metabolite | $\begin{aligned} & \hline \text { Total } \\ & \text { r2 } \\ & \hline \end{aligned}$ | Number of SNPS | Sample Size | statistic |
| :---: | :---: | :---: | :---: | :---: |
| Triglycerides to total lipids ratio in small LDL | 0.054 | 51 | 115078 | 129.6 |
| Cholesterol in small VLDL | 0.049 | 45 | 115078 | 132.1 |
| Cholesterol to total lipids ratio in small VLDL | 0.045 | 41 | 115078 | 132.5 |
| Cholesteryl esters in small VLDL | 0.053 | 49 | 115078 | 132.5 |
| Cholesteryl esters to total lipids ratio in small VLDL | 0.055 | 47 | 115078 | 143.7 |
| Free cholesterol in small VLDL | 0.043 | 40 | 115078 | 128.2 |
| Free cholesterol to total lipids ratio in small VLDL | 0.048 | 52 | 115078 | 112.2 |
| Total lipids in small VLDL | 0.054 | 49 | 115078 | 133.1 |
| Concentration of small VLDL particles | 0.056 | 48 | 115078 | 141.8 |
| Phospholipids in small VLDL | 0.049 | 51 | 115078 | 116.5 |
| Phospholipids to total lipids ratio in small VLDL | 0.047 | 51 | 115078 | 112.4 |
| Triglycerides in small VLDL | 0.055 | 58 | 115078 | 114.8 |
| Triglycerides to total lipids ratio in small VLDL | 0.043 | 44 | 115078 | 116.8 |
| Saturated fatty acids | 0.038 | 40 | 114999 | 114.7 |
| Ratio of saturated fatty acids to total fatty acids | 0.013 | 20 | 114999 | 77.2 |
| Sphingomyelins | 0.048 | 47 | 114999 | 123.9 |
| Ratio of triglycerides to phosphoglycerides | 0.052 | 60 | 114999 | 105.9 |
| Total concentration of branched-chain amino acids (leucine + isoleucine + valine) | 0.011 | 11 | 115047 | 112.8 |
| Total cholesterol | 0.043 | 45 | 115078 | 113.6 |
| Total esterified cholesterol | 0.042 | 47 | 115078 | 108.3 |
| Total fatty acids | 0.043 | 41 | 114999 | 125.6 |
| Total free cholesterol | 0.042 | 41 | 115078 | 124.1 |
| Total lipids in lipoprotein particles | 0.043 | 47 | 115078 | 110.9 |
| Total concentration of lipoprotein particles | 0.036 | 46 | 115078 | 93.2 |
| Total phospholipids in lipoprotein particles | 0.050 | 46 | 115078 | 130.8 |
| Total triglycerides | 0.043 | 47 | 115078 | 109.7 |
| Tyrosine | 0.018 | 23 | 114911 | 91.3 |
| Degree of unsaturation | 0.077 | 32 | 114999 | 301.2 |
| Valine | 0.013 | 13 | 115048 | 114.6 |
| VLDL cholesterol | 0.044 | 42 | 115078 | 126.0 |
| Cholesteryl esters in VLDL | 0.045 | 39 | 115078 | 139.8 |
| Free cholesterol in VLDL | 0.047 | 45 | 115078 | 125.1 |
| Total lipids in VLDL | 0.041 | 38 | 115078 | 128.7 |
| Concentration of VLDL particles | 0.050 | 45 | 115078 | 135.0 |
| Phospholipids in VLDL | 0.047 | 44 | 115078 | 127.7 |
| Average diameter for VLDL particles | 0.066 | 53 | 115078 | 152.6 |
| Triglycerides in VLDL | 0.042 | 44 | 115078 | 114.0 |
| Cholesterol in very large HDL | 0.071 | 61 | 115078 | 145.0 |
| Cholesterol to total lipids ratio in very large HDL | 0.070 | 37 | 115053 | 234.4 |
| Cholesteryl esters in very large HDL | 0.072 | 65 | 115078 | 136.7 |
| Cholesteryl esters to total lipids ratio in very large HDL | 0.036 | 29 | 115053 | 149.4 |
| Free cholesterol in very large HDL | 0.068 | 53 | 115078 | 159.5 |

Appendix B Table B 5: Mendelian Randomisation results for Prostate Cancer survival in the PRACTICAL consortium with instruments derived from UKBB

| Exposure (metabolic trait) | Method | Number of SNPs |  | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Acetate | MR Egger |  | 6 | 0.790 | 1.199 | 0.345 | 4.169 |
| Acetate | Weighted median |  | 6 | 0.927 | 1.026 | 0.596 | 1.765 |
| Acetate | Simple mode |  | 6 | 0.711 | 0.841 | §0.355 | 1.993 |
| Acetate | Weighted mode |  | 6 | 0.929 | 1.030 | 0.555 | 1.912 |
| Acetate | Inverse variance weighted |  | 6 | 0.807 | 1.056 | 0.683 | 1.634 |
| Acetoacetate | MR Egger |  | 4 | 0.641 | 0.632 | 0.121 | 3.295 |
| Acetoacetate | Weighted median |  | 4 | 0.606 | 1.201 | 0.599 | 2.405 |
| Acetoacetate | Simple mode |  | 4 | 0.649 | 1.254 | 0.520 | 3.027 |
| Acetoacetate | Weighted mode |  | 4 | 0.657 | 1.268 | 0.492 | 3.266 |
| Acetoacetate | Inverse variance weighted |  | 4 | 0.745 | 1.106 | 0.603 | 2.029 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Acetone | MR Egger | 9 | 0.162 | 2.761 | 0.773 | 9.860 |
| Acetone | Weighted median | 9 | 0.296 | 1.351 | 0.769 | 2.374 |
| Acetone | Simple mode | 9 | 0.368 | 1.485 | 0.659 | 3.343 |
| Acetone | Weighted mode | 9 | 0.381 | 1.436 | 0.668 | 3.089 |
| Acetone | Inverse variance weighted | 9 | 0.270 | 1.417 | 0.763 | 2.629 |
| Ala | MR Egger | 22 | 0.053 | 2.036 | 1.034 | 4.010 |
| Ala | Weighted median | 22 | 0.339 | 1.165 | 0.852 | 1.594 |
| Ala | Simple mode | 22 | 0.662 | 1.121 | 0.677 | 1.854 |
| Ala | Weighted mode | 22 | 0.513 | 1.138 | 0.777 | 1.666 |
| Ala | Inverse variance weighted | 22 | 0.057 | 1.240 | 0.994 | 1.547 |
| Albumin | MR Egger | 19 | 0.267 | 0.840 | 0.623 | 1.132 |
| Albumin | Weighted median | 19 | 0.156 | 0.836 | 0.652 | 1.071 |
| Albumin | Simple mode | 19 | 0.449 | 1.223 | 0.734 | 2.039 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Albumin | Inverse variance weighted | 19 | 0.250 | 0.892 | 0.734 | 1.084 |
| ApoA1 | MR Egger | 51 | 0.503 | 1.104 | 0.828 | 1.473 |
| ApoA1 | Weighted median | 51 | 0.839 | 1.019 | 0.849 | 1.223 |
| ApoA1 | Simple mode | 51 | 0.985 | 0.997 | 0.714 | 1.392 |
| ApoA1 | Weighted mode | 51 | 0.860 | 0.984 | 0.823 | 1.176 |
| ApoA1 | Inverse variance weighted | 51 | 0.821 | 1.019 | 0.866 | 1.200 |
| ApoB | MR Egger | 41 | 0.048 | 1.266 | 1.010 | 1.588 |
| ApoB | Weighted median | 41 | 0.112 | 1.160 | 0.966 | 1.393 |
| ApoB | Simple mode | 41 | 0.435 | 1.147 | 0.815 | 1.614 |
| ApoB | Weighted mode | 41 | 0.233 | 1.175 | 0.905 | 1.525 |
| ApoB | 41 | 0.044 | 1.138 | 1.003 | 1.291 |  |
| ApoB_by_ApoA1 | Mrverse variance weighted | 55 | 0.327 | 1.150 | 0.872 | 1.515 |
| ApoB_by_ApoA1 | Weighted median | 55 | 0.718 | 1.037 | 0.850 | 1.266 |
| ApoB_by_ApoA1 | Simple mode | 55 | 0.898 | 1.025 | 0.705 | 1.490 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| ApoB_by_ApoA1 | Weighted mode | 55 | 0.645 | 1.060 | 0.828 | 1.359 |
| ApoB_by_ApoA1 | Inverse variance weighted | 55 | 0.105 | 1.133 | 0.974 | 1.317 |
| bOHbutyrate | MR Egger | 10 | 0.322 | 2.013 | 0.550 | 7.366 |
| bOHbutyrate | Weighted median | 10 | 0.195 | 1.330 | 0.864 | 2.049 |
| bOHbutyrate | Simple mode | 10 | 0.464 | 1.283 | 0.677 | 2.430 |
| bOHbutyrate | Weighted mode | 10 | 0.380 | 1.365 | 0.705 | 2.640 |
| bOHbutyrate | Inverse variance weighted | 10 | 0.144 | 1.426 | 0.885 | 2.296 |
| Cholines | MR Egger | 46 | 0.930 | 1.009 | 0.823 | 1.237 |
| Cholines | Weighted median | 46 | 0.510 | 1.059 | 0.893 | 1.256 |
| Cholines | Simple mode | 46 | 0.724 | 1.055 | 0.785 | 1.419 |
| Cholines | Weighted mode | 46 | 0.486 | 1.063 | 0.897 | 1.259 |
| Cholines | Inverse variance weighted | 46 | 0.085 | 1.102 | 0.987 | 1.232 |
| Citrate | MR Egger | 25 | 0.278 | 0.829 | 0.595 | 1.154 |
| Citrate | Weighted median | 25 | 0.677 | 0.949 | 0.742 | 1.214 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Citrate | Simple mode | 25 | 0.182 | 0.757 | 0.509 | 1.126 |
| Citrate | Weighted mode | 25 | 0.588 | 0.928 | 0.712 | 1.211 |
| Citrate | Inverse variance weighted | 25 | 0.190 | 0.899 | 0.767 | 1.054 |
| Clinical_LDL_C | Inverse variance weighted | 31 | 0.007 | 1.209 | 1.053 | 1.388 |
| Clinical_LDL_C | MR Egger | 31 | 0.014 | 1.372 | 1.083 | 1.739 |
| Clinical_LDL_C | Weighted median | 31 | 0.003 | 1.336 | 1.102 | 1.621 |
| Clinical_LDL_C | Simple mode | 31 | 0.367 | 1.180 | 0.828 | 1.683 |
| Clinical_LDL_C | Weighted mode | 31 | 0.009 | 1.392 | 1.104 | 1.755 |
| Creatinine | MR Egger | 51 | 0.641 | 1.161 | 0.622 | 2.169 |
| Creatinine | Weighted median | 51 | 0.577 | 0.924 | 0.698 | 1.221 |
| Creatinine | Simple mode | 51 | 0.736 | 0.905 | 0.507 | 1.615 |
| Creatinine | Weighted mode | 51 | 0.872 | 0.966 | 0.632 | 1.475 |
| Creatinine | Inverse variance weighted | 51 | 0.463 | 0.929 | 0.763 | 1.131 |
| DHA | 29 | 0.225 | 1.088 | 0.952 | 1.243 |  |


| Exposure (metabolic <br> trait | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| DHA | Weighted median | 29 | 0.086 | 1.109 | 0.985 | 1.249 |
| DHA | Simple mode | 29 | 0.565 | 1.092 | 0.812 | 1.469 |
| DHA | Weighted mode | 29 | 0.083 | 1.107 | 0.991 | 1.236 |
| DHA | Inverse variance weighted | 29 | 0.038 | 1.107 | 1.006 | 1.219 |
| DHA_pct | MR Egger | 20 | 0.140 | 1.124 | 0.969 | 1.304 |
| DHA_pct | Weighted median | 20 | 0.079 | 1.115 | 0.987 | 1.259 |
| DHA_pct | Simple mode | 20 | 0.890 | 1.028 | 0.698 | 1.514 |
| DHA_pct | Weighted mode | 20 | 0.147 | 1.105 | 0.971 | 1.257 |
| DHA_pct | Inverse variance weighted | 20 | 0.274 | 1.064 | 0.952 | 1.189 |
| Gln | MR Egger | 30 | 0.572 | 0.926 | 0.712 | 1.205 |
| Gln | Weighted median | 30 | 0.639 | 1.038 | 0.887 | 1.215 |
| Gln | Simple mode | 30 | 0.873 | 1.024 | 0.767 | 1.368 |
| Gln | Weighted mode | 30 | 0.866 | 1.013 | 0.869 | 1.181 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Glucose | Odds ratio <br> upper 95\% CI |  |  |  |  |  |
| Glucose | MR Egger | 16 | 0.362 | 1.270 | 0.773 | 2.086 |
| Glucose | Weighted median | 16 | 0.731 | 1.054 | 0.781 | 1.422 |
| Glucose | Simple mode | 16 | 0.830 | 0.943 | 0.559 | 1.591 |
| Glucose | Weighted mode | 16 | 0.544 | 1.103 | 0.808 | 1.506 |
| Gly | Inverse variance weighted | 16 | 0.849 | 0.978 | 0.780 | 1.226 |
| Gly | MR Egger | 29 | 0.551 | 1.023 | 0.951 | 1.100 |
| Gly | Weighted median | 29 | 0.457 | 1.026 | 0.960 | 1.096 |
| Gly | Simple mode | 29 | 0.476 | 1.113 | 0.833 | 1.487 |
| Gly | Weighted mode | 29 | 0.434 | 1.028 | 0.960 | 1.100 |
| GlycA | Inverse variance weighted | 29 | 0.255 | 1.037 | 0.974 | 1.104 |
| GlycA | MR Egger | 41 | 0.149 | 1.169 | 0.949 | 1.440 |
| GlycA | Weighted median | 41 | 0.420 | 1.079 | 0.897 | 1.299 |
| GlycA | Simple mode | 41 | 0.562 | 1.111 | 0.781 | 1.582 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| GlycA | Inverse variance weighted | 41 | 0.833 | 1.013 | 0.896 | 1.146 |
| HDL_C | MR Egger | 69 | 0.794 | 1.035 | 0.800 | 1.340 |
| HDL_C | Weighted median | 69 | 0.870 | 1.014 | 0.855 | 1.203 |
| HDL_C | Simple mode | 69 | 0.870 | 0.974 | 0.711 | 1.334 |
| HDL_C | Weighted mode | 69 | 0.902 | 0.988 | 0.817 | 1.194 |
| HDL_C | Inverse variance weighted | 69 | 0.377 | 1.063 | 0.928 | 1.218 |
| HDL_CE | MR Egger | 67 | 0.692 | 1.054 | 0.812 | 1.369 |
| HDL_CE | Weighted median | 67 | 0.892 | 1.012 | 0.847 | 1.210 |
| HDL_CE | Simple mode | 67 | 0.704 | 0.939 | 0.681 | 1.296 |
| HDL_CE | Weighted mode | 67 | 0.957 | 0.994 | 0.809 | 1.222 |
| HDL_CE | Inverse variance weighted | 67 | 0.542 | 1.044 | 0.910 | 1.197 |
| HDL_FC | MR Egger | 62 | 0.850 | 1.020 | 0.829 | 1.256 |
| HDL_FC | Weighted median | 62 | 0.676 | 1.037 | 0.874 | 1.231 |
| HDL_FC | Simple mode | 0.643 | 0.931 | 0.690 | 1.257 |  |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| HDL_FC | Weighted mode | 62 | 0.818 | 1.021 | 0.855 | 1.219 |
| HDL_FC | Inverse variance weighted | 62 | 0.181 | 1.083 | 0.964 | 1.217 |
| HDL_L | MR Egger | 60 | 0.674 | 1.055 | 0.822 | 1.356 |
| HDL_L | Weighted median | 60 | 0.849 | 1.017 | 0.856 | 1.208 |
| HDL_L | Simple mode | 60 | 0.692 | 0.936 | 0.673 | 1.299 |
| HDL_L | Weighted mode | 60 | 0.777 | 0.974 | 0.812 | 1.168 |
| HDL_L | Inverse variance weighted | 60 | 0.743 | 1.023 | 0.892 | 1.174 |
| HDL_P | MR Egger | 47 | 0.685 | 1.076 | 0.756 | 1.531 |
| HDL_P | Weighted median | 47 | 0.835 | 1.021 | 0.837 | 1.246 |
| HDL_P | Simple mode | 0.951 | 0.989 | 0.690 | 1.417 |  |
| HDL_P | Weighted mode | 47 | 0.994 | 1.001 | 0.749 | 1.339 |
| HDL_P | Inverse variance weighted | 47 | 0.685 | 1.036 | 0.872 | 1.231 |
| HDL_PL | MR Egger | 52 | 0.992 | 0.999 | 0.787 | 1.267 |
| HDL_PL | Weighted median | 52 | 0.857 | 1.016 | 0.853 | 1.211 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| HDL_PL | Simple mode | 52 | 0.902 | 0.980 | 0.710 | 1.353 |
| HDL_PL | Weighted mode | 52 | 0.919 | 0.990 | 0.822 | 1.193 |
| HDL_PL | Inverse variance weighted | 52 | 0.823 | 1.016 | 0.886 | 1.165 |
| HDL_size | MR Egger | 66 | 0.924 | 0.991 | 0.831 | 1.183 |
| HDL_size | Weighted median | 66 | 0.740 | 1.025 | 0.888 | 1.183 |
| HDL_size | Simple mode | 66 | 0.877 | 0.974 | 0.703 | 1.350 |
| HDL_size | Weighted mode | 66 | 0.957 | 1.004 | 0.875 | 1.152 |
| HDL_size | Inverse variance weighted | 66 | 0.326 | 1.058 | 0.946 | 1.183 |
| HDL_TG | MR Egger | 43 | 0.712 | 1.032 | 0.875 | 1.217 |
| HDL_TG | Weighted median | 43 | 0.595 | 1.042 | 0.896 | 1.212 |
| HDL_TG | Simple mode | 43 | 0.702 | 1.053 | 0.808 | 1.373 |
| HDL_TG | Weighted mode | 43 | 0.543 | 1.043 | 0.911 | 1.195 |
| HDL_TG | Inverse variance weighted | 43 | 0.886 | 1.008 | 0.905 | 1.122 |
| His | 9 | 0.523 | 0.761 | 0.344 | 1.685 |  |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| His | Weighted median | 9 | 0.277 | 0.812 | 0.557 | 1.182 |
| His | Simple mode | 9 | 0.082 | 0.580 | 0.339 | 0.993 |
| His | Weighted mode | 9 | 0.230 | 0.764 | 0.509 | 1.147 |
| His | Inverse variance weighted | 9 | 0.047 | 0.740 | 0.550 | 0.996 |
| IDL_C | MR Egger | 51 | 0.003 | 1.399 | 1.133 | 1.726 |
| IDL_C | Weighted median | 51 | 0.165 | 1.131 | 0.950 | 1.346 |
| IDL_C | Simple mode | 0.784 | 1.051 | 0.738 | 1.498 |  |
| IDL_C | Weighted mode | 51 | 0.039 | 1.271 | 1.018 | 1.586 |
| IDL_C | Inverse variance weighted | 51 | 0.065 | 1.118 | 0.993 | 1.259 |
| IDL_C_pct | MR Egger | 54 | 0.282 | 1.106 | 0.922 | 1.327 |
| IDL_C_pct | Weighted median | 54 | 0.951 | 0.996 | 0.863 | 1.148 |
| IDL_C_pct | Simple mode | 54 | 0.858 | 0.975 | 0.739 | 1.286 |
| IDL_C_pct | Weighted mode | 54 | 0.945 | 1.005 | 0.872 | 1.158 |
| IDL_C_pct | Inverse variance weighted | 54 | 0.399 | 1.048 | 0.940 | 1.168 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| IDL_CE | MR Egger | 50 | 0.005 | 1.378 | 1.111 | 1.710 |
| IDL_CE | Weighted median | 50 | 0.151 | 1.141 | 0.953 | 1.366 |
| IDL_CE | Simple mode | 50 | 0.773 | 1.054 | 0.738 | 1.505 |
| IDL_CE | Weighted mode | 50 | 0.057 | 1.272 | 0.999 | 1.620 |
| IDL_CE | Inverse variance weighted | 50 | 0.055 | 1.124 | 0.997 | 1.267 |
| IDL_CE_pct | MR Egger | 41 | 0.362 | 1.127 | 0.874 | 1.455 |
| IDL_CE_pct | Weighted median | 41 | 0.524 | 1.062 | 0.882 | 1.280 |
| IDL_CE_pct | Simple mode | 41 | 0.641 | 0.920 | 0.650 | 1.302 |
| IDL_CE_pct | Weighted mode | 41 | 0.881 | 1.013 | 0.857 | 1.197 |
| IDL_CE_pct | Inverse variance weighted | 41 | 0.682 | 1.032 | 0.888 | 1.200 |
| IDL_FC | Inverse variance weighted | 49 | 0.023 | 1.145 | 1.018 | 1.288 |
| IDL_FC | MR Egger | 49 | 0.003 | 1.394 | 1.134 | 1.714 |
| IDL_FC | Weighted median | 49 | 0.027 | 1.212 | 1.022 | 1.436 |
| IDL_FC | Simple mode | 0.614 | 1.091 | 0.778 | 1.530 |  |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| IDL_FC | Weighted mode | 49 | 0.026 | 1.281 | 1.037 | 1.582 |
| IDL_FC_pct | MR Egger | 29 | 0.239 | 1.258 | 0.866 | 1.827 |
| IDL_FC_pct | Weighted median | 29 | 0.051 | 1.311 | 0.998 | 1.721 |
| IDL_FC_pct | Simple mode | 29 | 0.638 | 0.889 | 0.546 | 1.446 |
| IDL_FC_pct | Weighted mode | 29 | 0.153 | 1.250 | 0.928 | 1.685 |
| IDL_FC_pct | Inverse variance weighted | 29 | 0.119 | 1.151 | 0.964 | 1.374 |
| IDL_L | MR Egger | 47 | 0.007 | 1.330 | 1.090 | 1.622 |
| IDL_L | Weighted median | 47 | 0.318 | 1.092 | 0.919 | 1.297 |
| IDL_L | Simple mode | 0.837 | 1.037 | 0.735 | 1.464 |  |
| IDL_L | Weighted mode | 47 | 0.097 | 1.200 | 0.972 | 1.482 |
| IDL_L | Inverse variance weighted | 47 | 0.078 | 1.111 | 0.988 | 1.249 |
| IDL_P | MR Egger | 38 | 0.057 | 1.280 | 1.001 | 1.637 |
| IDL_P | Weighted median | 38 | 0.029 | 1.221 | 1.021 | 1.461 |
| IDL_P | Simple mode | 0.580 | 1.093 | 0.800 | 1.493 |  |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| IDL_P | Weighted mode | 38 | 0.041 | 1.272 | 1.018 | 1.590 |
| IDL_P | Inverse variance weighted | 38 | 0.044 | 1.154 | 1.004 | 1.325 |
| IDL_PL | Inverse variance weighted | 47 | 0.010 | 1.156 | 1.035 | 1.292 |
| IDL_PL | MR Egger | 47 | 0.026 | 1.243 | 1.032 | 1.496 |
| IDL_PL | Weighted median | 47 | 0.361 | 1.081 | 0.914 | 1.279 |
| IDL_PL | Simple mode | 47 | 0.196 | 1.242 | 0.899 | 1.716 |
| IDL_PL | Weighted mode | 37 | 0.072 | 1.194 | 0.989 | 1.443 |
| IDL_PL_pct | MR Egger | 0.371 | 0.891 | 0.694 | 1.143 |  |
| IDL_PL_pct | Weighted median | 33 | 0.732 | 0.971 | 0.819 | 1.151 |
| IDL_PL_pct | Simple mode | 0.831 | 1.029 | 0.794 | 1.333 |  |
| IDL_PL_pct | Weighted mode | 33 | 0.611 | 0.960 | 0.821 | 1.122 |
| IDL_PL_pct | Inverse variance weighted | 33 | 0.583 | 1.042 | 0.900 | 1.207 |
| IDL_TG | MR Egger | 43 | 0.947 | 0.996 | 0.873 | 1.135 |
| IDL_TG | Weighted median | 43 | 0.579 | 1.034 | 0.918 | 1.165 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| IDL_TG | Simple mode | 43 | 0.305 | 1.143 | 0.888 | 1.470 |
| IDL_TG | Weighted mode | 43 | 0.466 | 1.040 | 0.937 | 1.154 |
| IDL_TG | Inverse variance weighted | 43 | 0.453 | 1.034 | 0.947 | 1.130 |
| IDL_TG_pct | MR Egger | 57 | 0.346 | 0.915 | 0.762 | 1.099 |
| IDL_TG_pct | Weighted median | 57 | 0.954 | 1.004 | 0.865 | 1.166 |
| IDL_TG_pct | Simple mode | 57 | 0.853 | 1.025 | 0.787 | 1.336 |
| IDL_TG_pct | Weighted mode | 57 | 0.943 | 0.994 | 0.854 | 1.158 |
| IDL_TG_pct | Inverse variance weighted | 57 | 0.211 | 0.936 | 0.845 | 1.038 |
| Ile | MR Egger | 7 | 0.239 | 2.383 | 0.666 | 8.525 |
| Ile | Weighted median | 7 | 0.233 | 1.265 | 0.860 | 1.862 |
| Ile | Simple mode | 7 | 0.689 | 1.105 | 0.694 | 1.757 |
| Ile | Weighted mode | 7 | 0.318 | 1.250 | 0.836 | 1.869 |
| Ile | Inverse variance weighted | 73 | 0.108 | 1.397 | 0.930 | 2.099 |
| L_HDL_C | MR Egger |  | 0.752 | 0.973 | 0.821 | 1.153 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| L_HDL_C | Weighted median | 73 | 0.631 | 0.963 | 0.824 | 1.125 |
| L_HDL_C | Simple mode | 73 | 0.663 | 0.939 | 0.707 | 1.246 |
| L_HDL_C | Weighted mode | 73 | 0.621 | 0.964 | 0.833 | 1.115 |
| L_HDL_C | Inverse variance weighted | 73 | 0.690 | 1.021 | 0.921 | 1.132 |
| L_HDL_C_pct | MR Egger | 58 | 0.341 | 1.129 | 0.882 | 1.445 |
| L_HDL_C_pct | Weighted median | 58 | 0.844 | 0.979 | 0.790 | 1.212 |
| L_HDL_C_pct | Simple mode | 58 | 0.944 | 0.986 | 0.662 | 1.469 |
| L_HDL_C_pct | Weighted mode | 58 | 0.664 | 0.953 | 0.770 | 1.180 |
| L_HDL_C_pct | Inverse variance weighted | 58 | 0.165 | 1.101 | 0.961 | 1.261 |
| L_HDL_CE | MR Egger | 73 | 0.618 | 0.957 | 0.805 | 1.137 |
| L_HDL_CE | Weighted median | 73 | 0.883 | 1.012 | 0.862 | 1.188 |
| L_HDL_CE | Simple mode | 73 | 0.735 | 0.947 | 0.691 | 1.297 |
| L_HDL_CE | Weighted mode | 73 | 0.732 | 0.973 | 0.830 | 1.139 |
| L_HDL_CE | Inverse variance weighted | 73 | 0.622 | 1.026 | 0.926 | 1.138 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| L_HDL_CE_pct | MR Egger | 49 | 0.381 | 1.163 | 0.833 | 1.623 |
| L_HDL_CE_pct | Weighted median | 49 | 0.763 | 0.967 | 0.777 | 1.204 |
| L_HDL_CE_pct | Simple mode | 49 | 0.993 | 1.002 | 0.683 | 1.469 |
| L_HDL_CE_pct | Weighted mode | 49 | 0.529 | 0.932 | 0.751 | 1.158 |
| L_HDL_CE_pct | Inverse variance weighted | 49 | 0.243 | 1.108 | 0.933 | 1.317 |
| L_HDL_FC | MR Egger | 69 | 0.828 | 0.981 | 0.828 | 1.162 |
| L_HDL_FC | Weighted median | 69 | 0.596 | 0.961 | 0.831 | 1.112 |
| L_HDL_FC | Simple mode | 69 | 0.770 | 0.954 | 0.698 | 1.304 |
| L_HDL_FC | Weighted mode | 69 | 0.585 | 0.964 | 0.847 | 1.098 |
| L_HDL_FC | Inverse variance weighted | 69 | 0.640 | 0.974 | 0.874 | 1.086 |
| L_HDL_FC_pct | Inverse variance weighted | 53 | 0.039 | 1.124 | 1.006 | 1.257 |
| L_HDL_FC_pct | MR Egger | 53 | 0.390 | 1.079 | 0.909 | 1.282 |
| L_HDL_FC_pct | Weighted median | 53 | 0.993 | 1.001 | 0.863 | 1.160 |
| L_HDL_FC_pct | Simple mode | 53 | 0.798 | 1.036 | 0.794 | 1.351 |

$\left.\begin{array}{|l|l|l|l|l|l|l|}\hline \begin{array}{l}\text { Exposure (metabolic } \\ \text { trait) }\end{array} & \text { Method } & \begin{array}{l}\text { Number of } \\ \text { SNPs }\end{array} & \text { p-value } & \text { Odds ratio } & \begin{array}{l}\text { Odds ratio } \\ \text { lower 95\% CI }\end{array} \\ \hline \text { L_HDL_FC_pct } \\ \text { upper 95\% CI }\end{array}\right]$

| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| L_HDL_PL | Weighted mode | 63 | 0.664 | 0.968 | 0.838 | 1.119 |
| L_HDL_PL | Inverse variance weighted | 63 | 0.687 | 1.021 | 0.922 | 1.132 |
| L_HDL_PL_pct | Inverse variance weighted | 56 | 0.018 | 0.856 | 0.753 | 0.974 |
| L_HDL_PL_pct | MR Egger | 56 | 0.560 | 0.930 | 0.729 | 1.186 |
| L_HDL_PL_pct | Weighted median | 56 | 0.926 | 0.991 | 0.821 | 1.197 |
| L_HDL_PL_pct | Simple mode | 56 | 0.930 | 1.016 | 0.718 | 1.436 |
| L_HDL_PL_pct | Weighted mode | 56 | 0.883 | 1.016 | 0.826 | 1.249 |
| L_HDL_TG | MR Egger | 38 | 0.615 | 0.971 | 0.867 | 1.087 |
| L_HDL_TG | Weighted median | 38 | 0.740 | 1.018 | 0.914 | 1.134 |
| L_HDL_TG | Simple mode | 38 | 0.283 | 1.148 | 0.896 | 1.472 |
| L_HDL_TG | Weighted mode | 38 | 0.743 | 1.016 | 0.925 | 1.115 |
| L_HDL_TG | Inverse variance weighted | 38 | 0.637 | 1.020 | 0.939 | 1.109 |
| L_HDL_TG_pct | MR Egger | 57 | 0.469 | 0.915 | 0.720 | 1.162 |
| L_HDL_TG_pct | Weighted median | 57 | 0.627 | 1.046 | 0.872 | 1.255 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| L_HDL_TG_pct | Simple mode | 57 | 0.959 | 0.991 | 0.716 | 1.373 |
| L_HDL_TG_pct | Weighted mode | 57 | 0.705 | 1.038 | 0.858 | 1.255 |
| L_HDL_TG_pct | Inverse variance weighted | 57 | 0.724 | 0.977 | 0.859 | 1.111 |
| L_LDL_C | Inverse variance weighted | 37 | 0.000 | 1.250 | 1.104 | 1.416 |
| L_LDL_C | MR Egger | 37 | 0.022 | 1.317 | 1.051 | 1.651 |
| L_LDL_C | Weighted median | 37 | 0.003 | 1.335 | 1.104 | 1.614 |
| L_LDL_C | Simple mode | 37 | 0.308 | 1.198 | 0.851 | 1.686 |
| L_LDL_C | Weighted mode | 37 | 0.003 | 1.438 | 1.152 | 1.793 |
| L_LDL_C_pct | Inverse variance weighted | 43 | 0.001 | 1.312 | 1.116 | 1.544 |
| L_LDL_C_pct | MR Egger | 43 | 0.009 | 1.662 | 1.159 | 2.384 |
| L_LDL_C_pct | Weighted median | 43 | 0.254 | 1.137 | 0.912 | 1.417 |
| L_LDL_C_pct | Simple mode | 43 | 0.962 | 0.988 | 0.600 | 1.627 |
| L_LDL_C_pct | Weighted mode | 43 | 0.783 | 1.046 | 0.762 | 1.436 |
| L_LDL_CE | Inverse variance weighted | 35 | 0.001 | 1.235 | 1.085 | 1.406 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| L_LDL_CE | MR Egger | 35 | 0.033 | 1.310 | 1.032 | 1.662 |
| L_LDL_CE | Weighted median | 35 | 0.003 | 1.336 | 1.103 | 1.620 |
| L_LDL_CE | Simple mode | 35 | 0.404 | 1.170 | 0.813 | 1.683 |
| L_LDL_CE | Weighted mode | 35 | 0.025 | 1.353 | 1.050 | 1.743 |
| L_LDL_CE_pct | Inverse variance weighted | 26 | 0.008 | 1.317 | 1.073 | 1.617 |
| L_LDL_CE_pct | MR Egger | 26 | 0.189 | 1.394 | 0.861 | 2.257 |
| L_LDL_CE_pct | Weighted median | 26 | 0.186 | 1.187 | 0.921 | 1.530 |
| L_LDL_CE_pct | Simple mode | 26 | 0.353 | 1.301 | 0.755 | 2.243 |
| L_LDL_CE_pct | Weighted mode | 26 | 0.375 | 1.153 | 0.847 | 1.570 |
| L_LDL_FC | Inverse variance weighted | 43 | 0.005 | 1.196 | 1.055 | 1.355 |
| L_LDL_FC | MR Egger | 43 | 0.007 | 1.371 | 1.104 | 1.702 |
| L_LDL_FC | Weighted median | 43 | 0.005 | 1.310 | 1.085 | 1.582 |
| L_LDL_FC | Simple mode | 43 | 0.802 | 1.051 | 0.716 | 1.541 |
| L_LDL_FC | Weighted mode | 43 | 0.003 | 1.398 | 1.131 | 1.728 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| L_LDL_FC_pct | MR Egger | 58 | 0.922 | 0.989 | 0.789 | 1.239 |
| L_LDL_FC_pct | Weighted median | 58 | 0.611 | 0.954 | 0.794 | 1.145 |
| L_LDL_FC_pct | Simple mode | 58 | 0.818 | 0.962 | 0.693 | 1.335 |
| L_LDL_FC_pct | Weighted mode | 58 | 0.625 | 0.945 | 0.756 | 1.182 |
| L_LDL_FC_pct | Inverse variance weighted | 58 | 0.856 | 1.012 | 0.892 | 1.148 |
| L_LDL_L | Inverse variance weighted | 35 | 0.003 | 1.219 | 1.069 | 1.389 |
| L_LDL_L | MR Egger | 35 | 0.042 | 1.292 | 1.018 | 1.639 |
| L_LDL_L | Weighted median | 35 | 0.008 | 1.302 | 1.072 | 1.580 |
| L_LDL_L | Simple mode | 35 | 0.551 | 1.111 | 0.789 | 1.563 |
| L_LDL_L | Weighted mode | 35 | 0.009 | 1.383 | 1.099 | 1.741 |
| L_LDL_P | Inverse variance weighted | 42 | 0.009 | 1.175 | 1.041 | 1.327 |
| L_LDL_P | MR Egger | 42 | 0.011 | 1.326 | 1.078 | 1.631 |
| L_LDL_P | Weighted median | 42 | 0.003 | 1.303 | 1.091 | 1.555 |
| L_LDL_P | Simple mode | 0.391 | 1.155 | 0.834 | 1.601 |  |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| L_LDL_P | Weighted mode | 42 | 0.012 | 1.326 | 1.074 | 1.638 |
| L_LDL_PL | Inverse variance weighted | 35 | 0.009 | 1.197 | 1.047 | 1.369 |
| L_LDL_PL | MR Egger | 35 | 0.017 | 1.346 | 1.069 | 1.696 |
| L_LDL_PL | Weighted median | 35 | 0.009 | 1.298 | 1.067 | 1.579 |
| L_LDL_PL | Simple mode | 35 | 0.406 | 1.163 | 0.818 | 1.654 |
| L_LDL_PL | Weighted mode | 35 | 0.008 | 1.385 | 1.105 | 1.735 |
| L_LDL_PL_pct | MR Egger | 33 | 0.502 | 1.057 | 0.901 | 1.238 |
| L_LDL_PL_pct | Weighted median | 33 | 0.445 | 1.055 | 0.919 | 1.211 |
| L_LDL_PL_pct | Simple mode | 33 | 0.450 | 0.900 | 0.686 | 1.180 |
| L_LDL_PL_pct | Weighted mode | 33 | 0.490 | 1.046 | 0.922 | 1.187 |
| L_LDL_PL_pct | Inverse variance weighted | 33 | 0.901 | 0.994 | 0.899 | 1.098 |
| L_LDL_TG | MR Egger | 48 | 0.605 | 1.038 | 0.901 | 1.197 |
| L_LDL_TG | Weighted median | 48 | 0.593 | 1.039 | 0.904 | 1.193 |
| L_LDL_TG | Simple mode | 48 | 0.643 | 1.064 | 0.819 | 1.382 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| L_LDL_TG | Weighted mode | 48 | 0.449 | 1.045 | 0.933 | 1.170 |
| L_LDL_TG | Inverse variance weighted | 48 | 0.517 | 1.032 | 0.939 | 1.133 |
| L_LDL_TG_pct | MR Egger | 48 | 0.377 | 0.902 | 0.718 | 1.132 |
| L_LDL_TG_pct | Weighted median | 48 | 0.975 | 1.003 | 0.840 | 1.197 |
| L_LDL_TG_pct | Simple mode | 48 | 0.891 | 1.023 | 0.741 | 1.412 |
| L_LDL_TG_pct | Weighted mode | 48 | 0.881 | 1.014 | 0.845 | 1.218 |
| L_LDL_TG_pct | Inverse variance weighted | 48 | 0.337 | 0.940 | 0.830 | 1.066 |
| L_VLDL_C | MR Egger | 39 | 0.461 | 1.107 | 0.848 | 1.444 |
| L_VLDL_C | Weighted median | 39 | 0.375 | 1.078 | 0.913 | 1.274 |
| L_VLDL_C | Simple mode | 39 | 0.426 | 1.107 | 0.864 | 1.419 |
| L_VLDL_C | Weighted mode | 39 | 0.287 | 1.096 | 0.928 | 1.296 |
| L_VLDL_C | Inverse variance weighted | 39 | 0.349 | 1.068 | 0.930 | 1.227 |
| L_VLDL_C_pct | MR Egger | 42 | 0.791 | 0.968 | 0.761 | 1.230 |
| L_VLDL_C_pct | Weighted median | 42 | 0.948 | 0.994 | 0.840 | 1.178 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| L_VLDL_C_pct | Simple mode | 42 | 0.809 | 0.963 | 0.711 | 1.305 |
| L_VLDL_C_pct | Weighted mode | 42 | 0.828 | 0.983 | 0.840 | 1.150 |
| L_VLDL_C_pct | Inverse variance weighted | 42 | 0.793 | 0.981 | 0.848 | 1.134 |
| L_VLDL_CE | MR Egger | 36 | 0.768 | 0.959 | 0.727 | 1.264 |
| L_VLDL_CE | Weighted median | 36 | 0.175 | 1.129 | 0.947 | 1.346 |
| L_VLDL_CE | Simple mode | 36 | 0.429 | 1.112 | 0.857 | 1.444 |
| L_VLDL_CE | Weighted mode | 36 | 0.346 | 1.091 | 0.912 | 1.306 |
| L_VLDL_CE | Inverse variance weighted | 36 | 0.123 | 1.123 | 0.969 | 1.301 |
| L_VLDL_CE_pct | MR Egger | 46 | 0.591 | 0.938 | 0.743 | 1.183 |
| L_VLDL_CE_pct | Weighted median | 46 | 0.711 | 0.968 | 0.818 | 1.147 |
| L_VLDL_CE_pct | Simple mode | 46 | 0.706 | 0.944 | 0.699 | 1.273 |
| L_VLDL_CE_pct | Weighted mode | 46 | 0.613 | 0.960 | 0.821 | 1.123 |
| L_VLDL_CE_pct | Inverse variance weighted | 46 | 0.921 | 0.993 | 0.863 | 1.143 |
| L_VLDL_FC | MR Egger | 40 | 0.269 | 1.168 | 0.891 | 1.530 |

$\left.\begin{array}{|l|l|l|l|l|l|l|}\hline \begin{array}{l}\text { Exposure (metabolic } \\ \text { trait) }\end{array} & \text { Method } & \begin{array}{l}\text { Number of } \\ \text { SNPs }\end{array} & \text { p-value } & \text { Odds ratio } & \begin{array}{l}\text { Odds ratio } \\ \text { lower 95\% CI }\end{array} \\ \hline \text { L_VLDL_FC } \\ \text { upper 95\% CI }\end{array}\right]$

| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| L_VLDL_P | MR Egger | 41 | 0.401 | 1.106 | 0.877 | 1.394 |
| L_VLDL_P | Weighted median | 41 | 0.306 | 1.097 | 0.919 | 1.309 |
| L_VLDL_P | Simple mode | 41 | 0.483 | 1.117 | 0.823 | 1.515 |
| L_VLDL_P | Weighted mode | 41 | 0.318 | 1.107 | 0.909 | 1.349 |
| L_VLDL_P | Inverse variance weighted | 41 | 0.237 | 1.080 | 0.951 | 1.228 |
| L_VLDL_PL | MR Egger | 39 | 0.276 | 1.151 | 0.897 | 1.479 |
| L_VLDL_PL | Weighted median | 39 | 0.299 | 1.095 | 0.923 | 1.299 |
| L_VLDL_PL | Simple mode | 39 | 0.556 | 1.090 | 0.821 | 1.448 |
| L_VLDL_PL | Weighted mode | 39 | 0.326 | 1.108 | 0.906 | 1.355 |
| L_VLDL_PL | Inverse variance weighted | 39 | 0.399 | 1.059 | 0.927 | 1.211 |
| L_VLDL_PL_pct | MR Egger | 45 | 0.337 | 0.878 | 0.674 | 1.142 |
| L_VLDL_PL_pct | Weighted median | 45 | 0.854 | 1.018 | 0.838 | 1.238 |
| L_VLDL_PL_pct | Simple mode | 45 | 0.679 | 1.077 | 0.760 | 1.527 |
| L_VLDL_PL_pct | Weighted mode | 45 | 0.757 | 1.032 | 0.848 | 1.255 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| L_VLDL_PL_pct | Inverse variance weighted | 45 | 0.600 | 0.962 | 0.834 | 1.110 |
| L_VLDL_TG | MR Egger | 43 | 0.288 | 1.139 | 0.899 | 1.444 |
| L_VLDL_TG | Weighted median | 43 | 0.324 | 1.099 | 0.911 | 1.326 |
| L_VLDL_TG | Simple mode | 43 | 0.659 | 1.066 | 0.805 | 1.412 |
| L_VLDL_TG | Weighted mode | 43 | 0.276 | 1.122 | 0.915 | 1.376 |
| L_VLDL_TG | Inverse variance weighted | 43 | 0.397 | 1.063 | 0.923 | 1.224 |
| L_VLDL_TG_pct | MR Egger | 42 | 0.582 | 1.049 | 0.887 | 1.240 |
| L_VLDL_TG_pct | Weighted median | 42 | 0.411 | 0.941 | 0.815 | 1.087 |
| L_VLDL_TG_pct | Simple mode | 42 | 0.577 | 0.916 | 0.675 | 1.244 |
| L_VLDL_TG_pct | Weighted mode | 42 | 0.510 | 0.957 | 0.840 | 1.090 |
| L_VLDL_TG_pct | Inverse variance weighted | 42 | 0.218 | 0.932 | 0.833 | 1.042 |
| LA | MR Egger | 36 | 0.661 | 0.920 | 0.634 | 1.334 |
| LA | Weighted median | 36 | 0.731 | 1.033 | 0.860 | 1.240 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| LA | Weighted mode | 36 | 0.991 | 1.001 | 0.807 | 1.242 |
| LA | Inverse variance weighted | 36 | 0.529 | 0.949 | 0.806 | 1.117 |
| LA_pct | MR Egger | 24 | 0.996 | 0.999 | 0.610 | 1.635 |
| LA_pct | Weighted median | 24 | 0.503 | 0.893 | 0.641 | 1.243 |
| LA_pct | Simple mode | 24 | 0.442 | 0.799 | 0.455 | 1.402 |
| LA_pct | Weighted mode | 24 | 0.492 | 0.875 | 0.602 | 1.272 |
| LA_pct | Inverse variance weighted | 24 | 0.133 | 0.846 | 0.680 | 1.053 |
| Lactate | Inverse variance weighted | 6 | 0.086 | 1.488 | 0.945 | 2.343 |
| Lactate | MR Egger | 6 | 0.794 | 1.375 | 0.147 | 12.850 |
| Lactate | Weighted median | 6 | 0.132 | 1.520 | 0.881 | 2.624 |
| Lactate | Simple mode | 6 | 0.341 | 1.504 | 0.704 | 3.214 |
| Lactate | Weighted mode | 6 | 0.237 | 1.629 | 0.799 | 3.320 |
| LDL_C | Inverse variance weighted | 36 | 0.007 | 1.209 | 1.054 | 1.387 |
| LDL_C | MR Egger | 36 | 0.024 | 1.335 | 1.051 | 1.695 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| LDL_C | Weighted median | 36 | 0.003 | 1.335 | 1.103 | 1.615 |
| LDL_C | Simple mode | 36 | 0.198 | 1.256 | 0.894 | 1.765 |
| LDL_C | Weighted mode | 36 | 0.010 | 1.400 | 1.099 | 1.783 |
| LDL_CE | Inverse variance weighted | 33 | 0.008 | 1.214 | 1.052 | 1.401 |
| LDL_CE | MR Egger | 33 | 0.042 | 1.324 | 1.021 | 1.717 |
| LDL_CE | Weighted median | 33 | 0.006 | 1.320 | 1.084 | 1.608 |
| LDL_CE | Simple mode | 33 | 0.293 | 1.215 | 0.850 | 1.738 |
| LDL_CE | Weighted mode | 33 | 0.017 | 1.359 | 1.069 | 1.727 |
| LDL_FC | Inverse variance weighted | 39 | 0.004 | 1.203 | 1.062 | 1.362 |
| LDL_FC | MR Egger | 39 | 0.003 | 1.408 | 1.140 | 1.740 |
| LDL_FC | Weighted median | 39 | 0.004 | 1.326 | 1.095 | 1.607 |
| LDL_FC | Simple mode | 39 | 0.795 | 1.053 | 0.716 | 1.548 |
| LDL_FC | Weighted mode | 39 | 0.003 | 1.401 | 1.133 | 1.733 |
| LDL_L | Inverse variance weighted | 35 | 0.004 | 1.216 | 1.064 | 1.390 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| LDL_L | MR Egger | 35 | 0.034 | 1.309 | 1.031 | 1.663 |
| LDL_L | Weighted median | 35 | 0.009 | 1.302 | 1.070 | 1.585 |
| LDL_L | Simple mode | 35 | 0.315 | 1.203 | 0.844 | 1.714 |
| LDL_L | Weighted mode | 35 | 0.010 | 1.371 | 1.094 | 1.718 |
| LDL_P | Inverse variance weighted | 40 | 0.026 | 1.157 | 1.018 | 1.315 |
| LDL_P | MR Egger | 40 | 0.067 | 1.248 | 0.991 | 1.570 |
| LDL_P | Weighted median | 40 | 0.068 | 1.192 | 0.987 | 1.439 |
| LDL_P | Simple mode | 0.329 | 1.205 | 0.833 | 1.742 |  |
| LDL_P | Weighted mode | 40 | 0.119 | 1.214 | 0.956 | 1.541 |
| LDL_PL | Inverse variance weighted | 36 | 0.011 | 1.188 | 1.040 | 1.357 |
| LDL_PL | MR Egger | 36 | 0.027 | 1.306 | 1.041 | 1.638 |
| LDL_PL | Weighted median | 36 | 0.011 | 1.274 | 1.056 | 1.536 |
| LDL_PL | Simple mode | 36 | 0.631 | 1.098 | 0.752 | 1.605 |
| LDL_PL | Weighted mode | 0.011 | 1.324 | 1.078 | 1.628 |  |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| LDL_size | MR Egger | 27 | 0.295 | 1.204 | 0.857 | 1.690 |
| LDL_size | Weighted median | 27 | 0.400 | 1.113 | 0.867 | 1.429 |
| LDL_size | Simple mode | 27 | 0.865 | 0.959 | 0.594 | 1.549 |
| LDL_size | Weighted mode | 27 | 0.444 | 1.111 | 0.852 | 1.450 |
| LDL_size | Inverse variance weighted | 27 | 0.498 | 1.070 | 0.881 | 1.299 |
| LDL_TG | MR Egger | 53 | 0.359 | 1.074 | 0.923 | 1.250 |
| LDL_TG | Weighted median | 53 | 0.507 | 1.046 | 0.916 | 1.193 |
| LDL_TG | Simple mode | 53 | 0.532 | 1.081 | 0.849 | 1.376 |
| LDL_TG | Weighted mode | 53 | 0.455 | 1.047 | 0.929 | 1.179 |
| LDL_TG | Inverse variance weighted | 53 | 0.716 | 1.018 | 0.926 | 1.118 |
| Leu | MR Egger | 11 | 0.231 | 1.885 | 0.717 | 4.956 |
| Leu | Weighted median | 11 | 0.200 | 1.251 | 0.888 | 1.762 |
| Leu | Simple mode | 0.463 | 11 | 1.292 | 0.669 | 2.494 |
| Leu | Weighted mode | 0.218 | 1.267 | 0.890 | 1.801 |  |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Leu | Inverse variance weighted | 11 | 0.292 | 1.259 | 0.820 | 1.931 |
| M_HDL_C | MR Egger | 65 | 0.919 | 1.015 | 0.764 | 1.348 |
| M_HDL_C | Weighted median | 65 | 0.908 | 0.990 | 0.829 | 1.182 |
| M_HDL_C | Simple mode | 65 | 0.803 | 0.958 | 0.685 | 1.340 |
| M_HDL_C | Weighted mode | 65 | 0.822 | 0.970 | 0.745 | 1.264 |
| M_HDL_C | Inverse variance weighted | 65 | 0.848 | 1.014 | 0.879 | 1.170 |
| M_HDL_C_pct | MR Egger | 58 | 0.639 | 1.076 | 0.793 | 1.461 |
| M_HDL_C_pct | Weighted median | 58 | 0.800 | 0.973 | 0.790 | 1.200 |
| M_HDL_C_pct | Simple mode | 58 | 0.927 | 0.982 | 0.668 | 1.443 |
| M_HDL_C_pct | Weighted mode | 58 | 0.634 | 0.938 | 0.723 | 1.218 |
| M_HDL_C_pct | Inverse variance weighted | 58 | 0.214 | 1.097 | 0.948 | 1.270 |
| M_HDL_CE | MR Egger | 66 | 0.846 | 0.971 | 0.726 | 1.300 |
| M_HDL_CE | Weighted median | 66 | 0.864 | 0.985 | 0.829 | 1.171 |
| M_HDL_CE | Simple mode | 66 | 0.659 | 0.928 | 0.666 | 1.292 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| M_HDL_CE | Weighted mode | 66 | 0.770 | 0.962 | 0.744 | 1.245 |
| M_HDL_CE | Inverse variance weighted | 66 | 0.797 | 0.981 | 0.846 | 1.137 |
| M_HDL_CE_pct | MR Egger | 50 | 0.619 | 0.936 | 0.722 | 1.213 |
| M_HDL_CE_pct | Weighted median | 50 | 0.452 | 0.930 | 0.771 | 1.123 |
| M_HDL_CE_pct | Simple mode | 50 | 0.842 | 0.968 | 0.708 | 1.325 |
| M_HDL_CE_pct | Weighted mode | 50 | 0.519 | 0.936 | 0.765 | 1.144 |
| M_HDL_CE_pct | Inverse variance weighted | 50 | 0.539 | 1.044 | 0.911 | 1.195 |
| M_HDL_FC | MR Egger | 58 | 0.585 | 1.078 | 0.824 | 1.412 |
| M_HDL_FC | Weighted median | 58 | 0.849 | 1.018 | 0.845 | 1.227 |
| M_HDL_FC | Simple mode | 58 | 0.904 | 0.981 | 0.715 | 1.345 |
| M_HDL_FC | Weighted mode | 58 | 0.853 | 0.981 | 0.797 | 1.206 |
| M_HDL_FC | Inverse variance weighted | 58 | 0.821 | 1.017 | 0.877 | 1.179 |
| M_HDL_FC_pct | Inverse variance weighted | 61 | 0.034 | 1.128 | 1.009 | 1.262 |
| M_HDL_FC_pct | MR Egger | 61 | 0.213 | 1.136 | 0.931 | 1.386 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| M_HDL_FC_pct | Weighted median | 61 | 0.546 | 1.054 | 0.889 | 1.249 |
| M_HDL_FC_pct | Simple mode | 61 | 0.589 | 0.925 | 0.700 | 1.224 |
| M_HDL_FC_pct | Weighted mode | 61 | 0.686 | 1.035 | 0.877 | 1.221 |
| M_HDL_L | MR Egger | 49 | 0.699 | 1.064 | 0.779 | 1.452 |
| M_HDL_L | Weighted median | 49 | 0.813 | 1.023 | 0.847 | 1.236 |
| M_HDL_L | Simple mode | 49 | 0.850 | 0.967 | 0.680 | 1.374 |
| M_HDL_L | Weighted mode | 49 | 0.834 | 0.979 | 0.801 | 1.196 |
| M_HDL_L | Inverse variance weighted | 51 | 0.775 | 1.005 | 0.854 | 1.184 |
| M_HDL_P | MR Egger | 51 | 0.821 | 1.046 | 0.769 | 1.424 |
| M_HDL_P | Weighted median | 51 | 0.986 | 0.92 | 0.845 | 1.237 |
| M_HDL_P | Simple mode | 51 | 0.886 | 0.985 | 0.720 | 1.380 |
| M_HDL_P | Weighted mode | 51 | 0.860 | 1.015 | 0.798 | 1.215 |
| M_HDL_P | Inverse variance weighted | 47 | 0.655 | 0.938 | 0.711 | 1.196 |
| M_HDL_PL | MR Egger |  |  | 1.238 |  |  |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| M_HDL_PL | Weighted median | 47 | 0.832 | 1.020 | 0.852 | 1.221 |
| M_HDL_PL | Simple mode | 47 | 0.922 | 0.984 | 0.713 | 1.357 |
| M_HDL_PL | Weighted mode | 47 | 0.877 | 0.984 | 0.802 | 1.208 |
| M_HDL_PL | Inverse variance weighted | 47 | 0.866 | 0.988 | 0.856 | 1.139 |
| M_HDL_PL_pct | MR Egger | 55 | 0.229 | 0.783 | 0.528 | 1.161 |
| M_HDL_PL_pct | Weighted median | 55 | 0.883 | 0.985 | 0.805 | 1.206 |
| M_HDL_PL_pct | Simple mode | 55 | 0.982 | 0.995 | 0.649 | 1.525 |
| M_HDL_PL_pct | Weighted mode | 55 | 0.723 | 1.068 | 0.745 | 1.530 |
| M_HDL_PL_pct | Inverse variance weighted | 55 | 0.229 | 0.897 | 0.753 | 1.070 |
| M_HDL_TG | MR Egger | 40 | 0.817 | 1.024 | 0.839 | 1.249 |
| M_HDL_TG | Weighted median | 40 | 0.537 | 1.052 | 0.896 | 1.236 |
| M_HDL_TG | Simple mode | 40 | 0.463 | 1.120 | 0.830 | 1.510 |
| M_HDL_TG | Weighted mode | 40 | 0.456 | 1.058 | 0.914 | 1.223 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| M_HDL_TG_pct | MR Egger | 49 | 0.599 | 0.935 | 0.730 | 1.198 |
| M_HDL_TG_pct | Weighted median | 49 | 0.404 | 1.074 | 0.908 | 1.272 |
| M_HDL_TG_pct | Simple mode | 49 | 0.807 | 1.037 | 0.777 | 1.383 |
| M_HDL_TG_pct | Weighted mode | 49 | 0.524 | 1.063 | 0.881 | 1.283 |
| M_HDL_TG_pct | Inverse variance weighted | 49 | 0.598 | 0.965 | 0.847 | 1.100 |
| M_LDL_C | Inverse variance weighted | 35 | 0.034 | 1.160 | 1.012 | 1.330 |
| M_LDL_C | MR Egger | 35 | 0.104 | 1.246 | 0.963 | 1.613 |
| M_LDL_C | Weighted median | 35 | 0.141 | 1.158 | 0.952 | 1.407 |
| M_LDL_C | Simple mode | 35 | 0.841 | 1.038 | 0.725 | 1.485 |
| M_LDL_C | Weighted mode | 35 | 0.849 | 1.031 | 0.758 | 1.402 |
| M_LDL_C_pct | Inverse variance weighted | 28 | 0.012 | 1.209 | 1.043 | 1.402 |
| M_LDL_C_pct | MR Egger | 28 | 0.481 | 1.114 | 0.828 | 1.499 |
| M_LDL_C_pct | Weighted median | 28 | 0.190 | 1.147 | 0.934 | 1.408 |
| M_LDL_C_pct | Simple mode | 28 | 0.723 | 1.063 | 0.760 | 1.487 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| M_LDL_C_pct | Weighted mode | 28 | 0.348 | 1.109 | 0.896 | 1.373 |
| M_LDL_CE | Inverse variance weighted | 37 | 0.084 | 1.138 | 0.983 | 1.317 |
| M_LDL_CE | MR Egger | 37 | 0.258 | 1.178 | 0.891 | 1.556 |
| M_LDL_CE | Weighted median | 37 | 0.344 | 1.102 | 0.901 | 1.348 |
| M_LDL_CE | Simple mode | 37 | 0.730 | 1.065 | 0.747 | 1.519 |
| M_LDL_CE | Weighted mode | 37 | 0.744 | 1.050 | 0.785 | 1.404 |
| M_LDL_CE_pct | MR Egger | 32 | 0.967 | 0.995 | 0.783 | 1.265 |
| M_LDL_CE_pct | Weighted median | 32 | 0.886 | 1.014 | 0.842 | 1.221 |
| M_LDL_CE_pct | Simple mode | 32 | 0.303 | 1.176 | 0.868 | 1.593 |
| M_LDL_CE_pct | Weighted mode | 32 | 0.999 | 1.000 | 0.818 | 1.223 |
| M_LDL_CE_pct | Inverse variance weighted | 32 | 0.471 | 1.050 | 0.919 | 1.200 |
| M_LDL_FC | Inverse variance weighted | 38 | 0.006 | 1.193 | 1.052 | 1.353 |
| M_LDL_FC | MR Egger | 38 | 0.006 | 1.373 | 1.110 | 1.699 |
| M_LDL_FC | Weighted median | 38 | 0.007 | 1.300 | 1.075 | 1.573 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| M_LDL_FC | Simple mode | 38 | 0.456 | 1.153 | 0.796 | 1.671 |
| M_LDL_FC | Weighted mode | 38 | 0.005 | 1.389 | 1.121 | 1.722 |
| M_LDL_FC_pct | MR Egger | 58 | 0.387 | 0.918 | 0.756 | 1.113 |
| M_LDL_FC_pct | Weighted median | 58 | 0.458 | 0.939 | 0.794 | 1.109 |
| M_LDL_FC_pct | Simple mode | 58 | 0.818 | 0.965 | 0.710 | 1.310 |
| M_LDL_FC_pct | Weighted mode | 58 | 0.479 | 0.934 | 0.773 | 1.128 |
| M_LDL_FC_pct | Inverse variance weighted | 58 | 0.885 | 0.992 | 0.887 | 1.109 |
| M_LDL_L | Inverse variance weighted | 34 | 0.040 | 1.156 | 1.007 | 1.327 |
| M_LDL_L | MR Egger | 34 | 0.145 | 1.221 | 0.940 | 1.586 |
| M_LDL_L | Weighted median | 34 | 0.152 | 1.153 | 0.949 | 1.401 |
| M_LDL_L | Simple mode | 34 | 0.892 | 1.027 | 0.701 | 1.505 |
| M_LDL_L | Weighted mode | 34 | 0.825 | 1.034 | 0.770 | 1.388 |
| M_LDL_P | Inverse variance weighted | 35 | 0.021 | 1.186 | 1.026 | 1.371 |
| M_LDL_P | MR Egger | 35 | 0.158 | 1.229 | 0.929 | 1.626 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| M_LDL_P | Weighted median | 35 | 0.066 | 1.204 | 0.988 | 1.468 |
| M_LDL_P | Simple mode | 35 | 0.676 | 1.092 | 0.726 | 1.641 |
| M_LDL_P | Weighted mode | 35 | 0.437 | 1.124 | 0.839 | 1.506 |
| M_LDL_PL | Inverse variance weighted | 34 | 0.033 | 1.176 | 1.013 | 1.366 |
| M_LDL_PL | MR Egger | 34 | 0.188 | 1.204 | 0.919 | 1.579 |
| M_LDL_PL | Weighted median | 34 | 0.030 | 1.240 | 1.021 | 1.504 |
| M_LDL_PL | Simple mode | 34 | 0.569 | 1.112 | 0.775 | 1.595 |
| M_LDL_PL | Weighted mode | 34 | 0.068 | 1.259 | 0.992 | 1.598 |
| M_LDL_PL_pct | MR Egger | 30 | 0.628 | 1.049 | 0.866 | 1.272 |
| M_LDL_PL_pct | Weighted median | 30 | 0.892 | 0.990 | 0.851 | 1.151 |
| M_LDL_PL_pct | Simple mode | 30 | 0.554 | 0.920 | 0.700 | 1.209 |
| M_LDL_PL_pct | Weighted mode | 30 | 0.822 | 1.017 | 0.878 | 1.178 |
| M_LDL_PL_pct | Inverse variance weighted | 30 | 0.545 | 0.962 | 0.848 | 1.091 |
| M_LDL_TG | MR Egger | 51 | 0.714 | 1.032 | 0.872 | 1.222 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| M_LDL_TG | Weighted median | 51 | 0.473 | 1.056 | 0.911 | 1.224 |
| M_LDL_TG | Simple mode | 51 | 0.489 | 1.083 | 0.866 | 1.355 |
| M_LDL_TG | Weighted mode | 51 | 0.442 | 1.061 | 0.914 | 1.231 |
| M_LDL_TG | Inverse variance weighted | 51 | 0.329 | 1.050 | 0.952 | 1.160 |
| M_LDL_TG_pct | MR Egger | 49 | 0.230 | 0.894 | 0.745 | 1.071 |
| M_LDL_TG_pct | Weighted median | 49 | 0.439 | 0.948 | 0.829 | 1.084 |
| M_LDL_TG_pct | Simple mode | 49 | 0.907 | 1.017 | 0.773 | 1.337 |
| M_LDL_TG_pct | Weighted mode | 49 | 0.659 | 0.971 | 0.851 | 1.107 |
| M_LDL_TG_pct | Inverse variance weighted | 49 | 0.445 | 0.956 | 0.853 | 1.072 |
| M_VLDL_C | Inverse variance weighted | 38 | 0.039 | 1.156 | 1.007 | 1.328 |
| M_VLDL_C | MR Egger | 38 | 0.058 | 1.276 | 1.000 | 1.629 |
| M_VLDL_C | Weighted median | 38 | 0.075 | 1.190 | 0.983 | 1.442 |
| M_VLDL_C | Simple mode | 38 | 0.340 | 1.222 | 0.814 | 1.836 |
| M_VLDL_C | Weighted mode | 38 | 0.102 | 1.213 | 0.968 | 1.519 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| M_VLDL_C_pct | MR Egger | 56 | 0.344 | 1.133 | 0.877 | 1.463 |
| M_VLDL_C_pct | Weighted median | 56 | 0.690 | 1.039 | 0.863 | 1.251 |
| M_VLDL_C_pct | Simple mode | 56 | 0.885 | 1.030 | 0.694 | 1.527 |
| M_VLDL_C_pct | Weighted mode | 56 | 0.981 | 0.996 | 0.743 | 1.337 |
| M_VLDL_C_pct | Inverse variance weighted | 56 | 0.251 | 1.082 | 0.946 | 1.238 |
| M_VLDL_CE | Inverse variance weighted | 38 | 0.033 | 1.163 | 1.013 | 1.336 |
| M_VLDL_CE | MR Egger | 38 | 0.035 | 1.305 | 1.028 | 1.657 |
| M_VLDL_CE | Weighted median | 38 | 0.059 | 1.208 | 0.993 | 1.471 |
| M_VLDL_CE | Simple mode | 38 | 0.128 | 1.379 | 0.920 | 2.065 |
| M_VLDL_CE | Weighted mode | 38 | 0.023 | 1.344 | 1.052 | 1.717 |
| M_VLDL_CE_pct | MR Egger | 54 | 0.210 | 1.191 | 0.909 | 1.560 |
| M_VLDL_CE_pct | Weighted median | 54 | 0.998 | 1.000 | 0.827 | 1.210 |
| M_VLDL_CE_pct | Simple mode | 54 | 0.871 | 0.968 | 0.651 | 1.439 |
| M_VLDL_CE_pct | Weighted mode | 54 | 0.893 | 0.979 | 0.714 | 1.341 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| M_VLDL_CE_pct | Inverse variance weighted | 54 | 0.547 | 1.045 | 0.906 | 1.206 |
| M_VLDL_FC | Inverse variance weighted | 41 | 0.013 | 1.192 | 1.038 | 1.368 |
| M_VLDL_FC | MR Egger | 41 | 0.161 | 1.203 | 0.933 | 1.552 |
| M_VLDL_FC | Weighted median | 41 | 0.076 | 1.186 | 0.982 | 1.433 |
| M_VLDL_FC | Simple mode | 41 | 0.574 | 1.103 | 0.786 | 1.548 |
| M_VLDL_FC | Weighted mode | 41 | 0.451 | 1.103 | 0.857 | 1.420 |
| M_VLDL_FC_pct | MR Egger | 45 | 0.283 | 1.166 | 0.884 | 1.538 |
| M_VLDL_FC_pct | Weighted median | 45 | 0.638 | 1.044 | 0.872 | 1.251 |
| M_VLDL_FC_pct | Simple mode | 45 | 0.898 | 1.026 | 0.698 | 1.507 |
| M_VLDL_FC_pct | Weighted mode | 45 | 0.913 | 1.016 | 0.759 | 1.362 |
| M_VLDL_FC_pct | Inverse variance weighted | 45 | 0.433 | 1.060 | 0.916 | 1.227 |
| M_VLDL_L | Inverse variance weighted | 36 | 0.219 | 1.120 | 0.935 | 1.342 |
| M_VLDL_L | MR Egger | 36 | 0.984 | 1.004 | 0.703 | 1.433 |
| M_VLDL_L | Weighted median | 36 | 0.381 | 1.086 | 0.903 | 1.305 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| M_VLDL_L | Simple mode | 36 | 0.362 | 1.136 | 0.866 | 1.491 |
| M_VLDL_L | Weighted mode | 36 | 0.379 | 1.094 | 0.898 | 1.333 |
| M_VLDL_P | MR Egger | 43 | 0.607 | 1.099 | 0.770 | 1.569 |
| M_VLDL_P | Weighted median | 43 | 0.511 | 1.063 | 0.887 | 1.273 |
| M_VLDL_P | Simple mode | 43 | 0.484 | 1.101 | 0.843 | 1.438 |
| M_VLDL_P | Weighted mode | 43 | 0.484 | 1.079 | 0.874 | 1.330 |
| M_VLDL_P | Inverse variance weighted | 43 | 0.279 | 1.107 | 0.921 | 1.329 |
| M_VLDL_PL | Inverse variance weighted | 45 | 0.134 | 1.134 | 0.962 | 1.337 |
| M_VLDL_PL | MR Egger | 0.601 | 1.085 | 0.800 | 1.473 |  |
| M_VLDL_PL | Weighted median | 45 | 0.291 | 1.102 | 0.920 | 1.319 |
| M_VLDL_PL | Simple mode | 45 | 0.544 | 1.096 | 0.817 | 1.471 |
| M_VLDL_PL | Weighted mode | 45 | 0.516 | 1.073 | 0.868 | 1.327 |
| M_VLDL_PL_pct | MR Egger | 50 | 0.145 | 1.190 | 0.946 | 1.497 |
| M_VLDL_PL_pct | Weighted median | 50 | 0.642 | 1.042 | 0.875 | 1.242 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| M_VLDL_PL_pct | Simple mode | 50 | 0.984 | 0.997 | 0.722 | 1.375 |
| M_VLDL_PL_pct | Weighted mode | 50 | 0.802 | 1.033 | 0.802 | 1.331 |
| M_VLDL_PL_pct | Inverse variance weighted | 50 | 0.273 | 1.075 | 0.944 | 1.225 |
| M_VLDL_TG | Inverse variance weighted | 46 | 0.217 | 1.076 | 0.958 | 1.209 |
| M_VLDL_TG | MR Egger | 46 | 0.380 | 1.096 | 0.895 | 1.342 |
| M_VLDL_TG | Weighted median | 46 | 0.447 | 1.067 | 0.903 | 1.262 |
| M_VLDL_TG | Simple mode | 46 | 0.580 | 1.078 | 0.829 | 1.401 |
| M_VLDL_TG | Weighted mode | 46 | 0.418 | 1.078 | 0.901 | 1.290 |
| M_VLDL_TG_pct | MR Egger | 49 | 0.369 | 0.877 | 0.661 | 1.164 |
| M_VLDL_TG_pct | Weighted median | 49 | 0.753 | 0.969 | 0.795 | 1.180 |
| M_VLDL_TG_pct | Simple mode | 49 | 0.962 | 0.990 | 0.659 | 1.487 |
| M_VLDL_TG_pct | Weighted mode | 49 | 0.995 | 1.001 | 0.733 | 1.367 |
| M_VLDL_TG_pct | Inverse variance weighted | 49 | 0.416 | 0.938 | 0.805 | 1.094 |
| MUFA | Inverse variance weighted | 46 | 0.130 | 1.100 | 0.972 | 1.244 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| MUFA | MR Egger | 46 | 0.968 | 0.995 | 0.798 | 1.241 |
| MUFA | Weighted median | 46 | 0.367 | 1.088 | 0.906 | 1.308 |
| MUFA | Simple mode | 46 | 0.270 | 1.190 | 0.877 | 1.614 |
| MUFA | Weighted mode | 46 | 0.264 | 1.113 | 0.925 | 1.340 |
| MUFA_pct | MR Egger | 43 | 0.371 | 0.925 | 0.780 | 1.096 |
| MUFA_pct | Weighted median | 43 | 0.136 | 0.886 | 0.756 | 1.039 |
| MUFA_pct | Simple mode | 43 | 0.721 | 1.059 | 0.774 | 1.449 |
| MUFA_pct | Weighted mode | 43 | 0.295 | 0.918 | 0.784 | 1.075 |
| MUFA_pct | Inverse variance weighted | 43 | 0.646 | 0.975 | 0.873 | 1.088 |
| non_HDL_C | Inverse variance weighted | 35 | 0.012 | 1.200 | 1.041 | 1.384 |
| non_HDL_C | MR Egger | 35 | 0.039 | 1.319 | 1.024 | 1.699 |
| non_HDL_C | Weighted median | 35 | 0.015 | 1.275 | 1.049 | 1.550 |
| non_HDL_C | Simple mode | 35 | 0.405 | 1.164 | 0.818 | 1.656 |
| non_HDL_C | Weighted mode | 35 | 0.020 | 1.318 | 1.056 | 1.646 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Omega_3 | Inverse variance weighted | 37 | 0.031 | 1.091 | 1.008 | 1.180 |
| Omega_3 | MR Egger | 37 | 0.103 | 1.096 | 0.984 | 1.220 |
| Omega_3 | Weighted median | 37 | 0.077 | 1.095 | 0.990 | 1.211 |
| Omega_3 | Simple mode | 37 | 0.587 | 1.062 | 0.857 | 1.315 |
| Omega_3 | Weighted mode | 37 | 0.088 | 1.091 | 0.990 | 1.202 |
| Omega_3_pct | Inverse variance weighted | 26 | 0.073 | 1.075 | 0.993 | 1.164 |
| Omega_3_pct | MR Egger | 26 | 0.095 | 1.091 | 0.989 | 1.203 |
| Omega_3_pct | Weighted median | 26 | 0.064 | 1.082 | 0.995 | 1.176 |
| Omega_3_pct | Simple mode | 26 | 0.755 | 1.053 | 0.766 | 1.447 |
| Omega_3_pct | Weighted mode | 26 | 0.103 | 1.080 | 0.988 | 1.182 |
| Omega_6 | MR Egger | 42 | 0.426 | 1.137 | 0.831 | 1.555 |
| Omega_6 | Weighted median | 42 | 0.407 | 1.075 | 0.906 | 1.277 |
| Omega_6 | Simple mode | 42 | 0.467 | 1.114 | 0.835 | 1.486 |
| Omega_6 | Weighted mode | 42 | 0.476 | 1.076 | 0.881 | 1.314 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | P-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Omega_6 | Inverse variance weighted | 42 | 0.631 | 1.038 | 0.893 | 1.206 |
| Omega_6_by_Omega_3 | Inverse variance weighted | 27 | 0.044 | 0.920 | 0.849 | 0.998 |
| Omega_6_by_Omega_3 | MR Egger | 27 | 0.136 | 0.922 | 0.831 | 1.023 |
| Omega_6_by_Omega_3 | Weighted median | 27 | 0.070 | 0.921 | 0.842 | 1.007 |
| Omega_6_by_Omega_3 | Simple mode | 27 | 0.567 | 0.922 | 0.700 | 1.214 |
| Omega_6_by_Omega_3 | Weighted mode | 27 | 0.089 | 0.922 | 0.842 | 1.009 |
| Omega_6_pct | Inverse variance weighted | 36 | 0.183 | 0.898 | 0.768 | 1.052 |
| Omega_6_pct | MR Egger | 36 | 0.246 | 0.842 | 0.633 | 1.120 |
| Omega_6_pct | Weighted median | 36 | 0.461 | 0.916 | 0.725 | 1.157 |
| Omega_6_pct | Simple mode | 36 | 0.386 | 0.864 | 0.624 | 1.197 |
| Omega_6_pct | Weighted mode | 36 | 0.427 | 0.907 | 0.716 | 1.150 |
| Phe | 5 | 0.317 | 1.381 | 0.815 | 2.342 |  |
| Phe | MR Egger | 5 | 0.359 | 1.135 | 0.866 | 1.487 |
| Phe | Weighted median | 0.880 | 1.034 | 0.687 | 1.557 |  |


| Exposure (metabolic trait) | Method | Number of SNPs |  | p-value | Odds ratio | Odds ratio lower 95\% CI | Odds ratio upper 95\% CI |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Phe | Weighted mode |  | 5 | 0.387 | 1.150 | 0.868 | 1.523 |
| Phe | Inverse variance weighted |  | 5 | 0.410 | 1.110 | 0.866 | 1.421 |
| Phosphatidylc | MR Egger |  | 43 | 0.949 | 1.006 | 0.840 | 1.205 |
| Phosphatidylc | Weighted median |  | 43 | 0.537 | 1.052 | 0.895 | 1.236 |
| Phosphatidylc | Simple mode |  | 43 | 0.590 | 1.071 | 0.836 | 1.371 |
| Phosphatidylc | Weighted mode |  | 43 | 0.611 | 1.040 | 0.895 | 1.208 |
| Phosphatidylc | Inverse variance weighted |  | 43 | 0.424 | 1.045 | 0.939 | 1.162 |
| Phosphoglyc | Inverse variance weighted |  | 43 | 0.202 | 1.073 | 0.963 | 1.196 |
| Phosphoglyc | MR Egger |  | 43 | 0.687 | 1.041 | 0.858 | 1.262 |
| Phosphoglyc | Weighted median |  | 43 | 0.517 | 1.055 | 0.896 | 1.243 |
| Phosphoglyc | Simple mode |  | 43 | 0.667 | 1.056 | 0.826 | 1.350 |
| Phosphoglyc | Weighted mode |  | 43 | 0.535 | 1.056 | 0.891 | 1.252 |
| PUFA | Inverse variance weighted |  | 44 | 0.114 | 1.094 | 0.979 | 1.222 |
| PUFA | MR Egger |  | 44 | 0.131 | 1.197 | 0.952 | 1.505 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| PUFA | Weighted median | 44 | 0.306 | 1.081 | 0.931 | 1.255 |
| PUFA | Simple mode | 44 | 0.257 | 1.151 | 0.905 | 1.465 |
| PUFA | Weighted mode | 44 | 0.204 | 1.112 | 0.946 | 1.306 |
| PUFA_by_MUFA | MR Egger | 38 | 0.391 | 1.089 | 0.898 | 1.321 |
| PUFA_by_MUFA | Weighted median | 38 | 0.631 | 1.042 | 0.880 | 1.234 |
| PUFA_by_MUFA | Simple mode | 38 | 0.530 | 0.900 | 0.649 | 1.247 |
| PUFA_by_MUFA | Weighted mode | 38 | 0.474 | 1.063 | 0.900 | 1.257 |
| PUFA_by_MUFA | Inverse variance weighted | 38 | 0.909 | 1.007 | 0.893 | 1.135 |
| PUFA_pct | MR Egger | 35 | 0.562 | 1.078 | 0.838 | 1.387 |
| PUFA_pct | Weighted median | 35 | 0.730 | 0.962 | 0.774 | 1.197 |
| PUFA_pct | Simple mode | 35 | 0.945 | 0.987 | 0.675 | 1.442 |
| PUFA_pct | Weighted mode | 35 | 0.699 | 1.046 | 0.833 | 1.314 |
| PUFA_pct | Inverse variance weighted | 35 | 0.649 | 0.967 | 0.839 | 1.115 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Pyruvate | Weighted median | 15 | 0.477 | 1.107 | 0.837 | 1.464 |
| Pyruvate | Simple mode | 15 | 0.465 | 1.217 | 0.729 | 2.030 |
| Pyruvate | Weighted mode | 15 | 0.312 | 1.154 | 0.883 | 1.509 |
| Pyruvate | Inverse variance weighted | 15 | 0.695 | 0.948 | 0.728 | 1.236 |
| Remnant_C | Inverse variance weighted | 40 | 0.081 | 1.131 | 0.985 | 1.300 |
| Remnant_C | MR Egger | 40 | 0.066 | 1.279 | 0.992 | 1.650 |
| Remnant_C | Weighted median | 40 | 0.205 | 1.129 | 0.936 | 1.363 |
| Remnant_C | Simple mode | 40 | 0.290 | 1.214 | 0.852 | 1.730 |
| Remnant_C | Weighted mode | 40 | 0.211 | 1.166 | 0.921 | 1.476 |
| S_HDL_C | MR Egger | 36 | 0.959 | 1.007 | 0.778 | 1.304 |
| S_HDL_C | Weighted median | 36 | 0.729 | 1.033 | 0.860 | 1.241 |
| S_HDL_C | Simple mode | 36 | 0.706 | 1.051 | 0.812 | 1.362 |
| S_HDL_C | Weighted mode | 36 | 0.565 | 1.051 | 0.888 | 1.245 |
| S_HDL_C | Inverse variance weighted | 36 | 0.548 | 1.043 | 0.910 | 1.196 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| S_HDL_C_pct | MR Egger | 42 | 0.674 | 1.052 | 0.831 | 1.332 |
| S_HDL_C_pct | Weighted median | 42 | 0.568 | 0.951 | 0.799 | 1.131 |
| S_HDL_C_pct | Simple mode | 42 | 0.604 | 0.922 | 0.680 | 1.251 |
| S_HDL_C_pct | Weighted mode | 42 | 0.553 | 0.949 | 0.800 | 1.126 |
| S_HDL_C_pct | Inverse variance weighted | 42 | 0.688 | 1.029 | 0.895 | 1.183 |
| S_HDL_CE | MR Egger | 34 | 0.957 | 0.994 | 0.800 | 1.235 |
| S_HDL_CE | Weighted median | 34 | 0.821 | 1.019 | 0.868 | 1.195 |
| S_HDL_CE | Simple mode | 34 | 0.671 | 1.053 | 0.831 | 1.336 |
| S_HDL_CE | Weighted mode | 34 | 0.563 | 1.045 | 0.902 | 1.210 |
| S_HDL_CE | Inverse variance weighted | 34 | 0.791 | 1.016 | 0.901 | 1.146 |
| S_HDL_CE_pct | MR Egger | 43 | 0.796 | 1.021 | 0.873 | 1.194 |
| S_HDL_CE_pct | Weighted median | 43 | 0.582 | 0.962 | 0.838 | 1.104 |
| S_HDL_CE_pct | Simple mode | 43 | 0.318 | 0.859 | 0.641 | 1.153 |
| S_HDL_CE_pct | Weighted mode | 43 | 0.592 | 0.967 | 0.856 | 1.092 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| S_HDL_CE_pct | Inverse variance weighted | 43 | 0.459 | 0.960 | 0.860 | 1.070 |
| S_HDL_FC | MR Egger | 37 | 0.819 | 1.034 | 0.776 | 1.379 |
| S_HDL_FC | Weighted median | 37 | 0.440 | 1.075 | 0.895 | 1.290 |
| S_HDL_FC | Simple mode | 37 | 0.604 | 1.070 | 0.830 | 1.380 |
| S_HDL_FC | Weighted mode | 37 | 0.496 | 1.070 | 0.882 | 1.299 |
| S_HDL_FC | Inverse variance weighted | 37 | 0.453 | 1.057 | 0.915 | 1.220 |
| S_HDL_FC_pct | Inverse variance weighted | 55 | 0.152 | 1.073 | 0.975 | 1.180 |
| S_HDL_FC_pct | MR Egger | 55 | 0.926 | 0.993 | 0.864 | 1.142 |
| S_HDL_FC_pct | Weighted median | 55 | 0.386 | 0.946 | 0.834 | 1.073 |
| S_HDL_FC_pct | Simple mode | 55 | 0.551 | 1.098 | 0.809 | 1.490 |
| S_HDL_FC_pct | Weighted mode | 55 | 0.540 | 0.962 | 0.852 | 1.087 |
| S_HDL_L | MR Egger | 37 | 0.986 | 1.002 | 0.791 | 1.270 |
| S_HDL_L | Weighted median | 37 | 0.501 | 1.059 | 0.896 | 1.251 |
| S_HDL_L | Simple mode | 37 | 0.482 | 1.100 | 0.846 | 1.431 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| S_HDL_L | Weighted mode | 37 | 0.461 | 1.063 | 0.905 | 1.250 |
| S_HDL_L | Inverse variance weighted | 37 | 0.675 | 1.028 | 0.903 | 1.170 |
| S_HDL_P | MR Egger | 31 | 0.556 | 1.075 | 0.848 | 1.362 |
| S_HDL_P | Weighted median | 31 | 0.770 | 1.026 | 0.863 | 1.221 |
| S_HDL_P | Simple mode | 31 | 0.730 | 1.049 | 0.800 | 1.377 |
| S_HDL_P | Weighted mode | 31 | 0.522 | 1.058 | 0.893 | 1.253 |
| S_HDL_P | Inverse variance weighted | 31 | 0.869 | 0.989 | 0.866 | 1.129 |
| S_HDL_PL | MR Egger | 43 | 0.536 | 1.088 | 0.835 | 1.417 |
| S_HDL_PL | Weighted median | 43 | 0.559 | 1.051 | 0.889 | 1.244 |
| S_HDL_PL | Simple mode | 43 | 0.821 | 1.033 | 0.780 | 1.369 |
| S_HDL_PL | Weighted mode | 43 | 0.501 | 1.054 | 0.905 | 1.228 |
| S_HDL_PL | Inverse variance weighted | 43 | 0.996 | 1.000 | 0.865 | 1.157 |
| S_HDL_PL_pct | Inverse variance weighted | 44 | 0.035 | 0.857 | 0.742 | 0.989 |
| S_HDL_PL_pct | MR Egger | 44 | 0.228 | 0.862 | 0.679 | 1.094 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| S_HDL_PL_pct | Weighted median | 44 | 0.339 | 0.918 | 0.771 | 1.094 |
| S_HDL_PL_pct | Simple mode | 44 | 0.841 | 0.968 | 0.702 | 1.333 |
| S_HDL_PL_pct | Weighted mode | 44 | 0.657 | 0.958 | 0.792 | 1.157 |
| S_HDL_TG | MR Egger | 47 | 0.510 | 1.085 | 0.853 | 1.381 |
| S_HDL_TG | Weighted median | 47 | 0.559 | 1.053 | 0.886 | 1.252 |
| S_HDL_TG | Simple mode | 47 | 0.715 | 1.053 | 0.799 | 1.389 |
| S_HDL_TG | Weighted mode | 47 | 0.442 | 1.072 | 0.899 | 1.277 |
| S_HDL_TG | Inverse variance weighted | 47 | 0.527 | 0.959 | 0.841 | 1.093 |
| S_HDL_TG_pct | MR Egger | 60 | 0.719 | 1.045 | 0.824 | 1.324 |
| S_HDL_TG_pct | Weighted median | 60 | 0.433 | 1.066 | 0.909 | 1.250 |
| S_HDL_TG_pct | Simple mode | 60 | 0.943 | 1.009 | 0.780 | 1.307 |
| S_HDL_TG_pct | Weighted mode | 60 | 0.560 | 1.057 | 0.879 | 1.270 |
| S_HDL_TG_pct | Inverse variance weighted | 60 | 0.843 | 0.987 | 0.871 | 1.120 |
| S_LDL_C | Inverse variance weighted | 38 | 0.070 | 1.133 | 0.990 | 1.296 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| S_LDL_C | MR Egger | 38 | 0.041 | 1.287 | 1.019 | 1.626 |
| S_LDL_C | Weighted median | 38 | 0.129 | 1.164 | 0.957 | 1.415 |
| S_LDL_C | Simple mode | 38 | 0.745 | 1.061 | 0.744 | 1.514 |
| S_LDL_C | Weighted mode | 38 | 0.272 | 1.152 | 0.898 | 1.477 |
| S_LDL_C_pct | Inverse variance weighted | 27 | 0.234 | 1.101 | 0.940 | 1.289 |
| S_LDL_C_pct | MR Egger | 27 | 0.386 | 1.161 | 0.834 | 1.616 |
| S_LDL_C_pct | Weighted median | 27 | 0.214 | 1.147 | 0.924 | 1.423 |
| S_LDL_C_pct | Simple mode | 27 | 0.942 | 0.988 | 0.718 | 1.360 |
| S_LDL_C_pct | Weighted mode | 27 | 0.469 | 1.089 | 0.867 | 1.369 |
| S_LDL_CE | Inverse variance weighted | 36 | 0.116 | 1.120 | 0.973 | 1.290 |
| S_LDL_CE | MR Egger | 36 | 0.062 | 1.282 | 0.996 | 1.648 |
| S_LDL_CE | Weighted median | 36 | 0.410 | 1.089 | 0.888 | 1.336 |
| S_LDL_CE | Simple mode | 36 | 0.780 | 1.054 | 0.730 | 1.523 |
| S_LDL_CE | Weighted mode | 36 | 0.473 | 1.109 | 0.838 | 1.467 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| S_LDL_CE_pct | MR Egger | 34 | 0.141 | 1.229 | 0.940 | 1.606 |
| S_LDL_CE_pct | Weighted median | 34 | 0.562 | 1.054 | 0.883 | 1.257 |
| S_LDL_CE_pct | Simple mode | 34 | 0.651 | 1.064 | 0.815 | 1.390 |
| S_LDL_CE_pct | Weighted mode | 34 | 0.316 | 1.088 | 0.925 | 1.281 |
| S_LDL_CE_pct | Inverse variance weighted | 34 | 0.974 | 1.002 | 0.866 | 1.160 |
| S_LDL_FC | Inverse variance weighted | 41 | 0.007 | 1.189 | 1.048 | 1.349 |
| S_LDL_FC | MR Egger | 41 | 0.005 | 1.360 | 1.108 | 1.669 |
| S_LDL_FC | Weighted median | 41 | 0.002 | 1.327 | 1.110 | 1.587 |
| S_LDL_FC | Simple mode | 41 | 0.163 | 1.282 | 0.910 | 1.805 |
| S_LDL_FC | Weighted mode | 41 | 0.002 | 1.370 | 1.133 | 1.657 |
| S_LDL_FC_pct | MR Egger | 53 | 0.502 | 0.930 | 0.752 | 1.149 |
| S_LDL_FC_pct | Weighted median | 53 | 0.372 | 0.927 | 0.785 | 1.095 |
| S_LDL_FC_pct | Simple mode | 53 | 0.485 | 0.902 | 0.678 | 1.202 |
| S_LDL_FC_pct | Weighted mode | 53 | 0.405 | 0.927 | 0.776 | 1.107 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| S_LDL_FC_pct | Inverse variance weighted | 53 | 0.958 | 0.997 | 0.881 | 1.128 |
| S_LDL_L | Inverse variance weighted | 37 | 0.086 | 1.131 | 0.983 | 1.302 |
| S_LDL_L | MR Egger | 37 | 0.142 | 1.211 | 0.944 | 1.556 |
| S_LDL_L | Weighted median | 37 | 0.095 | 1.167 | 0.973 | 1.398 |
| S_LDL_L | Simple mode | 37 | 0.781 | 1.054 | 0.729 | 1.525 |
| S_LDL_L | Weighted mode | 37 | 0.298 | 1.139 | 0.894 | 1.451 |
| S_LDL_P | Inverse variance weighted | 40 | 0.034 | 1.156 | 1.011 | 1.323 |
| S_LDL_P | MR Egger | 0.103 | 1.232 | 0.965 | 1.573 |  |
| S_LDL_P | Weighted median | 40 | 0.118 | 1.168 | 0.962 | 1.418 |
| S_LDL_P | Simple mode | 40 | 0.413 | 1.166 | 0.810 | 1.679 |
| S_LDL_P | Weighted mode | 40 | 0.260 | 1.166 | 0.896 | 1.518 |
| S_LDL_PL | Inverse variance weighted | 36 | 0.020 | 1.175 | 1.026 | 1.346 |
| S_LDL_PL | MR Egger | 36 | 0.038 | 1.284 | 1.024 | 1.609 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| S_LDL_PL | Simple mode | 36 | 0.316 | 1.161 | 0.870 | 1.550 |
| S_LDL_PL | Weighted mode | 36 | 0.003 | 1.322 | 1.115 | 1.568 |
| S_LDL_PL_pct | MR Egger | 35 | 0.134 | 0.830 | 0.654 | 1.053 |
| S_LDL_PL_pct | Weighted median | 35 | 0.637 | 0.959 | 0.805 | 1.141 |
| S_LDL_PL_pct | Simple mode | 35 | 0.846 | 0.971 | 0.726 | 1.300 |
| S_LDL_PL_pct | Weighted mode | 35 | 0.443 | 0.929 | 0.770 | 1.120 |
| S_LDL_PL_pct | Inverse variance weighted | 35 | 0.646 | 0.971 | 0.856 | 1.101 |
| S_LDL_TG | Inverse variance weighted | 47 | 0.312 | 1.061 | 0.946 | 1.189 |
| S_LDL_TG | MR Egger | 47 | 0.426 | 1.088 | 0.886 | 1.335 |
| S_LDL_TG | Weighted median | 47 | 0.370 | 1.076 | 0.917 | 1.262 |
| S_LDL_TG | Simple mode | 47 | 0.532 | 1.080 | 0.849 | 1.374 |
| S_LDL_TG | Weighted mode | 47 | 0.363 | 1.080 | 0.916 | 1.274 |
| S_LDL_TG_pct | MR Egger | 51 | 0.876 | 0.985 | 0.817 | 1.188 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| S_LDL_TG_pct | Simple mode | 51 | 0.923 | 1.015 | 0.756 | 1.361 |
| S_LDL_TG_pct | Weighted mode | 51 | 0.945 | 0.994 | 0.839 | 1.178 |
| S_LDL_TG_pct | Inverse variance weighted | 51 | 0.695 | 0.979 | 0.879 | 1.090 |
| S_VLDL_C | Inverse variance weighted | 45 | 0.045 | 1.138 | 1.003 | 1.291 |
| S_VLDL_C | MR Egger | 45 | 0.347 | 1.121 | 0.886 | 1.420 |
| S_VLDL_C | Weighted median | 45 | 0.177 | 1.126 | 0.948 | 1.337 |
| S_VLDL_C | Simple mode | 45 | 0.464 | 1.109 | 0.843 | 1.458 |
| S_VLDL_C | Weighted mode | 45 | 0.439 | 1.077 | 0.894 | 1.299 |
| S_VLDL_C_pct | Inverse variance weighted | 41 | 0.235 | 1.084 | 0.949 | 1.238 |
| S_VLDL_C_pct | MR Egger | 41 | 0.358 | 1.118 | 0.884 | 1.413 |
| S_VLDL_C_pct | Weighted median | 41 | 0.489 | 1.068 | 0.886 | 1.287 |
| S_VLDL_C_pct | Simple mode | 41 | 0.998 | 0.999 | 0.687 | 1.455 |
| S_VLDL_C_pct | Weighted mode | 41 | 0.748 | 1.040 | 0.819 | 1.323 |
| S_VLDL_CE | Inverse variance weighted | 49 | 0.122 | 1.103 | 0.974 | 1.249 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| S_VLDL_CE | MR Egger | 49 | 0.373 | 1.109 | 0.886 | 1.389 |
| S_VLDL_CE | Weighted median | 49 | 0.215 | 1.107 | 0.943 | 1.299 |
| S_VLDL_CE | Simple mode | 49 | 0.453 | 1.116 | 0.840 | 1.483 |
| S_VLDL_CE | Weighted mode | 49 | 0.473 | 1.064 | 0.899 | 1.259 |
| S_VLDL_CE_pct | Inverse variance weighted | 47 | 0.412 | 1.054 | 0.929 | 1.195 |
| S_VLDL_CE_pct | MR Egger | 47 | 0.561 | 1.062 | 0.868 | 1.300 |
| S_VLDL_CE_pct | Weighted median | 47 | 0.411 | 1.075 | 0.905 | 1.276 |
| S_VLDL_CE_pct | Simple mode | 47 | 0.925 | 1.018 | 0.706 | 1.467 |
| S_VLDL_CE_pct | Weighted mode | 47 | 0.485 | 1.058 | 0.904 | 1.239 |
| S_VLDL_FC | Inverse variance weighted | 40 | 0.034 | 1.157 | 1.011 | 1.326 |
| S_VLDL_FC | MR Egger | 40 | 0.212 | 1.177 | 0.915 | 1.513 |
| S_VLDL_FC | Weighted median | 40 | 0.184 | 1.133 | 0.942 | 1.362 |
| S_VLDL_FC | Simple mode | 40 | 0.837 | 1.038 | 0.731 | 1.473 |
| S_VLDL_FC | Weighted mode | 40 | 0.402 | 1.098 | 0.885 | 1.363 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| S_VLDL_FC_pct | Inverse variance weighted | 52 | 0.082 | 1.109 | 0.987 | 1.246 |
| S_VLDL_FC_pct | MR Egger | 52 | 0.144 | 1.195 | 0.944 | 1.512 |
| S_VLDL_FC_pct | Weighted median | 52 | 0.868 | 1.015 | 0.848 | 1.217 |
| S_VLDL_FC_pct | Simple mode | 52 | 0.896 | 0.976 | 0.682 | 1.397 |
| S_VLDL_FC_pct | Weighted mode | 52 | 0.908 | 0.983 | 0.739 | 1.307 |
| S_VLDL_L | Inverse variance weighted | 49 | 0.306 | 1.062 | 0.947 | 1.191 |
| S_VLDL_L | MR Egger | 49 | 0.876 | 0.982 | 0.788 | 1.225 |
| S_VLDL_L | Weighted median | 49 | 0.389 | 1.069 | 0.918 | 1.245 |
| S_VLDL_L | Simple mode | 49 | 0.425 | 1.101 | 0.870 | 1.394 |
| S_VLDL_L | Weighted mode | 49 | 0.462 | 1.069 | 0.896 | 1.276 |
| S_VLDL_P | Inverse variance weighted | 48 | 0.374 | 1.052 | 0.941 | 1.176 |
| S_VLDL_P | MR Egger | 48 | 0.617 | 0.947 | 0.766 | 1.171 |
| S_VLDL_P | Weighted median | 48 | 0.508 | 1.049 | 0.911 | 1.207 |
| S_VLDL_P | Simple mode | 48 | 0.445 | 1.098 | 0.865 | 1.394 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| S_VLDL_P | Weighted mode | 48 | 0.598 | 1.045 | 0.887 | 1.231 |
| S_VLDL_PL | Inverse variance weighted | 51 | 0.112 | 1.120 | 0.974 | 1.287 |
| S_VLDL_PL | MR Egger | 51 | 0.659 | 1.062 | 0.813 | 1.388 |
| S_VLDL_PL | Weighted median | 51 | 0.271 | 1.100 | 0.929 | 1.302 |
| S_VLDL_PL | Simple mode | 51 | 0.652 | 1.072 | 0.794 | 1.448 |
| S_VLDL_PL | Weighted mode | 51 | 0.613 | 1.052 | 0.866 | 1.277 |
| S_VLDL_PL_pct | Inverse variance weighted | 51 | 0.100 | 1.106 | 0.981 | 1.247 |
| S_VLDL_PL_pct | MR Egger | 51 | 0.092 | 1.233 | 0.971 | 1.566 |
| S_VLDL_PL_pct | Weighted median | 51 | 0.719 | 1.034 | 0.863 | 1.239 |
| S_VLDL_PL_pct | Simple mode | 51 | 0.917 | 0.981 | 0.689 | 1.397 |
| S_VLDL_PL_pct | Weighted mode | 51 | 0.938 | 0.989 | 0.744 | 1.314 |
| S_VLDL_TG | Inverse variance weighted | 57 | 0.497 | 1.038 | 0.932 | 1.155 |
| S_VLDL_TG | MR Egger | 57 | 0.194 | 1.133 | 0.940 | 1.366 |
| S_VLDL_TG | Weighted median | 57 | 0.419 | 1.067 | 0.912 | 1.247 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| S_VLDL_TG | Simple mode | 57 | 0.741 | 1.045 | 0.805 | 1.357 |
| S_VLDL_TG | Weighted mode | 57 | 0.355 | 1.083 | 0.916 | 1.282 |
| S_VLDL_TG_pct | Inverse variance weighted | 44 | 0.171 | 0.905 | 0.785 | 1.044 |
| S_VLDL_TG_pct | MR Egger | 44 | 0.197 | 0.841 | 0.649 | 1.089 |
| S_VLDL_TG_pct | Weighted median | 44 | 0.476 | 0.934 | 0.775 | 1.126 |
| S_VLDL_TG_pct | Simple mode | 44 | 0.916 | 0.980 | 0.673 | 1.427 |
| S_VLDL_TG_pct | Weighted mode | 44 | 0.532 | 0.917 | 0.699 | 1.202 |
| SFA | Inverse variance weighted | 40 | 0.156 | 1.105 | 0.962 | 1.270 |
| SFA | MR Egger | 40 | 0.618 | 1.066 | 0.830 | 1.369 |
| SFA | Weighted median | 40 | 0.366 | 1.087 | 0.907 | 1.303 |
| SFA | Simple mode | 40 | 0.348 | 1.155 | 0.858 | 1.553 |
| SFA | Weighted mode | 40 | 0.351 | 1.094 | 0.908 | 1.318 |
| SFA_pct | 20 | 0.728 | 0.852 | 0.351 | 2.067 |  |
| SFA_pct | MR Egger | Weighted median | 0.840 | 1.034 | 0.748 | 1.429 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| SFA_pct | Simple mode | 20 | 0.717 | 0.916 | 0.573 | 1.464 |
| SFA_pct | Weighted mode | 20 | 0.880 | 0.968 | 0.637 | 1.470 |
| SFA_pct | Inverse variance weighted | 20 | 0.796 | 1.038 | 0.780 | 1.382 |
| Sphingomyelins | Inverse variance weighted | 47 | 0.011 | 1.170 | 1.037 | 1.321 |
| Sphingomyelins | MR Egger | 47 | 0.003 | 1.431 | 1.144 | 1.791 |
| Sphingomyelins | Weighted median | 47 | 0.323 | 1.092 | 0.917 | 1.302 |
| Sphingomyelins | Simple mode | 47 | 0.831 | 1.037 | 0.745 | 1.444 |
| Sphingomyelins | Weighted mode | 47 | 0.203 | 1.145 | 0.932 | 1.405 |
| TG_by_PG | MR Egger | 60 | 0.357 | 1.099 | 0.900 | 1.343 |
| TG_by_PG | Weighted median | 60 | 0.467 | 1.064 | 0.900 | 1.258 |
| TG_by_PG | Simple mode | 60 | 0.690 | 1.061 | 0.795 | 1.415 |
| TG_by_PG | Weighted mode | 60 | 0.553 | 1.061 | 0.874 | 1.288 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | P-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Total_BCAA | MR Egger | 11 | 0.038 | 2.055 | 1.149 | 3.676 |
| Total_BCAA | Weighted median | 11 | 0.165 | 1.236 | 0.916 | 1.668 |
| Total_BCAA | Simple mode | 11 | 0.502 | 1.210 | 0.708 | 2.069 |
| Total_BCAA | Weighted mode | 11 | 0.234 | 1.230 | 0.893 | 1.694 |
| Total_C | Inverse variance weighted | 45 | 0.022 | 1.160 | 1.021 | 1.317 |
| Total_C | MR Egger | 45 | 0.005 | 1.435 | 1.130 | 1.821 |
| Total_C | Weighted median | 45 | 0.226 | 1.120 | 0.932 | 1.345 |
| Total_C | Simple mode | 0.711 | 1.071 | 0.748 | 1.533 |  |
| Total_C | Weighted mode | 45 | 0.209 | 1.170 | 0.919 | 1.489 |
| Total_CE | Inverse variance weighted | 47 | 0.036 | 1.149 | 1.009 | 1.308 |
| Total_CE | MR Egger | 47 | 0.001 | 1.543 | 1.215 | 1.958 |
| Total_CE | Weighted median | 47 | 0.132 | 1.144 | 0.960 | 1.363 |
| Total_CE | Simple mode | 47 | 0.455 | 1.138 | 0.813 | 1.593 |
| Total_CE | Weighted mode | 0.135 | 1.215 | 0.945 | 1.561 |  |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Total_FA | Inverse variance weighted | 41 | 0.066 | 1.125 | 0.992 | 1.276 |
| Total_FA | MR Egger | 41 | 0.957 | 1.006 | 0.803 | 1.261 |
| Total_FA | Weighted median | 41 | 0.322 | 1.087 | 0.922 | 1.282 |
| Total_FA | Simple mode | 41 | 0.279 | 1.163 | 0.888 | 1.525 |
| Total_FA | Weighted mode | 41 | 0.284 | 1.098 | 0.928 | 1.300 |
| Total_FC | Inverse variance weighted | 41 | 0.025 | 1.162 | 1.019 | 1.325 |
| Total_FC | MR Egger | 41 | 0.019 | 1.356 | 1.063 | 1.731 |
| Total_FC | Weighted median | 41 | 0.139 | 1.144 | 0.957 | 1.369 |
| Total_FC | Simple mode | 0.500 | 1.123 | 0.804 | 1.568 |  |
| Total_FC | Weighted mode | 41 | 0.210 | 1.163 | 0.922 | 1.467 |
| Total_L | Inverse variance weighted | 47 | 0.161 | 1.112 | 0.959 | 1.290 |
| Total_L | MR Egger | 47 | 0.514 | 1.113 | 0.810 | 1.529 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | P-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Total_L | Weighted mode | 47 | 0.355 | 1.095 | 0.905 | 1.325 |
| Total_P | Inverse variance weighted | 46 | 0.330 | 1.084 | 0.922 | 1.274 |
| Total_P | MR Egger | 46 | 0.362 | 1.169 | 0.838 | 1.630 |
| Total_P | Weighted median | 46 | 0.609 | 1.052 | 0.867 | 1.276 |
| Total_P | Simple mode | 46 | 0.832 | 1.039 | 0.733 | 1.472 |
| Total_P | Weighted mode | 46 | 0.637 | 1.065 | 0.822 | 1.379 |
| Total_PL | Inverse variance weighted | 46 | 0.134 | 1.095 | 0.973 | 1.232 |
| Total_PL | MR Egger | 0.670 | 1.052 | 0.833 | 1.330 |  |
| Total_PL | Weighted median | 46 | 0.489 | 1.063 | 0.894 | 1.265 |
| Total_PL | Simple mode | 0.709 | 1.053 | 0.803 | 1.382 |  |
| Total_PL | Weighted mode | 46 | 0.585 | 1.053 | 0.875 | 1.269 |
| Total_TG | Inverse variance weighted | 46 | 0.350 | 1.063 | 0.936 | 1.207 |
| Total_TG | MR Egger | 46 | 0.299 | 1.125 | 0.903 | 1.402 |
| Total_TG | Weighted median | 46 | 0.345 | 1.091 | 0.911 | 1.306 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Total_TG | Simple mode | 46 | 0.600 | 1.089 | 0.794 | 1.493 |
| Total_TG | Weighted mode | 46 | 0.293 | 1.119 | 0.910 | 1.377 |
| Tyr | Inverse variance weighted | 23 | 0.135 | 0.869 | 0.723 | 1.045 |
| Tyr | MR Egger | 23 | 0.573 | 0.917 | 0.682 | 1.234 |
| Tyr | Weighted median | 23 | 0.492 | 0.918 | 0.721 | 1.171 |
| Tyr | Simple mode | 23 | 0.392 | 0.790 | 0.465 | 1.342 |
| Tyr | Weighted mode | 23 | 0.365 | 0.895 | 0.708 | 1.132 |
| Unsaturation | Inverse variance weighted | 32 | 0.132 | 1.074 | 0.979 | 1.178 |
| Unsaturation | MR Egger | 32 | 0.163 | 1.091 | 0.968 | 1.228 |
| Unsaturation | Weighted median | 32 | 0.091 | 1.091 | 0.986 | 1.206 |
| Unsaturation | Simple mode | 32 | 0.608 | 0.920 | 0.670 | 1.262 |
| Unsaturation | Weighted mode | 32 | 0.140 | 1.079 | 0.978 | 1.190 |
| Val | Inverse variance weighted | 13 | 0.012 | 1.368 | 1.072 | 1.745 |
| Val | MR Egger | 13 | 0.012 | 1.864 | 1.240 | 2.801 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Val | Weighted median | 13 | 0.160 | 1.224 | 0.923 | 1.622 |
| Val | Simple mode | 13 | 0.441 | 1.228 | 0.741 | 2.033 |
| Val | Weighted mode | 13 | 0.203 | 1.220 | 0.914 | 1.628 |
| VLDL_C | Inverse variance weighted | 42 | 0.232 | 1.093 | 0.945 | 1.265 |
| VLDL_C | MR Egger | 42 | 0.658 | 1.068 | 0.799 | 1.428 |
| VLDL_C | Weighted median | 42 | 0.192 | 1.118 | 0.945 | 1.323 |
| VLDL_C | Simple mode | 42 | 0.470 | 1.102 | 0.849 | 1.429 |
| VLDL_C | Weighted mode | 42 | 0.412 | 1.081 | 0.899 | 1.300 |
| VLDL_CE | Inverse variance weighted | 39 | 0.167 | 1.119 | 0.954 | 1.312 |
| VLDL_CE | MR Egger | 39 | 0.477 | 1.116 | 0.828 | 1.503 |
| VLDL_CE | Weighted median | 39 | 0.213 | 1.119 | 0.938 | 1.335 |
| VLDL_CE | Simple mode | 39 | 0.637 | 1.071 | 0.807 | 1.421 |
| VLDL_CE | Weighted mode | 39 | 0.367 | 1.094 | 0.902 | 1.327 |
| VLDL_FC | Inverse variance weighted | 45 | 0.061 | 1.130 | 0.995 | 1.283 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| VLDL_FC | MR Egger | 45 | 0.945 | 0.991 | 0.766 | 1.282 |
| VLDL_FC | Weighted median | 45 | 0.466 | 1.060 | 0.906 | 1.241 |
| VLDL_FC | Simple mode | 45 | 0.426 | 1.108 | 0.862 | 1.424 |
| VLDL_FC | Weighted mode | 45 | 0.461 | 1.070 | 0.896 | 1.278 |
| VLDL_L | Inverse variance weighted | 38 | 0.150 | 1.105 | 0.965 | 1.266 |
| VLDL_L | MR Egger | 38 | 0.375 | 1.121 | 0.874 | 1.438 |
| VLDL_L | Weighted median | 38 | 0.252 | 1.106 | 0.931 | 1.313 |
| VLDL_L | Simple mode | 38 | 0.343 | 1.154 | 0.861 | 1.547 |
| VLDL_L | Weighted mode | 0.376 | 1.092 | 0.901 | 1.324 |  |
| VLDL_P | Inverse variance weighted | 45 | 0.125 | 1.094 | 0.975 | 1.227 |
| VLDL_P | MR Egger | 45 | 0.912 | 0.987 | 0.785 | 1.241 |
| VLDL_P | Weighted median | 45 | 0.362 | 1.076 | 0.919 | 1.259 |
| VLDL_P | Simple mode | 45 | 0.571 | 1.069 | 0.850 | 1.344 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| VLDL_PL | Odds ratio <br> upper 95\% CI |  |  |  |  |  |
| VLDL_PL | Inverse variance weighted | 44 | 0.115 | 1.107 | 0.975 | 1.257 |
| VLDL_PL | MR Egger | 44 | 0.842 | 0.974 | 0.753 | 1.261 |
| VLDL_PL | Weighted median | 44 | 0.511 | 1.056 | 0.898 | 1.242 |
| VLDL_PL | Simple mode | 44 | 0.468 | 1.094 | 0.860 | 1.392 |
| VLDL_size | Weighted mode | 44 | 0.511 | 1.056 | 0.898 | 1.242 |
| VLDL_size | Inverse variance weighted | 53 | 0.146 | 1.083 | 0.973 | 1.205 |
| VLDL_size | MR Egger | 53 | 0.256 | 1.109 | 0.929 | 1.324 |
| VLDL_size | Weighted median | 53 | 0.395 | 1.067 | 0.919 | 1.240 |
| VLDL_size | Simple mode | 53 | 0.548 | 1.079 | 0.843 | 1.380 |
| VLDL_TG | Weighted mode | 53 | 0.346 | 1.079 | 0.923 | 1.261 |
| VLDL_TG | Inverse variance weighted | 43 | 0.253 | 1.076 | 0.949 | 1.219 |
| VLDL_TG | MR Egger | 43 | 0.205 | 1.146 | 0.931 | 1.412 |
| VLDL_TG | Weighted median | 43 | 0.333 | 1.095 | 0.911 | 1.316 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| VLDL_TG | Weighted mode | 43 | 0.298 | 1.111 | 0.913 | 1.352 |
| XL_HDL_C | MR Egger | 61 | 0.510 | 1.063 | 0.887 | 1.274 |
| XL_HDL_C | Weighted median | 61 | 0.830 | 0.984 | 0.849 | 1.140 |
| XL_HDL_C | Simple mode | 61 | 0.973 | 0.994 | 0.723 | 1.367 |
| XL_HDL_C | Weighted mode | 61 | 0.677 | 0.974 | 0.860 | 1.103 |
| XL_HDL_C | Inverse variance weighted | 61 | 0.788 | 1.016 | 0.903 | 1.144 |
| XL_HDL_C_pct | MR Egger | 37 | 0.728 | 1.027 | 0.883 | 1.195 |
| XL_HDL_C_pct | Weighted median | 37 | 0.623 | 0.969 | 0.853 | 1.100 |
| XL_HDL_C_pct | Simple mode | 37 | 0.830 | 0.970 | 0.739 | 1.274 |
| XL_HDL_C_pct | Weighted mode | 37 | 0.811 | 0.987 | 0.888 | 1.097 |
| XL_HDL_C_pct | Inverse variance weighted | 37 | 0.928 | 0.99 | 0.893 | 1.109 |
| XL_HDL_CE | MR Egger | 65 | 0.496 | 1.064 | 0.890 | 1.272 |
| XL_HDL_CE | Weighted median | 65 | 0.828 | 0.984 | 0.847 | 1.142 |
| XL_HDL_CE | Simple mode | 65 | 0.909 | 0.982 | 0.718 | 1.342 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| XL_HDL_CE | Weighted mode | 65 | 0.544 | 0.962 | 0.850 | 1.089 |
| XL_HDL_CE | Inverse variance weighted | 65 | 0.776 | 1.017 | 0.905 | 1.142 |
| XL_HDL_CE_pct | Inverse variance weighted | 29 | 0.332 | 1.065 | 0.938 | 1.208 |
| XL_HDL_CE_pct | MR Egger | 29 | 0.801 | 1.029 | 0.825 | 1.283 |
| XL_HDL_CE_pct | Weighted median | 29 | 0.947 | 1.006 | 0.836 | 1.211 |
| XL_HDL_CE_pct | Simple mode | 29 | 0.798 | 0.962 | 0.719 | 1.287 |
| XL_HDL_CE_pct | Weighted mode | 29 | 0.908 | 0.990 | 0.834 | 1.175 |
| XL_HDL_FC | MR Egger | 53 | 0.651 | 1.037 | 0.886 | 1.215 |
| XL_HDL_FC | Weighted median | 53 | 0.829 | 0.985 | 0.857 | 1.132 |
| XL_HDL_FC | Simple mode | 53 | 0.754 | 0.956 | 0.721 | 1.266 |
| XL_HDL_FC | Weighted mode | 53 | 0.582 | 0.965 | 0.849 | 1.096 |
| XL_HDL_FC | Inverse variance weighted | 53 | 0.731 | 1.018 | 0.918 | 1.130 |
| XL_HDL_FC_pct | Inverse variance weighted | 52 | 0.624 | 0.969 | 0.852 | 1.101 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| XL_HDL_FC_pct | Weighted median | 52 | 0.764 | 0.975 | 0.828 | 1.148 |
| XL_HDL_FC_pct | Simple mode | 52 | 0.696 | 1.070 | 0.762 | 1.503 |
| XL_HDL_FC_pct | Weighted mode | 52 | 0.879 | 0.988 | 0.852 | 1.147 |
| XL_HDL_L | MR Egger | 55 | 0.665 | 1.040 | 0.873 | 1.238 |
| XL_HDL_L | Weighted median | 55 | 0.677 | 1.030 | 0.895 | 1.185 |
| XL_HDL_L | Simple mode | 55 | 0.959 | 1.008 | 0.740 | 1.373 |
| XL_HDL_L | Weighted mode | 55 | 0.892 | 1.008 | 0.896 | 1.134 |
| XL_HDL_L | Inverse variance weighted | 55 | 0.792 | 1.016 | 0.904 | 1.142 |
| XL_HDL_P | Inverse variance weighted | 67 | 0.456 | 1.040 | 0.938 | 1.152 |
| XL_HDL_P | MR Egger | 67 | 0.151 | 1.120 | 0.961 | 1.306 |
| XL_HDL_P | Weighted median | 67 | 0.592 | 1.039 | 0.902 | 1.197 |
| XL_HDL_P | Simple mode | 67 | 0.754 | 1.053 | 0.763 | 1.454 |
| XL_HDL_P | Weighted mode | 67 | 0.460 | 1.044 | 0.932 | 1.168 |
| XL_HDL_PL | MR Egger | 53 | 0.771 | 1.026 | 0.862 | 1.221 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| XL_HDL_PL | Weighted median | 53 | 0.614 | 1.037 | 0.901 | 1.192 |
| XL_HDL_PL | Simple mode | 53 | 0.955 | 1.009 | 0.750 | 1.357 |
| XL_HDL_PL | Weighted mode | 53 | 0.890 | 1.009 | 0.893 | 1.139 |
| XL_HDL_PL | Inverse variance weighted | 53 | 0.894 | 1.008 | 0.898 | 1.132 |
| XL_HDL_PL_pct | MR Egger | 41 | 0.845 | 0.982 | 0.816 | 1.180 |
| XL_HDL_PL_pct | Weighted median | 41 | 0.792 | 1.019 | 0.883 | 1.176 |
| XL_HDL_PL_pct | Simple mode | 41 | 0.843 | 1.034 | 0.745 | 1.434 |
| XL_HDL_PL_pct | Weighted mode | 41 | 0.871 | 1.012 | 0.876 | 1.169 |
| XL_HDL_PL_pct | Inverse variance weighted | 41 | 0.875 | 1.009 | 0.898 | 1.135 |
| XL_HDL_TG | Inverse variance weighted | 46 | 0.110 | 1.081 | 0.983 | 1.189 |
| XL_HDL_TG | MR Egger | 46 | 0.873 | 0.989 | 0.860 | 1.136 |
| XL_HDL_TG | Weighted median | 46 | 0.573 | 1.034 | 0.920 | 1.164 |
| XL_HDL_TG | Simple mode | 46 | 0.128 | 1.212 | 0.950 | 1.545 |
| XL_HDL_TG | Weighted mode | 46 | 0.515 | 1.036 | 0.932 | 1.151 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| XL_HDL_TG_pct | MR Egger | 51 | 0.552 | 1.065 | 0.866 | 1.311 |
| XL_HDL_TG_pct | Weighted median | 51 | 0.667 | 1.035 | 0.884 | 1.213 |
| XL_HDL_TG_pct | Simple mode | 51 | 0.660 | 1.058 | 0.823 | 1.361 |
| XL_HDL_TG_pct | Weighted mode | 51 | 0.590 | 1.050 | 0.880 | 1.253 |
| XL_HDL_TG_pct | Inverse variance weighted | 51 | 0.785 | 0.984 | 0.877 | 1.104 |
| XL_VLDL_C | Inverse variance weighted | 37 | 0.372 | 1.073 | 0.920 | 1.251 |
| XL_VLDL_C | MR Egger | 37 | 0.222 | 1.195 | 0.902 | 1.584 |
| XL_VLDL_C | Weighted median | 37 | 0.148 | 1.136 | 0.956 | 1.351 |
| XL_VLDL_C | Simple mode | 37 | 0.248 | 1.168 | 0.901 | 1.514 |
| XL_VLDL_C | Weighted mode | 37 | 0.218 | 1.130 | 0.933 | 1.369 |
| XL_VLDL_C_pct | Inverse variance weighted | 51 | 0.723 | 1.028 | 0.881 | 1.199 |
| XL_VLDL_C_pct | MR Egger | 51 | 0.960 | 0.993 | 0.756 | 1.304 |
| XL_VLDL_C_pct | Weighted median | 51 | 0.740 | 0.966 | 0.788 | 1.184 |
| XL_VLDL_C_pct | Simple mode | 51 | 0.942 | 0.986 | 0.679 | 1.431 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| XL_VLDL_C_pct | Weighted mode | 51 | 0.664 | 0.952 | 0.765 | 1.186 |
| XL_VLDL_CE | Inverse variance weighted | 42 | 0.369 | 1.067 | 0.926 | 1.229 |
| XL_VLDL_CE | MR Egger | 42 | 0.767 | 1.043 | 0.793 | 1.371 |
| XL_VLDL_CE | Weighted median | 42 | 0.311 | 1.094 | 0.920 | 1.301 |
| XL_VLDL_CE | Simple mode | 42 | 0.506 | 1.098 | 0.836 | 1.440 |
| XL_VLDL_CE | Weighted mode | 42 | 0.389 | 1.087 | 0.901 | 1.311 |
| XL_VLDL_CE_pct | MR Egger | 54 | 0.994 | 1.001 | 0.778 | 1.287 |
| XL_VLDL_CE_pct | Weighted median | 54 | 0.796 | 0.975 | 0.805 | 1.180 |
| XL_VLDL_CE_pct | Simple mode | 54 | 0.919 | 0.982 | 0.692 | 1.393 |
| XL_VLDL_CE_pct | Weighted mode | 54 | 0.543 | 0.939 | 0.767 | 1.150 |
| XL_VLDL_CE_pct | Inverse variance weighted | 54 | 0.915 | 0.992 | 0.856 | 1.150 |
| XL_VLDL_FC | Inverse variance weighted | 34 | 0.079 | 1.121 | 0.987 | 1.273 |
| XL_VLDL_FC | MR Egger | 34 | 0.155 | 1.182 | 0.944 | 1.480 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| XL_VLDL_FC | Simple mode | 34 | 0.383 | 1.131 | 0.861 | 1.486 |
| XL_VLDL_FC | Weighted mode | 34 | 0.349 | 1.104 | 0.900 | 1.355 |
| XL_VLDL_FC_pct | MR Egger | 44 | 0.637 | 0.931 | 0.695 | 1.249 |
| XL_VLDL_FC_pct | Weighted median | 44 | 0.693 | 0.964 | 0.806 | 1.154 |
| XL_VLDL_FC_pct | Simple mode | 44 | 0.989 | 0.998 | 0.733 | 1.359 |
| XL_VLDL_FC_pct | Weighted mode | 44 | 0.639 | 0.951 | 0.770 | 1.173 |
| XL_VLDL_FC_pct | Inverse variance weighted | 44 | 0.802 | 1.019 | 0.877 | 1.184 |
| XL_VLDL_L | Inverse variance weighted | 41 | 0.221 | 1.080 | 0.955 | 1.221 |
| XL_VLDL_L | MR Egger | 41 | 0.165 | 1.165 | 0.943 | 1.439 |
| XL_VLDL_L | Weighted median | 41 | 0.298 | 1.099 | 0.920 | 1.313 |
| XL_VLDL_L | Simple mode | 41 | 0.476 | 1.121 | 0.822 | 1.528 |
| XL_VLDL_L | Weighted mode | 41 | 0.294 | 1.111 | 0.915 | 1.348 |
| XL_VLDL_P | Inverse variance weighted | 43 | 0.219 | 1.076 | 0.957 | 1.209 |
| XL_VLDL_P | MR Egger | 43 | 0.164 | 1.161 | 0.944 | 1.428 |

$\left.\begin{array}{|l|l|l|l|l|l|l|}\hline \begin{array}{l}\text { Exposure (metabolic } \\ \text { trait) }\end{array} & \text { Method } & \begin{array}{l}\text { Number of } \\ \text { SNPs }\end{array} & \text { p-value } & \text { Odds ratio } & \begin{array}{l}\text { Odds ratio } \\ \text { lower 95\% CI }\end{array} \\ \hline \text { XL_VLDL_P } \\ \text { upper 95\% CI }\end{array}\right]$

| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| XL_VLDL_TG | Odds ratio <br> upper 95\% CI |  |  |  |  |  |
| XL_VLDL_TG | MR Egger | Weighted median | 48 | 0.230 | 1.147 | 0.920 |
| XL_VLDL_TG | Simple mode | 48 | 0.305 | 1.098 | 0.918 | 1.329 |
| XL_VLDL_TG | Weighted mode | 48 | 0.344 | 1.161 | 0.855 | 1.577 |
| XL_VLDL_TG_pct | MR Egger | 48 | 0.296 | 1.113 | 0.913 | 1.356 |
| XL_VLDL_TG_pct | Weighted median | 42 | 0.754 | 0.958 | 0.733 | 1.252 |
| XL_VLDL_TG_pct | Simple mode | 42 | 0.825 | 1.022 | 0.844 | 1.237 |
| XL_VLDL_TG_pct | Weighted mode | 42 | 0.924 | 1.016 | 0.741 | 1.393 |
| XL_VLDL_TG_pct | Inverse variance weighted | 42 | 0.959 | 1.006 | 0.812 | 1.245 |
| XS_VLDL_C | Inverse variance weighted | 42 | 0.933 | 1.006 | 0.867 | 1.169 |
| XS_VLDL_C | MR Egger | 44 | 0.108 | 1.112 | 0.977 | 1.266 |
| XS_VLDL_C | Weighted median | 44 | 0.146 | 1.173 | 0.949 | 1.450 |
| XS_VLDL_C | Simple mode | 44 | 0.538 | 1.050 | 0.899 | 1.226 |
| XS_VLDL_C | Weighted mode | 44 | 0.089 | 1.304 | 0.967 | 1.758 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| XS_VLDL_C_pct | Inverse variance weighted | 51 | 0.114 | 1.111 | 0.975 | 1.266 |
| XS_VLDL_C_pct | MR Egger | 51 | 0.712 | 1.050 | 0.811 | 1.360 |
| XS_VLDL_C_pct | Weighted median | 51 | 0.754 | 0.971 | 0.808 | 1.167 |
| XS_VLDL_C_pct | Simple mode | 51 | 0.824 | 0.961 | 0.681 | 1.357 |
| XS_VLDL_C_pct | Weighted mode | 51 | 0.738 | 0.961 | 0.764 | 1.209 |
| XS_VLDL_CE | Inverse variance weighted | 44 | 0.062 | 1.119 | 0.995 | 1.259 |
| XS_VLDL_CE | MR Egger | 44 | 0.037 | 1.235 | 1.019 | 1.496 |
| XS_VLDL_CE | Weighted median | 44 | 0.526 | 1.053 | 0.898 | 1.234 |
| XS_VLDL_CE | Simple mode | 44 | 0.495 | 1.102 | 0.837 | 1.450 |
| XS_VLDL_CE | Weighted mode | 44 | 0.208 | 1.112 | 0.945 | 1.307 |
| XS_VLDL_CE_pct | Inverse variance weighted | 53 | 0.192 | 1.091 | 0.957 | 1.244 |
| XS_VLDL_CE_pct | MR Egger | 53 | 0.840 | 0.974 | 0.752 | 1.260 |
| XS_VLDL_CE_pct | Weighted median | 53 | 0.738 | 0.970 | 0.812 | 1.159 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| XS_VLDL_CE_pct | Weighted mode | 53 | 0.779 | 0.966 | 0.760 | 1.228 |
| XS_VLDL_FC | Inverse variance weighted | 45 | 0.267 | 1.072 | 0.948 | 1.213 |
| XS_VLDL_FC | MR Egger | 45 | 0.556 | 1.062 | 0.870 | 1.296 |
| XS_VLDL_FC | Weighted median | 45 | 0.588 | 1.039 | 0.905 | 1.193 |
| XS_VLDL_FC | Simple mode | 45 | 0.822 | 1.033 | 0.781 | 1.365 |
| XS_VLDL_FC | Weighted mode | 45 | 0.500 | 1.044 | 0.922 | 1.184 |
| XS_VLDL_FC_pct | Inverse variance weighted | 36 | 0.003 | 1.209 | 1.069 | 1.368 |
| XS_VLDL_FC_pct | MR Egger | 36 | 0.013 | 1.367 | 1.081 | 1.729 |
| XS_VLDL_FC_pct | Weighted median | 36 | 0.024 | 1.236 | 1.029 | 1.485 |
| XS_VLDL_FC_pct | Simple mode | 36 | 0.878 | 1.027 | 0.735 | 1.434 |
| XS_VLDL_FC_pct | Weighted mode | 36 | 0.171 | 1.202 | 0.929 | 1.557 |
| XS_VLDL_L | Inverse variance weighted | 50 | 0.245 | 1.078 | 0.950 | 1.224 |
| XS_VLDL_L | MR Egger | 50 | 0.703 | 1.040 | 0.852 | 1.270 |

$\left.\begin{array}{|l|l|l|l|l|l|l|}\hline \begin{array}{l}\text { Exposure (metabolic } \\ \text { trait) }\end{array} & \text { Method } & \begin{array}{l}\text { Number of } \\ \text { SNPs }\end{array} & \text { p-value } & \text { Odds ratio } & \begin{array}{l}\text { Odds ratio } \\ \text { lower 95\% CI }\end{array} \\ \hline \text { XS_VLDL_L } \\ \text { upper 95\% CI }\end{array}\right]$.

| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| XS_VLDL_PL_pct | Weighted median | 47 | 0.830 | 0.978 | 0.796 | 1.200 |
| XS_VLDL_PL_pct | Simple mode | 47 | 0.991 | 1.002 | 0.681 | 1.476 |
| XS_VLDL_PL_pct | Weighted mode | 47 | 0.928 | 0.986 | 0.720 | 1.349 |
| XS_VLDL_TG | Inverse variance weighted | 57 | 0.684 | 1.020 | 0.927 | 1.122 |
| XS_VLDL_TG | MR Egger | 57 | 0.569 | 1.043 | 0.902 | 1.206 |
| XS_VLDL_TG | Weighted median | 57 | 0.577 | 1.039 | 0.907 | 1.191 |
| XS_VLDL_TG | Simple mode | 57 | 0.781 | 1.033 | 0.820 | 1.302 |
| XS_VLDL_TG | Weighted mode | 57 | 0.508 | 1.043 | 0.922 | 1.181 |
| XS_VLDL_TG_pct | Inverse variance weighted | 53 | 0.049 | 0.890 | 0.792 | 1.000 |
| XS_VLDL_TG_pct | MR Egger | 53 | 0.333 | 0.890 | 0.705 | 1.124 |
| XS_VLDL_TG_pct | Weighted median | 53 | 0.755 | 0.972 | 0.811 | 1.164 |
| XS_VLDL_TG_pct | Simple mode | 53 | 0.966 | 1.007 | 0.727 | 1.396 |
| XS_VLDL_TG_pct | Weighted mode | 53 | 0.996 | 0.999 | 0.787 | 1.269 |

$\left.\begin{array}{|l|l|l|l|l|l|l|}\hline \begin{array}{l}\text { Exposure (metabolic } \\ \text { trait) }\end{array} & \text { Method } & \begin{array}{l}\text { Number of } \\ \text { SNPs }\end{array} & \text { p-value } & \text { Odds ratio } & \begin{array}{l}\text { Odds ratio } \\ \text { lower 95\% CI }\end{array} \\ \hline \text { XXL_VLDL_C } \\ \text { upper 95\% CI }\end{array}\right]$

| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| XXL_VLDL_CE_pct | Inverse variance weighted | 17 | 0.813 | 1.028 | 0.816 | 1.296 |
| XXL_VLDL_CE_pct | MR Egger | 17 | 0.393 | 0.857 | 0.608 | 1.208 |
| XXL_VLDL_CE_pct | Weighted median | 17 | 0.839 | 1.023 | 0.820 | 1.276 |
| XXL_VLDL_CE_pct | Simple mode | 17 | 0.390 | 1.206 | 0.796 | 1.827 |
| XXL_VLDL_CE_pct | Weighted mode | 17 | 0.776 | 1.033 | 0.828 | 1.289 |
| XXL_VLDL_FC | Inverse variance weighted | 44 | 0.090 | 1.109 | 0.984 | 1.251 |
| XXL_VLDL_FC | MR Egger | 44 | 0.105 | 1.174 | 0.971 | 1.418 |
| XXL_VLDL_FC | Weighted median | 44 | 0.393 | 1.085 | 0.900 | 1.307 |
| XXL_VLDL_FC | Simple mode | 44 | 0.227 | 1.226 | 0.885 | 1.699 |
| XXL_VLDL_FC | Weighted mode | 44 | 0.260 | 1.116 | 0.924 | 1.347 |
| XXL_VLDL_FC_pct | Inverse variance weighted | 20 | 0.362 | 0.922 | 0.773 | 1.099 |
| XXL_VLDL_FC_pct | MR Egger | 20 | 0.066 | 0.793 | 0.629 | 1.000 |
| XXL_VLDL_FC_pct | Weighted median | 20 | 0.703 | 0.961 | 0.783 | 1.180 |
| XXL_VLDL_FC_pct | Simple mode | 20 | 0.608 | 0.910 | 0.640 | 1.296 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| XXL_VLDL_FC_pct | Weighted mode | 20 | 0.397 | 0.919 | 0.759 | 1.113 |
| XXL_VLDL_L | Inverse variance weighted | 45 | 0.080 | 1.111 | 0.987 | 1.251 |
| XXL_VLDL_L | MR Egger | 45 | 0.089 | 1.176 | 0.980 | 1.410 |
| XXL_VLDL_L | Weighted median | 45 | 0.491 | 1.070 | 0.883 | 1.297 |
| XXL_VLDL_L | Simple mode | 45 | 0.474 | 1.123 | 0.820 | 1.538 |
| XXL_VLDL_L | Weighted mode | 45 | 0.412 | 1.089 | 0.890 | 1.334 |
| XXL_VLDL_P | Inverse variance weighted | 42 | 0.139 | 1.096 | 0.970 | 1.238 |
| XXL_VLDL_P | MR Egger | 42 | 0.077 | 1.193 | 0.986 | 1.443 |
| XXL_VLDL_P | Weighted median | 42 | 0.451 | 1.075 | 0.891 | 1.297 |
| XXL_VLDL_P | Simple mode | 42 | 0.211 | 1.220 | 0.898 | 1.659 |
| XXL_VLDL_P | Weighted mode | 42 | 0.342 | 1.103 | 0.903 | 1.349 |
| XXL_VLDL_PL | Inverse variance weighted | 45 | 0.040 | 1.139 | 1.006 | 1.290 |
| XXL_VLDL_PL | MR Egger | 45 | 0.209 | 1.138 | 0.933 | 1.389 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI | Odds ratio <br> upper 95\% CI |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| XXL_VLDL_PL | Simple mode | 45 | 0.266 | 1.205 | 0.871 | 1.666 |
| XXL_VLDL_PL | Weighted mode | 45 | 0.297 | 1.118 | 0.908 | 1.377 |
| XXL_VLDL_PL_pct | Inverse variance weighted | 29 | 0.391 | 0.942 | 0.821 | 1.080 |
| XXL_VLDL_PL_pct | MR Egger | 29 | 0.362 | 0.908 | 0.740 | 1.113 |
| XXL_VLDL_PL_pct | Weighted median | 29 | 0.428 | 0.937 | 0.797 | 1.101 |
| XXL_VLDL_PL_pct | Simple mode | 29 | 0.191 | 1.226 | 0.910 | 1.653 |
| XXL_VLDL_PL_pct | Weighted mode | 29 | 0.624 | 0.962 | 0.827 | 1.120 |
| XXL_VLDL_TG | Inverse variance weighted | 42 | 0.176 | 1.083 | 0.965 | 1.217 |
| XXL_VLDL_TG | MR Egger | 42 | 0.047 | 1.199 | 1.008 | 1.426 |
| XXL_VLDL_TG | Weighted median | 42 | 0.800 | 1.024 | 0.853 | 1.230 |
| XXL_VLDL_TG | Simple mode | 42 | 0.581 | 1.092 | 0.801 | 1.488 |
| XXL_VLDL_TG | Weighted mode | 42 | 0.375 | 1.092 | 0.901 | 1.323 |
| XXL_VLDL_TG_pct | Inverse variance weighted | 13 | 0.439 | 1.084 | 0.884 | 1.329 |


| Exposure (metabolic <br> trait) | Method | Number of <br> SNPs | p-value | Odds ratio | Odds ratio <br> lower 95\% CI |
| :--- | :--- | :--- | :---: | :---: | :---: |
| XXL_VLDL_TG_pct | Odds ratio <br> upper 95\% CI |  |  |  |  |
| XXL_VLDL_TG_pct | Weighted median | 13 | 0.508 | 1.066 | 0.882 |
| Simple mode | 13 | 0.850 | 0.973 | 0.735 | 1.287 |
| XXL_VLDL_TG_pct | Weighted mode | 13 | 0.631 | 1.050 | 0.865 |

Appendix B Table B 6:Odds ratios for metastses and PCa death, in the minimally adjusted model, in the ProtecT trial and PCa death in the MR analysis, in the PRACTICAL consortium.

| Metabolite | p-value | Odds ratio | Odds ratio lower $95 \% \text { CI }$ | Odds ratio upper $95 \% \text { CI }$ | model |
| :---: | :---: | :---: | :---: | :---: | :---: |
| acace | 0.745 | 1.106 | 0.603 | 2.029 | UKBB |
| acace | 0.019 | 1.345 | 1.049 | 1.724 | ProtecT |
| ace | 0.927 | 1.012 | 0.779 | 1.315 | ProtecT |
| ace | 0.807 | 1.056 | 0.683 | 1.634 | UKBB |
| ala | 0.057 | 1.240 | 0.994 | 1.547 | UKBB |
| ala | 0.372 | 0.890 | 0.689 | 1.150 | ProtecT |
| alb | 0.250 | 0.892 | 0.734 | 1.084 | UKBB |
| alb | 0.686 | 1.055 | 0.813 | 1.369 | ProtecT |
| apoa1 | 0.412 | 0.899 | 0.697 | 1.160 | ProtecT |
| apoa1 | 0.821 | 1.019 | 0.866 | 1.200 | UKBB |
| apob | 0.044 | 1.138 | 1.003 | 1.291 | UKBB |
| apob | 0.767 | 1.040 | 0.802 | 1.348 | ProtecT |
| apob_apoa1 | 0.105 | 1.133 | 0.974 | 1.317 | UKBB |
| apob_apoa1 | 0.520 | 1.089 | 0.840 | 1.411 | ProtecT |
| bohbut | 0.293 | 1.151 | 0.885 | 1.498 | ProtecT |
| bohbut | 0.144 | 1.426 | 0.885 | 2.296 | UKBB |
| cit | 0.422 | 0.899 | 0.693 | 1.166 | ProtecT |
| cit | 0.190 | 0.899 | 0.767 | 1.054 | UKBB |
| crea | 0.463 | 0.929 | 0.763 | 1.131 | UKBB |
| crea | 1.000 | 1.000 | 0.775 | 1.291 | ProtecT |


| Metabolite | p-value | Odds ratio | Odds ratio lower $95 \% \mathrm{CI}$ | Odds ratio upper 95\%CI | model |
| :---: | :---: | :---: | :---: | :---: | :---: |
| dha | 0.919 | 1.013 | 0.783 | 1.312 | ProtecT |
| dha | 0.038 | 1.107 | 1.006 | 1.219 | UKBB |
| dha_fa | 0.274 | 1.064 | 0.952 | 1.189 | UKBB |
| dha_fa | 0.642 | 0.941 | 0.727 | 1.217 | ProtecT |
| estc | 0.036 | 1.149 | 1.009 | 1.308 | UKBB |
| estc | 0.352 | 0.884 | 0.682 | 1.146 | ProtecT |
| faw3 | 0.609 | 1.070 | 0.827 | 1.384 | ProtecT |
| faw3 | 0.031 | 1.091 | 1.008 | 1.180 | UKBB |
| faw3_fa | 0.073 | 1.075 | 0.993 | 1.164 | UKBB |
| faw3_fa | 0.900 | 0.984 | 0.760 | 1.273 | ProtecT |
| faw6 | 0.942 | 1.010 | 0.779 | 1.309 | ProtecT |
| faw6_fa | 0.183 | 0.898 | 0.768 | 1.052 | UKBB |
| faw6_fa | 0.327 | 0.881 | 0.683 | 1.135 | ProtecT |
| freec | 0.431 | 0.901 | 0.695 | 1.168 | ProtecT |
| freec | 0.025 | 1.162 | 1.019 | 1.325 | UKBB |
| glc | 0.849 | 0.978 | 0.780 | 1.226 | UKBB |
| glc | 0.735 | 1.045 | 0.811 | 1.347 | ProtecT |
| gln | 0.612 | 0.935 | 0.722 | 1.212 | ProtecT |
| gln | 0.646 | 0.961 | 0.811 | 1.139 | UKBB |
| glol | 0.707 | 1.151 | 0.552 | 2.400 | ProtecT |
| gly | 0.523 | 0.919 | 0.711 | 1.190 | ProtecT |
| gly | 0.255 | 1.037 | 0.974 | 1.104 | UKBB |


| Metabolite | p-value | Odds ratio | Odds ratio lower $95 \% \text { CI }$ | Odds ratio upper $95 \% \text { CI }$ | model |
| :---: | :---: | :---: | :---: | :---: | :---: |
| gp | 0.833 | 1.013 | 0.896 | 1.146 | UKBB |
| gp | 0.031 | 1.329 | 1.026 | 1.720 | ProtecT |
| hdl2_c | 0.154 | 0.830 | 0.642 | 1.072 | ProtecT |
| hdl3_c | 0.134 | 0.823 | 0.637 | 1.062 | ProtecT |
| hdl_c | 0.377 | 1.063 | 0.928 | 1.218 | UKBB |
| hdl_c | 0.152 | 0.829 | 0.642 | 1.071 | ProtecT |
| hdl_d | 0.441 | 0.904 | 0.699 | 1.169 | ProtecT |
| hdl_d | 0.326 | 1.058 | 0.946 | 1.183 | UKBB |
| hdl_tg | 0.352 | 1.128 | 0.875 | 1.455 | ProtecT |
| hdl_tg | 0.886 | 1.008 | 0.905 | 1.122 | UKBB |
| his | 0.047 | 0.740 | 0.550 | 0.996 | UKBB |
| his | 0.580 | 1.076 | 0.830 | 1.394 | ProtecT |
| idl_c | 0.435 | 0.902 | 0.696 | 1.169 | ProtecT |
| idl_c | 0.065 | 1.118 | 0.993 | 1.259 | UKBB |
| idl_ce | 0.055 | 1.124 | 0.997 | 1.267 | UKBB |
| idl_ce | 0.615 | 0.936 | 0.722 | 1.213 | ProtecT |
| idl_fc | 0.023 | 1.145 | 1.018 | 1.288 | UKBB |
| idl_fc | 0.170 | 0.835 | 0.645 | 1.080 | ProtecT |
| idl_1 | 0.078 | 1.111 | 0.988 | 1.249 | UKBB |
| idl_1 | 0.420 | 0.898 | 0.692 | 1.166 | ProtecT |
| idl_p | 0.044 | 1.154 | 1.004 | 1.325 | UKBB |
| idl_p | 0.466 | 0.908 | 0.699 | 1.178 | ProtecT |


| Metabolite | p-value | Odds ratio | Odds ratio lower $95 \% \text { CI }$ | Odds ratio upper $95 \% \text { CI }$ | model |
| :---: | :---: | :---: | :---: | :---: | :---: |
| idl_pl | 0.243 | 0.857 | 0.661 | 1.111 | ProtecT |
| idl_pl | 0.010 | 1.156 | 1.035 | 1.292 | UKBB |
| idl_tg | 0.691 | 1.054 | 0.813 | 1.367 | ProtecT |
| idl_tg | 0.453 | 1.034 | 0.947 | 1.130 | UKBB |
| ile | 0.108 | 1.397 | 0.930 | 2.099 | UKBB |
| ile | 0.279 | 1.151 | 0.892 | 1.486 | ProtecT |
| 1_hdl_c | 0.265 | 0.864 | 0.668 | 1.118 | ProtecT |
| 1_hdl_c | 0.690 | 1.021 | 0.921 | 1.132 | UKBB |
| l_hdl_ce | 0.316 | 0.877 | 0.678 | 1.134 | ProtecT |
| l_hdl_ce | 0.622 | 1.026 | 0.926 | 1.138 | UKBB |
| 1_hdl_fc | 0.640 | 0.974 | 0.874 | 1.086 | UKBB |
| 1_hdl_fc | 0.170 | 0.835 | 0.645 | 1.080 | ProtecT |
| 1_hdl_1 | 0.272 | 0.866 | 0.669 | 1.120 | ProtecT |
| 1_hdl_1 | 0.690 | 1.021 | 0.922 | 1.130 | UKBB |
| 1_hdl_p | 0.831 | 1.012 | 0.907 | 1.129 | UKBB |
| 1_hdl_p | 0.290 | 0.870 | 0.673 | 1.126 | ProtecT |
| 1_hdl_pl | 0.687 | 1.021 | 0.922 | 1.132 | UKBB |
| 1_hdl_pl | 0.214 | 0.850 | 0.657 | 1.099 | ProtecT |
| 1_hdl_tg | 0.637 | 1.020 | 0.939 | 1.109 | UKBB |
| 1_hdl_tg | 0.989 | 1.002 | 0.778 | 1.291 | ProtecT |
| 1_ldl_c | 0.000 | 1.250 | 1.104 | 1.416 | UKBB |
| 1_1dl_c | 0.266 | 0.863 | 0.665 | 1.119 | ProtecT |


| Metabolite | p-value | Odds ratio | Odds ratio lower $95 \% \text { CI }$ | Odds ratio upper 95\%CI | model |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 1_ldl_ce | 0.309 | 0.873 | 0.673 | 1.133 | ProtecT |
| 1_ldl_ce | 0.001 | 1.235 | 1.085 | 1.406 | UKBB |
| 1_ldl_fc | 0.168 | 0.834 | 0.645 | 1.079 | ProtecT |
| 1_ldl_fc | 0.005 | 1.196 | 1.055 | 1.355 | UKBB |
| 1_1d1_1 | 0.302 | 0.872 | 0.672 | 1.131 | ProtecT |
| 1_1dl_1 | 0.003 | 1.219 | 1.069 | 1.389 | UKBB |
| 1_ldl_p | 0.334 | 0.879 | 0.678 | 1.141 | ProtecT |
| 1_ldl_p | 0.009 | 1.175 | 1.041 | 1.327 | UKBB |
| 1_ldl_pl | 0.009 | 1.197 | 1.047 | 1.369 | UKBB |
| 1_ldl_pl | 0.308 | 0.873 | 0.673 | 1.133 | ProtecT |
| 1_ldl_tg | 0.517 | 1.032 | 0.939 | 1.133 | UKBB |
| 1_ldl_tg | 0.923 | 1.013 | 0.780 | 1.315 | ProtecT |
| 1_vldl_c | 0.425 | 1.110 | 0.859 | 1.435 | ProtecT |
| 1_vldl_c | 0.349 | 1.068 | 0.930 | 1.227 | UKBB |
| 1_vldl_ce | 0.123 | 1.123 | 0.969 | 1.301 | UKBB |
| 1_vldl_ce | 0.446 | 1.105 | 0.855 | 1.430 | ProtecT |
| 1_vldl_fc | 0.406 | 1.115 | 0.863 | 1.440 | ProtecT |
| 1_vldl_ff | 0.367 | 1.070 | 0.924 | 1.240 | UKBB |
| 1_vldl_1 | 0.180 | 1.089 | 0.961 | 1.234 | UKBB |
| 1_vldl_l | 0.461 | 1.101 | 0.852 | 1.423 | ProtecT |
| 1_vldl_p | 0.237 | 1.080 | 0.951 | 1.228 | UKBB |
| 1_vldl_p | 0.458 | 1.102 | 0.853 | 1.424 | ProtecT |


| Metabolite | p-value | Odds ratio | Odds ratio lower $95 \% \text { CI }$ | Odds ratio upper 95\%CI | model |
| :---: | :---: | :---: | :---: | :---: | :---: |
| l_vldl_pl | 0.399 | 1.059 | 0.927 | 1.211 | UKBB |
| 1_vldl_pl | 0.459 | 1.102 | 0.853 | 1.423 | ProtecT |
| 1_vldl_tg | 0.397 | 1.063 | 0.923 | 1.224 | UKBB |
| 1_vldl_tg | 0.460 | 1.101 | 0.853 | 1.423 | ProtecT |
| la | 0.529 | 0.949 | 0.806 | 1.117 | UKBB |
| la | 0.859 | 1.024 | 0.790 | 1.327 | ProtecT |
| la_fa | 0.133 | 0.846 | 0.680 | 1.053 | UKBB |
| la_fa | 0.512 | 0.918 | 0.711 | 1.185 | ProtecT |
| lac | 0.522 | 1.090 | 0.837 | 1.419 | ProtecT |
| ldl_c | 0.188 | 0.840 | 0.648 | 1.089 | ProtecT |
| ldl_c | 0.007 | 1.209 | 1.054 | 1.387 | UKBB |
| ldl_d | 0.056 | 1.276 | 0.993 | 1.640 | ProtecT |
| ldl_d | 0.498 | 1.070 | 0.881 | 1.299 | UKBB |
| ldl_tg | 0.716 | 1.018 | 0.926 | 1.118 | UKBB |
| ldl_tg | 0.922 | 1.013 | 0.781 | 1.314 | ProtecT |
| leu | 0.182 | 1.190 | 0.922 | 1.537 | ProtecT |
| leu | 0.292 | 1.259 | 0.820 | 1.931 | UKBB |
| m_hdl_c | 0.362 | 0.887 | 0.686 | 1.148 | ProtecT |
| m_hdl_c | 0.848 | 1.014 | 0.879 | 1.170 | UKBB |
| m_hdl_ce | 0.797 | 0.981 | 0.846 | 1.137 | UKBB |
| m_hdl_ce | 0.383 | 0.892 | 0.689 | 1.154 | ProtecT |
| m_hdl_fc | 0.247 | 0.859 | 0.665 | 1.111 | ProtecT |


| Metabolite | p-value | Odds ratio | Odds ratio lower 95\%CI | Odds ratio upper 95\%CI | model |
| :---: | :---: | :---: | :---: | :---: | :---: |
| m_hdl_fc | 0.821 | 1.017 | 0.877 | 1.179 | UKBB |
| m_hdl_1 | 0.949 | 1.005 | 0.854 | 1.184 | UKBB |
| m_hdl_1 | 0.501 | 0.916 | 0.709 | 1.183 | ProtecT |
| m_hdl_p | 0.567 | 0.928 | 0.719 | 1.198 | ProtecT |
| m_hdl_p | 0.860 | 1.015 | 0.861 | 1.196 | UKBB |
| m_hdl_pl | 0.542 | 0.924 | 0.716 | 1.192 | ProtecT |
| m_hdl_pl | 0.866 | 0.988 | 0.856 | 1.139 | UKBB |
| m_hdl_tg | 0.783 | 1.017 | 0.902 | 1.147 | UKBB |
| m_hdl_tg | 0.284 | 1.147 | 0.892 | 1.476 | ProtecT |
| m_ldl_c | 0.034 | 1.160 | 1.012 | 1.330 | UKBB |
| m_ldl_c | 0.145 | 0.825 | 0.636 | 1.069 | ProtecT |
| m_ldl_ce | 0.084 | 1.138 | 0.983 | 1.317 | UKBB |
| m_ldl_ce | 0.133 | 0.820 | 0.632 | 1.062 | ProtecT |
| m_ldl_fc | 0.196 | 0.843 | 0.652 | 1.092 | ProtecT |
| m_ldl_fc | 0.006 | 1.193 | 1.052 | 1.353 | UKBB |
| m_ldl_1 | 0.199 | 0.843 | 0.650 | 1.094 | ProtecT |
| m_ldl_1 | 0.040 | 1.156 | 1.007 | 1.327 | UKBB |
| m_ldl_p | 0.208 | 0.846 | 0.652 | 1.098 | ProtecT |
| m_ldl_p | 0.021 | 1.186 | 1.026 | 1.371 | UKBB |
| m_ldl_pl | 0.033 | 1.176 | 1.013 | 1.366 | UKBB |
| m_ldl_pl | 0.398 | 0.894 | 0.690 | 1.159 | ProtecT |
| m_ldl_tg | 0.329 | 1.050 | 0.952 | 1.160 | UKBB |


| Metabolite | p-value | Odds ratio | Odds ratio lower 95\%CI | Odds ratio upper 95\%CI | model |
| :---: | :---: | :---: | :---: | :---: | :---: |
| m_ldl_tg | 0.601 | 0.933 | 0.720 | 1.210 | ProtecT |
| m_vldl_c | 0.282 | 1.150 | 0.892 | 1.482 | ProtecT |
| m_vldl_c | 0.039 | 1.156 | 1.007 | 1.328 | UKBB |
| m_vldl_ce | 0.033 | 1.163 | 1.013 | 1.336 | UKBB |
| m_vldl_ce | 0.335 | 1.134 | 0.878 | 1.465 | ProtecT |
| m_vldl_fc | 0.013 | 1.192 | 1.038 | 1.368 | UKBB |
| m_vldl_fc | 0.504 | 1.091 | 0.845 | 1.409 | ProtecT |
| m_vldl_1 | 0.219 | 1.120 | 0.935 | 1.342 | UKBB |
| m_vldl_1 | 0.505 | 1.091 | 0.844 | 1.410 | ProtecT |
| m_vldl_p | 0.510 | 1.090 | 0.844 | 1.408 | ProtecT |
| m_vldl_p | 0.279 | 1.107 | 0.921 | 1.329 | UKBB |
| m_vldl_pl | 0.300 | 1.143 | 0.888 | 1.472 | ProtecT |
| m_vldl_pl | 0.134 | 1.134 | 0.962 | 1.337 | UKBB |
| m_vldl_tg | 0.217 | 1.076 | 0.958 | 1.209 | UKBB |
| m_vldl_tg | 0.553 | 1.080 | 0.836 | 1.396 | ProtecT |
| mufa | 0.414 | 1.112 | 0.862 | 1.434 | ProtecT |
| mufa | 0.130 | 1.100 | 0.972 | 1.244 | UKBB |
| mufa_fa | 0.583 | 1.074 | 0.833 | 1.384 | ProtecT |
| mufa_fa | 0.646 | 0.975 | 0.873 | 1.088 | UKBB |
| pc | 0.424 | 1.045 | 0.939 | 1.162 | UKBB |
| pc | 0.916 | 1.014 | 0.786 | 1.308 | ProtecT |
| phe | 0.680 | 1.055 | 0.819 | 1.359 | ProtecT |


| Metabolite | p-value | Odds ratio | Odds ratio lower $95 \% \text { CI }$ | Odds ratio upper $95 \% \text { CI }$ | model |
| :---: | :---: | :---: | :---: | :---: | :---: |
| phe | 0.410 | 1.110 | 0.866 | 1.421 | UKBB |
| pufa | 0.114 | 1.094 | 0.979 | 1.222 | UKBB |
| pufa | 0.868 | 1.022 | 0.788 | 1.325 | ProtecT |
| pufa_fa | 0.649 | 0.967 | 0.839 | 1.115 | UKBB |
| pufa_fa | 0.350 | 0.886 | 0.686 | 1.143 | ProtecT |
| pyr | 0.312 | 0.878 | 0.683 | 1.130 | ProtecT |
| pyr | 0.695 | 0.948 | 0.728 | 1.236 | UKBB |
| remnant_c | 0.631 | 1.066 | 0.821 | 1.383 | ProtecT |
| remnant_c | 0.081 | 1.131 | 0.985 | 1.300 | UKBB |
| s_hdl_c | 0.548 | 1.043 | 0.910 | 1.196 | UKBB |
| s_hdl_c | 0.066 | 0.781 | 0.600 | 1.017 | ProtecT |
| s_hdl_ce | 0.791 | 1.016 | 0.901 | 1.146 | UKBB |
| s_hdl_ce | 0.073 | 0.786 | 0.605 | 1.023 | ProtecT |
| s_hdl_fc | 0.453 | 1.057 | 0.915 | 1.220 | UKBB |
| s_hdl_fc | 0.956 | 0.993 | 0.770 | 1.280 | ProtecT |
| s_hdl_l | 0.285 | 0.869 | 0.671 | 1.124 | ProtecT |
| s_hdl_1 | 0.675 | 1.028 | 0.903 | 1.170 | UKBB |
| s_hdl_p | 0.365 | 0.888 | 0.686 | 1.149 | ProtecT |
| s_hdl_p | 0.869 | 0.989 | 0.866 | 1.129 | UKBB |
| s_hdl_pl | 0.996 | 1.000 | 0.865 | 1.157 | UKBB |
| s_hdl_pl | 0.558 | 1.078 | 0.838 | 1.387 | ProtecT |
| s_hdl_tg | 0.295 | 1.145 | 0.888 | 1.477 | ProtecT |


| Metabolite | p-value | Odds ratio | Odds ratio lower 95\%CI | Odds ratio upper 95\%CI | model |
| :---: | :---: | :---: | :---: | :---: | :---: |
| s_hdl_tg | 0.527 | 0.959 | 0.841 | 1.093 | UKBB |
| s_ldl_c | 0.070 | 1.133 | 0.990 | 1.296 | UKBB |
| s_ldl_c | 0.119 | 0.814 | 0.628 | 1.054 | ProtecT |
| s_ldl_ce | 0.116 | 1.120 | 0.973 | 1.290 | UKBB |
| s_ldl_ce | 0.113 | 0.811 | 0.626 | 1.051 | ProtecT |
| s_ldl_fc | 0.151 | 0.828 | 0.640 | 1.071 | ProtecT |
| s_ldl_fc | 0.007 | 1.189 | 1.048 | 1.349 | UKBB |
| s_ldl_1 | 0.086 | 1.131 | 0.983 | 1.302 | UKBB |
| s_ldl_1 | 0.181 | 0.838 | 0.646 | 1.086 | ProtecT |
| s_ldl_p | 0.034 | 1.156 | 1.011 | 1.323 | UKBB |
| s_ldl_p | 0.194 | 0.842 | 0.649 | 1.092 | ProtecT |
| s_ldl_pl | 0.367 | 0.888 | 0.686 | 1.149 | ProtecT |
| s_ldl_pl | 0.020 | 1.175 | 1.026 | 1.346 | UKBB |
| s_ldl_tg | 0.312 | 1.061 | 0.946 | 1.189 | UKBB |
| s_ldl_tg | 0.920 | 0.987 | 0.764 | 1.275 | ProtecT |
| s_vldl_c | 0.045 | 1.138 | 1.003 | 1.291 | UKBB |
| s_vldl_c | 0.586 | 1.075 | 0.829 | 1.395 | ProtecT |
| s_vldl_ce | 0.122 | 1.103 | 0.974 | 1.249 | UKBB |
| s_vldl_ce | 0.726 | 1.048 | 0.807 | 1.361 | ProtecT |
| s_vldl_fc | 0.034 | 1.157 | 1.011 | 1.326 | UKBB |
| s_vldl_fc | 0.438 | 1.107 | 0.857 | 1.429 | ProtecT |
| s_vldl_1 | 0.306 | 1.062 | 0.947 | 1.191 | UKBB |


| Metabolite | p-value | Odds ratio | Odds ratio lower $95 \% \text { CI }$ | Odds ratio upper 95\%CI | model |
| :---: | :---: | :---: | :---: | :---: | :---: |
| s_vldl_1 | 0.414 | 1.112 | 0.862 | 1.435 | ProtecT |
| s_vldl_p | 0.402 | 1.115 | 0.864 | 1.437 | ProtecT |
| s_vldl_p | 0.374 | 1.052 | 0.941 | 1.176 | UKBB |
| s_vldl_pl | 0.112 | 1.120 | 0.974 | 1.287 | UKBB |
| s_vldl_pl | 0.399 | 1.116 | 0.865 | 1.440 | ProtecT |
| s_vldl_tg | 0.497 | 1.038 | 0.932 | 1.155 | UKBB |
| s_vldl_tg | 0.381 | 1.119 | 0.870 | 1.441 | ProtecT |
| serum_c | 0.383 | 0.891 | 0.687 | 1.155 | ProtecT |
| serum_tg | 0.350 | 1.063 | 0.936 | 1.207 | UKBB |
| serum_tg | 0.325 | 1.136 | 0.881 | 1.463 | ProtecT |
| sfa | 0.156 | 1.105 | 0.962 | 1.270 | UKBB |
| sfa | 0.331 | 1.135 | 0.879 | 1.464 | ProtecT |
| sfa_fa | 0.233 | 1.174 | 0.902 | 1.527 | ProtecT |
| sfa_fa | 0.796 | 1.038 | 0.780 | 1.382 | UKBB |
| sm | 0.011 | 1.170 | 1.037 | 1.321 | UKBB |
| sm | 0.915 | 0.986 | 0.762 | 1.275 | ProtecT |
| tg_pg | 0.757 | 1.018 | 0.909 | 1.140 | UKBB |
| tg_pg | 0.235 | 1.167 | 0.905 | 1.504 | ProtecT |
| totcho | 0.022 | 1.160 | 1.021 | 1.317 | UKBB |
| totcho | 0.948 | 1.008 | 0.781 | 1.303 | ProtecT |
| totfa | 0.066 | 1.125 | 0.992 | 1.276 | UKBB |
| totfa | 0.414 | 1.112 | 0.862 | 1.436 | ProtecT |


| Metabolite | p-value | Odds ratio | Odds ratio lower 95\%CI | Odds ratio upper 95\%CI | model |
| :---: | :---: | :---: | :---: | :---: | :---: |
| totpg | 0.786 | 1.036 | 0.803 | 1.336 | ProtecT |
| totpg | 0.202 | 1.073 | 0.963 | 1.196 | UKBB |
| tyr | 0.135 | 0.869 | 0.723 | 1.045 | UKBB |
| tyr | 0.601 | 0.933 | 0.721 | 1.209 | ProtecT |
| unsat | 0.212 | 0.848 | 0.655 | 1.098 | ProtecT |
| unsat | 0.132 | 1.074 | 0.979 | 1.178 | UKBB |
| val | 0.797 | 1.034 | 0.801 | 1.335 | ProtecT |
| val | 0.012 | 1.368 | 1.072 | 1.745 | UKBB |
| vldl_c | 0.275 | 1.154 | 0.892 | 1.493 | ProtecT |
| vldl_c | 0.232 | 1.093 | 0.945 | 1.265 | UKBB |
| vldl_d | 0.352 | 1.128 | 0.875 | 1.453 | ProtecT |
| vldl_d | 0.146 | 1.083 | 0.973 | 1.205 | UKBB |
| vldl_tg | 0.333 | 1.133 | 0.880 | 1.459 | ProtecT |
| vldl_tg | 0.253 | 1.076 | 0.949 | 1.219 | UKBB |
| xl_hdl_c | 0.788 | 1.016 | 0.903 | 1.144 | UKBB |
| xl_hdl_c | 0.964 | 1.006 | 0.781 | 1.296 | ProtecT |
| xl_hdl_ce | 0.976 | 1.004 | 0.779 | 1.293 | ProtecT |
| xl_hdl_ce | 0.776 | 1.017 | 0.905 | 1.142 | UKBB |
| xl_hdl_fc | 0.985 | 1.002 | 0.778 | 1.292 | ProtecT |
| xl_hdl_fc | 0.731 | 1.018 | 0.918 | 1.130 | UKBB |
| xl_hdl_l | 0.749 | 0.959 | 0.743 | 1.238 | ProtecT |
| xl_hdl_l | 0.792 | 1.016 | 0.904 | 1.142 | UKBB |


| Metabolite | p-value | Odds ratio | Odds ratio lower 95\%CI | Odds ratio upper $95 \% \text { CI }$ | model |
| :---: | :---: | :---: | :---: | :---: | :---: |
| xl_hdl_p | 0.456 | 1.040 | 0.938 | 1.152 | UKBB |
| xl_hdl_p | 0.735 | 0.957 | 0.741 | 1.235 | ProtecT |
| xl_hdl_pl | 0.412 | 0.898 | 0.695 | 1.161 | ProtecT |
| xl_hdl_pl | 0.894 | 1.008 | 0.898 | 1.132 | UKBB |
| xl_hdl_tg | 0.110 | 1.081 | 0.983 | 1.189 | UKBB |
| xl_hdl_tg | 0.356 | 1.128 | 0.873 | 1.459 | ProtecT |
| xl_vldl_c | 0.373 | 1.124 | 0.869 | 1.453 | ProtecT |
| xl_vldl_c | 0.372 | 1.073 | 0.920 | 1.251 | UKBB |
| xl_vldl_ce | 0.393 | 1.119 | 0.865 | 1.447 | ProtecT |
| xl_vldl_ce | 0.369 | 1.067 | 0.926 | 1.229 | UKBB |
| xl_vldl_fc | 0.299 | 1.146 | 0.886 | 1.481 | ProtecT |
| xl_vldl_fc | 0.079 | 1.121 | 0.987 | 1.273 | UKBB |
| xl_vldl_1 | 0.353 | 1.129 | 0.874 | 1.459 | ProtecT |
| xl_vldl_1 | 0.221 | 1.080 | 0.955 | 1.221 | UKBB |
| xl_vldl_p | 0.219 | 1.076 | 0.957 | 1.209 | UKBB |
| xl_vldl_p | 0.373 | 1.124 | 0.870 | 1.452 | ProtecT |
| xl_vldl_pl | 0.313 | 1.141 | 0.883 | 1.474 | ProtecT |
| xl_vldl_pl | 0.159 | 1.093 | 0.966 | 1.236 | UKBB |
| xl_vldl_tg | 0.361 | 1.063 | 0.932 | 1.213 | UKBB |
| xl_vldl_tg | 0.360 | 1.127 | 0.873 | 1.455 | ProtecT |
| xs_vldl_c | 0.911 | 1.015 | 0.782 | 1.318 | ProtecT |
| xs_vldl_c | 0.108 | 1.112 | 0.977 | 1.266 | UKBB |


| Metabolite | p-value | Odds ratio | Odds ratio lower $95 \% \text { CI }$ | Odds ratio upper $95 \% \text { CI }$ | model |
| :---: | :---: | :---: | :---: | :---: | :---: |
| xs_vldl_ce | 0.668 | 1.059 | 0.816 | 1.375 | ProtecT |
| xs_vldl_ce | 0.062 | 1.119 | 0.995 | 1.259 | UKBB |
| xs_vldl_fc | 0.267 | 1.072 | 0.948 | 1.213 | UKBB |
| xs_vldl_fc | 0.541 | 0.922 | 0.710 | 1.197 | ProtecT |
| xs_vldl_1 | 0.956 | 1.007 | 0.775 | 1.309 | ProtecT |
| xs_vldl_1 | 0.245 | 1.078 | 0.950 | 1.224 | UKBB |
| xs_vldl_p | 0.148 | 1.103 | 0.966 | 1.260 | UKBB |
| xs_vldl_p | 0.873 | 1.022 | 0.786 | 1.328 | ProtecT |
| xs_vldl_pl | 0.423 | 1.050 | 0.932 | 1.183 | UKBB |
| xs_vldl_pl | 0.388 | 0.891 | 0.686 | 1.157 | ProtecT |
| xs_vldl_tg | 0.684 | 1.020 | 0.927 | 1.122 | UKBB |
| xs_vldl_tg | 0.536 | 1.084 | 0.840 | 1.399 | ProtecT |
| xxl_vldl_c | 0.089 | 1.108 | 0.984 | 1.247 | UKBB |
| xxl_vldl_c | 0.276 | 1.153 | 0.892 | 1.490 | ProtecT |
| xxl_vldl_ce | 0.318 | 1.140 | 0.881 | 1.476 | ProtecT |
| xxl_vldl_ce | 0.164 | 1.085 | 0.967 | 1.217 | UKBB |
| xxl_vldl_fc | 0.090 | 1.109 | 0.984 | 1.251 | UKBB |
| xxl_vldl_fc | 0.232 | 1.169 | 0.905 | 1.509 | ProtecT |
| xxl_vldl_1 | 0.080 | 1.111 | 0.987 | 1.251 | UKBB |
| xxl_vldl_1 | 0.415 | 1.114 | 0.860 | 1.443 | ProtecT |
| xxl_vldl_p | 0.140 | 1.209 | 0.939 | 1.558 | ProtecT |
| xxl_vldl_p | 0.139 | 1.096 | 0.970 | 1.238 | UKBB |


| Metabolite | p-value | Odds ratio | Odds ratio lower <br> 95\% CI | Odds ratio upper <br> $\mathbf{9 5 \%} \mathbf{C I}$ | model |
| :--- | ---: | ---: | :--- | :--- | :--- |
| xxl_vldl_pl | 0.040 | 1.139 | 1.006 | 1.290 | UKBB |
| xxl_vldl_pl | 0.382 | 1.122 | 0.867 | 1.452 | ProtecT |
| xxl_vldl_tg | 0.176 | 1.083 | 0.965 | 1.217 | UKBB |
| xxl_vldl_tg | 0.394 | 1.119 | 0.864 | 1.450 | ProtecT |

Appendix B Figure B 1: Forest plot of hazard ratios for clinical progression and metastases or death, in the fully adjusted Cox models.

Lipoprotein subclasses


Lipoprotein particle size


Cholesterol


Glycerides and phospholipids


## Fatty acids



## Fatty acids ratios



## Glycolysis related metabolites



Branched-chain amino acids


Aromatic amino acids


Ketone bodies


## Fluid balance



## Inflammation


$\rightarrow$ Clinical Progression $\rightarrow-$ Metastases or death

Appendix B Figure B 2: Forest plot of hazard ratios for clinical progression, in the minimally and fully adjusted Cox models.

Lipoprotein subclasses

| Extremely large VLDL | 40 |
| :---: | :---: |
| Very large VLDL | $\bigcirc$ |
| Large VLDL | $\bigcirc$ |
| Medium VLDL | $\square$ |
| Small VLDL | $\bigcirc$ |
| Very Small VLDL | $\square$ |
| IDL | $\rightarrow$ |
| Large LDL | $\bigcirc$ |
| Medium LDL | $\bigcirc$ |
| Small LDL | $\bigcirc$ |
| Very large HDL | $\bigcirc$ |
| Large HDL | $\bigcirc$ |
| Medium HDL | $\square$ |
| Small HDL | $\square$ |

Lipoprotein particle size

| VLDL particles size | $\xrightarrow{+0-}$ |
| :---: | :---: |
| LDL particles size | $\bigcirc$ |
| HDL particles size | －0 |
| Cholesterol |  |
| Total C | $\square$ |
| VLDL C | $\xrightarrow[\square-]{0}$ |
| Remnant C | $\square$ |
| LDL C | $\cdots$ |
| HDLC | $\bigcirc$ |
| $\mathrm{HDL}_{2} \mathrm{C}$ | $\underline{\square-1}$ |
| $\mathrm{HDL}_{3} \mathrm{C}$ | $\bigcirc$ |
| Esterified C | $\cdots$ |
| Free C | $\bigcirc$ |

## Glycerides and phospholipids

| Triglycerides | －0－ |
| :---: | :---: |
| VLDL triglycerides | －0－ |
| LDL triglycerides | $\square$ |
| HDL triglycerides | －0－ |
| Phosphoglycerides | $\cdots$ |
| Phosphatidylcholine | $\square$ |
| Sphingomyelins | $\square$ |
| Cholines | $\bigcirc$ |

Apolipoproteins


## Fatty acids

| Total fatty acids | $\square$ |
| :---: | :---: |
| Degree of unsaturation | －8－ |
| Docosahexaenoic acid | 二8－ |
| Linoleic acid | $\bigcirc$ |
| $\mathrm{n}-3$ fatty acids | －0－ |
| n－6 fatty acids | $\rightarrow$ |
| PUFA | $\square$ |
| MUFA | $\square$ |
| Saturated fatty acids | $\square$ |

## Fatty acids ratios

| Docosahexaenoic acid（\％） | $-{ }_{-}^{-1}$ |
| :---: | :---: |
| Linoleic acid（\％） | －0， |
| $\mathrm{n}-3$ fatty acids（\％） | －0－ |
| n－6 fatty acids（\％） | －0－ |
| PUFA（\％） | －8－ |
| MUFA（\％） | －8－ |
| Saturated fatty acids（\％） | 二－ |

## Glycolysis related metabolites



## Amino acids



Branched－chain amino acids

| Isoleucine | －0－ |
| :---: | :---: |
| Leucine | －0 |
| Valine | $\bigcirc$ |

Aromatic amino acids


## Ketone bodies

| Acetate | 工二－ |  |
| :--- | :--- | :--- |
| Beta－hydroxybutyrate | －0－ |  |

Fluid balance

Inflammation


Open symbols： $\mathrm{P}>=0.003$
$\rightarrow$ Main $\quad-\square$ Fully Adjusted

Appendix B Figure B 3：Forest plot of hazard ratios for metastases or PCa death，in the minimally and fully adjusted Cox models．

Lipoprotein subclasses


Lipoprotein particle size


Cholesterol


Glycerides and phospholipids


Apolipoproteins


## Fatty acids



## Fatty acids ratios



## Glycolysis related metabolites



Branched-chain amino acids


Aromatic amino acids


## Ketone bodies



Fluid balance


## Inflammation


$\rightarrow$ Main $\quad \square-$ Fully Adjusted

Appendix B Figure B 4: Forest plot of hazard ratios for clinical progression and metastases or death, in the minimally adjusted Cox models, censored at 10 years of follow-up.

Lipoprotein subclasses


Lipoprotein particle size


Cholesterol


Glycerides and phospholipids


## Apolipoproteins



Closed symblols: $\mathrm{P}<0.003$; Open symbols: $\mathrm{P}>=0.003$

Appendix B Figure B 5：Forest plot of hazard ratios for clinical progression，in the minimally and fully adjusted Cox models，censored at 10 years of follow－up．

Lipoprotein subclasses

| Extremely large VLDL | －8－ |  |
| :---: | :---: | :---: |
| Very large VLDL | －0－ |  |
| Large VLDL | －0－ |  |
| Medium VLDL | －0－ |  |
| Small VLDL | －8－ |  |
| Very Small VLDL | 二－0 |  |
| IDL | －0－ |  |
| Large LDL | 二〇〇－ |  |
| Medium LDL | －0－ |  |
| Small LDL | 二－ |  |
| Very large HDL | －0－ |  |
| Large HDL | 二－8＋ |  |
| Medium HDL | －0－ |  |
| Small HDL | －0－ | 1 |

Lipoprotein particle size

| VLDL particles size |  |
| :--- | :---: |
| LDL particles size |  |
| HDL particles size | O－O－ |

## Cholesterol

| Total C | －00 |
| :--- | ---: |
| VLDL C | $=0$ |
| Remnant C | $=0$ |


| LDL C | －0－ |
| :---: | :---: |
| HDLC | 二ロー |
| $\mathrm{HDL}_{2} \mathrm{C}$ | 二ロー |
| $\mathrm{HDL}_{3} \mathrm{C}$ | 二－ |
| Esterified C | 二8－ |
| Free C | 二8－ |

Glycerides and phospholipids

| Triglycerides | －0－ |
| :---: | :---: |
| VLDL triglycerides | －0－ |
| LDL triglycerides | －0－ |
| HDL triglycerides | ＋\％－ |
| Phosphoglycerides | 二－ |
| Phosphatidylcholine | 二－0 |
| Sphingomyelins | －0－ |
| Cholines | －0－ |

## Apolipoproteins

| Apolipoprotein A－I |  | －0－ |  |  |
| :---: | :---: | :---: | :---: | :---: |
| Apolipoprotein B |  | －8－ |  |  |
| 0 | 0.5 | 1 | 1.5 |  |

Fatty acids

| Total fatty acids | 二8三 |
| :---: | :---: |
| Degree of unsaturation | 二8－ |
| Docosahexaenoic acid | －0－ |
| Linoleic acid | 二口－ |
| $\mathrm{n}-3$ fatty acids | 二0三 |
| $n-6$ fatty acids | 二8－ |
| PUFA | 二8三 |
| MUFA | 7－8－ |
| Saturated fatty acids | －0－ |

## Fatty acids ratios

| Docosahexaenoic acid（\％） | －8－ |
| :---: | :---: |
| Linoleic acid（\％） | 二8－ |
| $\mathrm{n}-3$ fatty acids（\％） | 二8－ |
| n －6 fatty acids（\％） | 二－ |
| PUFA（\％） | 二8－ |
| MUFA（\％） | － |
| Saturated fatty acids（\％） | 二8三 |

## Glycolysis related metabolites

| Glucose | －8－ |
| :---: | :---: |
| Lactate | 二－ |
| Pyruvate | 二－ |
| Citrate Amino acids | －0－ |
| Alanine | －0－ |
| Glutamine | － |
| Glycine | －8－ |
| Histidine | －0－ |

Branched－chain amino acids

| Isoleucine | 二人日 |
| :--- | :--- |
| Leucine | $=0=$ |
| Valine | $=0=$ |

Aromatic amino acids

| Phenylalanine | －0－ |  |
| :---: | :---: | :---: |
| Tyrosine | －0－ |  |
| Ketone bodies |  |  |
| Acetate | －0－ |  |
| Beta－hydroxybutyrate | ＝0－ |  |

## Fluid balance



Inflammation

| Glycoprotein acetyls | 二人日 |  |  |  |
| :--- | :--- | :--- | :--- | :--- |
| 0 | 0.5 | 1 | 1.5 | 2 |
| Open symbols： $\mathrm{P}>=0.003$ |  |  |  |  |

Appendix B Figure B 6: Forest plot of hazard ratios for metastases or PCa death, in the minimally and fully adjusted Cox models, censored at 10 years of follow-up.

Lipoprotein subclasses


Lipoprotein particle size


Cholesterol


Glycerides and phospholipids


Apolipoproteins

$\rightarrow-$ Non-censored $\rightarrow-$ Censored

## Fatty acids



## Fatty acids ratios



## Glycolysis related metabolites



Branched-chain amino acids


Aromatic amino acids


Ketone bodies


Fluid balance


Inflammation


### 9.3. Appendix C

Appendix C Table C 1 Linear regression results of baseline metabolic trait levels in the plant-based diet arm, compared to control.

| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit $95 \% \text { CI }$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| mufa_fa |  | 80 | -1.386 | $-2.874$ | 0.102 | 0.067 |
| vldl_d |  | 80 | -0.917 | -1.607 | -0.227 | 0.01 |
| totfa |  | 80 | -0.383 | -1.624 | 0.858 | 0.541 |
| serum_tg |  | 80 | -0.349 | -0.673 | -0.024 | 0.036 |
| sfa_fa |  | 80 | -0.336 | -1.248 | 0.576 | 0.466 |
| glc |  | 80 | -0.316 | -1.238 | 0.605 | 0.496 |
| vldl_tg |  | 80 | -0.312 | -0.603 | -0.021 | 0.036 |
| mufa |  | 80 | -0.271 | -0.709 | 0.167 | 0.222 |
| m_vldl_l |  | 80 | -0.2 | -0.392 | -0.007 | 0.042 |
| sfa |  | 80 | -0.182 | -0.652 | 0.289 | 0.444 |
| 1_vldl_1 |  | 80 | -0.141 | -0.274 | -0.009 | 0.037 |
| vldl_c |  | 80 | -0.118 | -0.264 | 0.028 | 0.112 |
| tg_pg |  | 80 | -0.115 | -0.251 | 0.022 | 0.098 |
| m_vldl_tg |  | 80 | -0.113 | -0.221 | -0.005 | 0.04 |
| s_vldl_1 |  | 80 | -0.112 | -0.223 | -0.001 | 0.049 |
| remnant_c |  | 80 | -0.089 | -0.319 | 0.141 | 0.445 |
| l_vldl_tg |  | 80 | -0.083 | -0.162 | -0.005 | 0.037 |
| lac |  | 80 | -0.065 | -0.247 | 0.117 | 0.478 |
| s_vldl_tg |  | 80 | -0.062 | -0.118 | -0.006 | 0.031 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit 95\%CI | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| apob_apoa1 |  | 80 | -0.052 | -0.129 | 0.025 | 0.181 |
| m_vldl_c |  | 80 | -0.049 | -0.098 | 0 | 0.051 |
| apob |  | 80 | -0.047 | -0.164 | 0.071 | 0.432 |
| gp |  | 80 | -0.042 | -0.147 | 0.063 | 0.429 |
| xl_vldl_1 |  | 80 | -0.04 | -0.079 | -0.002 | 0.042 |
| m_vldl_pl |  | 80 | -0.038 | -0.075 | -0.001 | 0.043 |
| 1_vldl_c |  | 80 | -0.033 | -0.063 | -0.002 | 0.036 |
| ala |  | 80 | -0.031 | -0.064 | 0.002 | 0.061 |
| xl_vldl_tg |  | 80 | -0.026 | -0.05 | -0.001 | 0.042 |
| s_vldl_c |  | 80 | -0.026 | -0.065 | 0.012 | 0.176 |
| 1_vldl_pl |  | 80 | -0.025 | -0.049 | -0.001 | 0.038 |
| m_vldl_fc |  | 80 | -0.025 | -0.048 | -0.001 | 0.044 |
| s_hdl_pl |  | 80 | -0.025 | -0.059 | 0.009 | 0.151 |
| s_vldl_pl |  | 80 | -0.024 | -0.046 | -0.002 | 0.034 |
| m_vldl_ce |  | 80 | -0.024 | -0.05 | 0.001 | 0.064 |
| s_hdl_1 |  | 80 | -0.021 | -0.077 | 0.035 | 0.457 |
| hdl_tg |  | 80 | -0.018 | -0.033 | -0.003 | 0.018 |
| l_vldl_ce |  | 80 | -0.017 | -0.032 | -0.002 | 0.031 |
| xs_vldl_tg |  | 80 | -0.017 | -0.033 | -0.001 | 0.038 |
| $g 1 n$ |  | 80 | -0.017 | -0.042 | 0.008 | 0.178 |
| l_vldl_fc |  | 80 | -0.016 | -0.032 | -0.001 | 0.041 |
| xs_vldl_1 |  | 80 | -0.016 | -0.08 | 0.049 | 0.631 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit $95 \% \text { CI }$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xxl_vldl_1 |  | 80 | -0.014 | -0.029 | 0 | 0.047 |
| s_vldl_fc |  | 80 | -0.014 | -0.029 | 0 | 0.057 |
| s_vldl_ce |  | 80 | -0.012 | -0.038 | 0.013 | 0.343 |
| xxl_vldl_tg |  | 80 | -0.01 | -0.02 | 0 | 0.047 |
| ldl_tg |  | 80 | -0.01 | -0.03 | 0.009 | 0.301 |
| m_hdl_tg |  | 80 | -0.009 | -0.015 | -0.003 | 0.006 |
| s_hdl_tg |  | 80 | -0.009 | -0.016 | -0.002 | 0.011 |
| xl_vldl_c |  | 80 | -0.008 | -0.016 | 0 | 0.04 |
| idl_tg |  | 80 | -0.008 | -0.02 | 0.003 | 0.158 |
| xl_vldl_pl |  | 80 | -0.007 | -0.013 | 0 | 0.046 |
| xl_vldl_ce |  | 80 | -0.005 | -0.009 | 0 | 0.038 |
| glol |  | 80 | -0.005 | -0.015 | 0.004 | 0.268 |
| ile |  | 80 | -0.005 | -0.015 | 0.005 | 0.298 |
| leu |  | 80 | -0.005 | -0.014 | 0.005 | 0.317 |
| cit |  | 80 | -0.005 | -0.015 | 0.006 | 0.355 |
| 1_ldl_tg |  | 80 | -0.005 | -0.015 | 0.006 | 0.384 |
| xl_vldl_fc |  | 80 | -0.004 | -0.007 | 0 | 0.044 |
| phe |  | 80 | -0.004 | -0.009 | 0 | 0.073 |
| s_ldl_tg |  | 80 | -0.004 | -0.008 | 0.001 | 0.116 |
| gly |  | 80 | -0.004 | -0.02 | 0.011 | 0.59 |
| xxl_vldl_c |  | 80 | -0.003 | -0.005 | 0 | 0.044 |
| tyr |  | 80 | -0.003 | -0.008 | 0.003 | 0.372 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit 95\%CI | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_hdl_fc |  | 80 | -0.003 | -0.009 | 0.003 | 0.382 |
| xxl_vldl_ce |  | 80 | -0.002 | -0.003 | 0 | 0.045 |
| xxl_vldl_pl |  | 80 | -0.002 | -0.004 | 0 | 0.05 |
| m_ldl_tg |  | 80 | -0.002 | -0.007 | 0.003 | 0.405 |
| his |  | 80 | -0.002 | -0.007 | 0.003 | 0.455 |
| val |  | 80 | -0.002 | -0.017 | 0.014 | 0.833 |
| pyr |  | 80 | -0.002 | -0.019 | 0.015 | 0.84 |
| xxl_vldl_fc |  | 80 | -0.001 | -0.002 | 0 | 0.049 |
| bohbut |  | 80 | -0.001 | -0.021 | 0.018 | 0.893 |
| 1_vldl_p |  | 80 | 0 | 0 | 0 | 0.037 |
| m_vldl_p |  | 80 | 0 | 0 | 0 | 0.041 |
| xl_vldl_p |  | 80 | 0 | 0 | 0 | 0.042 |
| s_vldl_p |  | 80 | 0 | 0 | 0 | 0.044 |
| xxl_vldl_p |  | 80 | 0 | 0 | 0 | 0.046 |
| xl_hdl_p |  | 80 | 0 | 0 | 0 | 0.092 |
| 1_hdl_p |  | 80 | 0 | 0 | 0 | 0.107 |
| s_hdl_p |  | 80 | 0 | 0 | 0 | 0.364 |
| xs_vldl_p |  | 80 | 0 | 0 | 0 | 0.51 |
| 1_ldl_p |  | 80 | 0 | 0 | 0 | 0.661 |
| m_ldl_p |  | 80 | 0 | 0 | 0 | 0.676 |
| s_ldl_p |  | 80 | 0 | 0 | 0 | 0.694 |
| idl_p |  | 80 | 0 | 0 | 0 | 0.742 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit $95 \% \text { CI }$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_hdl_tg |  | 80 | 0 | -0.004 | 0.003 | 0.783 |
| m_hdl_p |  | 80 | 0 | 0 | 0 | 0.85 |
| s_ldl_pl |  | 80 | 0 | -0.016 | 0.017 | 0.956 |
| crea |  | 80 | 0 | -0.006 | 0.006 | 0.968 |
| xs_vldl_ce |  | 80 | 0 | -0.023 | 0.023 | 0.983 |
| xs_vldl_pl |  | 80 | 0 | -0.022 | 0.022 | 0.986 |
| m_ldl_pl |  | 80 | 0 | -0.024 | 0.024 | 0.991 |
| alb |  | 80 | 0.001 | -0.002 | 0.004 | 0.561 |
| xs_vldl_fc |  | 80 | 0.001 | -0.01 | 0.011 | 0.877 |
| dha |  | 80 | 0.001 | -0.018 | 0.02 | 0.927 |
| xs_vldl_c |  | 80 | 0.001 | -0.032 | 0.034 | 0.949 |
| faw3 |  | 80 | 0.001 | -0.052 | 0.053 | 0.982 |
| 1_hdl_tg |  | 80 | 0.002 | -0.003 | 0.007 | 0.363 |
| m_hdl_pl |  | 80 | 0.002 | -0.035 | 0.039 | 0.919 |
| s_ldl_fc |  | 80 | 0.003 | -0.008 | 0.014 | 0.562 |
| m_hdl_fc |  | 80 | 0.004 | -0.007 | 0.014 | 0.499 |
| m_ldl_fc |  | 80 | 0.005 | -0.012 | 0.022 | 0.549 |
| xl_hdl_fc |  | 80 | 0.008 | -0.002 | 0.018 | 0.123 |
| 1_ldl_pl |  | 80 | 0.009 | -0.031 | 0.05 | 0.652 |
| ldl_d |  | 80 | 0.009 | -0.065 | 0.083 | 0.812 |
| pc |  | 80 | 0.009 | -0.183 | 0.201 | 0.928 |
| ace |  | 80 | 0.01 | 0.001 | 0.018 | 0.024 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | $\begin{aligned} & \hline \text { Upper Limit } \\ & 95 \% \text { CI } \end{aligned}$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| idl_pl |  | 80 | 0.012 | -0.028 | 0.051 | 0.559 |
| hdl3_c |  | 80 | 0.013 | -0.003 | 0.029 | 0.103 |
| s_hdl_c |  | 80 | 0.013 | -0.022 | 0.048 | 0.462 |
| m_hdl_1 |  | 80 | 0.013 | -0.073 | 0.098 | 0.766 |
| idl_fc |  | 80 | 0.014 | -0.016 | 0.044 | 0.354 |
| totpg |  | 80 | 0.014 | -0.156 | 0.184 | 0.871 |
| idl_ce |  | 80 | 0.015 | -0.063 | 0.093 | 0.7 |
| s_hdl_ce |  | 80 | 0.016 | -0.017 | 0.048 | 0.343 |
| 1_ldl_fc |  | 80 | 0.016 | -0.019 | 0.05 | 0.363 |
| m_hdl_ce |  | 80 | 0.016 | -0.022 | 0.054 | 0.405 |
| 1_hdl_fc |  | 80 | 0.017 | -0.003 | 0.036 | 0.088 |
| s_ldl_ce |  | 80 | 0.017 | -0.028 | 0.062 | 0.452 |
| s_ldl_1 |  | 80 | 0.017 | -0.058 | 0.092 | 0.649 |
| sm |  | 80 | 0.02 | -0.02 | 0.061 | 0.314 |
| m_hdl_c |  | 80 | 0.02 | -0.029 | 0.068 | 0.423 |
| s_ldl_c |  | 80 | 0.02 | -0.035 | 0.076 | 0.471 |
| xl_hdl_ce |  | 80 | 0.022 | -0.003 | 0.047 | 0.087 |
| m_ldl_ce |  | 80 | 0.025 | -0.049 | 0.098 | 0.506 |
| m_ldl_1 |  | 80 | 0.027 | -0.091 | 0.145 | 0.645 |
| 1_ldl_ce |  | 80 | 0.028 | -0.081 | 0.137 | 0.613 |
| freec |  | 80 | 0.028 | -0.124 | 0.179 | 0.719 |
| idl_c |  | 80 | 0.029 | -0.078 | 0.137 | 0.589 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit 95\%CI | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_hdl_c |  | 80 | 0.03 | -0.005 | 0.065 | 0.093 |
| m_ldl_c |  | 80 | 0.03 | -0.06 | 0.12 | 0.514 |
| idl_1 |  | 80 | 0.033 | -0.123 | 0.188 | 0.677 |
| la |  | 80 | 0.038 | -0.289 | 0.366 | 0.816 |
| unsat |  | 80 | 0.039 | 0.009 | 0.068 | 0.011 |
| xl_hdl_pl |  | 80 | 0.043 | -0.006 | 0.092 | 0.087 |
| 1_ldl_c |  | 80 | 0.044 | -0.099 | 0.187 | 0.544 |
| totcho |  | 80 | 0.047 | -0.16 | 0.254 | 0.653 |
| 1_1dl_1 |  | 80 | 0.049 | -0.144 | 0.241 | 0.617 |
| 1_hdl_ce |  | 80 | 0.051 | -0.009 | 0.11 | 0.096 |
| apoa1 |  | 80 | 0.051 | -0.048 | 0.15 | 0.306 |
| 1_hdl_pl |  | 80 | 0.057 | -0.014 | 0.129 | 0.114 |
| dha_fa |  | 80 | 0.057 | -0.096 | 0.211 | 0.462 |
| 1_hdl_c |  | 80 | 0.067 | -0.012 | 0.146 | 0.094 |
| faw6 |  | 80 | 0.068 | -0.303 | 0.44 | 0.714 |
| pufa |  | 80 | 0.069 | -0.341 | 0.479 | 0.738 |
| xl_hdl_1 |  | 80 | 0.071 | -0.012 | 0.154 | 0.091 |
| ldl_c |  | 80 | 0.094 | -0.195 | 0.382 | 0.52 |
| estc |  | 80 | 0.107 | -0.286 | 0.5 | 0.589 |
| faw3_fa |  | 80 | 0.108 | -0.255 | 0.47 | 0.555 |
| hdl_d |  | 80 | 0.115 | -0.008 | 0.237 | 0.066 |
| hdl2_c |  | 80 | 0.117 | -0.035 | 0.268 | 0.129 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \mathrm{CI}$ | Upper Limit $95 \% \text { CI }$ | $P$-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| l_hdl_l |  | 80 | 0.125 | -0.026 | 0.276 | 0.104 |
| hdl_c |  | 80 | 0.13 | -0.035 | 0.294 | 0.12 |
| serum_c |  | 80 | 0.134 | -0.409 | 0.678 | 0.624 |
| la_fa |  | 80 | 1.168 | -0.083 | 2.418 | 0.067 |
| faw6_fa |  | 80 | 1.613 | 0.385 | 2.841 | 0.011 |
| pufa_fa |  | 80 | 1.719 | 0.337 | 3.101 | 0.015 |

Appendix C Table C 2: Linear regression results of baseline metabolic trait levels in the lycopene arm, compared to control.

| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit 95\%CI | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| mufa_fa |  | 80 | -0.97 | -2.388 | 0.449 | 0.177 |
| sfa_fa |  | 80 | -0.788 | -1.634 | 0.059 | 0.068 |
| vldl_d |  | 80 | -0.718 | -1.357 | -0.079 | 0.028 |
| totfa |  | 80 | -0.353 | -1.742 | 1.037 | 0.615 |
| serum_tg |  | 80 | -0.287 | -0.616 | 0.043 | 0.087 |
| vldl_tg |  | 80 | -0.256 | -0.544 | 0.033 | 0.081 |
| sfa |  | 80 | -0.221 | -0.747 | 0.305 | 0.405 |
| mufa |  | 80 | -0.209 | -0.683 | 0.264 | 0.382 |
| glc |  | 80 | -0.17 | -1.104 | 0.764 | 0.718 |
| m_vldl_1 |  | 80 | -0.168 | -0.36 | 0.024 | 0.086 |
| 1_vldı_1 |  | 80 | -0.119 | -0.251 | 0.012 | 0.075 |
| vldl_c |  | 80 | -0.108 | -0.263 | 0.047 | 0.168 |
| remnant_c |  | 80 | -0.104 | -0.356 | 0.147 | 0.41 |
| m_vldl_tg |  | 80 | -0.093 | -0.199 | 0.013 | 0.083 |
| tg_pg |  | 80 | -0.09 | -0.216 | 0.036 | 0.161 |
| s_vldl_l |  | 80 | -0.086 | -0.199 | 0.028 | 0.136 |
| l_vldl_tg |  | 80 | -0.07 | -0.147 | 0.007 | 0.075 |
| gp |  | 80 | -0.055 | -0.162 | 0.052 | 0.308 |
| apob_apoa1 |  | 80 | -0.053 | -0.13 | 0.024 | 0.177 |
| apob |  | 80 | -0.05 | -0.178 | 0.077 | 0.434 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit $95 \% \text { CI }$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_vldl_tg |  | 80 | -0.048 | -0.103 | 0.007 | 0.089 |
| m_vldl_c |  | 80 | -0.043 | -0.093 | 0.008 | 0.096 |
| xl_vldl_1 |  | 80 | -0.035 | -0.074 | 0.004 | 0.079 |
| m_vldl_pl |  | 80 | -0.032 | -0.068 | 0.005 | 0.091 |
| 1_vldl_c |  | 80 | -0.028 | -0.059 | 0.002 | 0.071 |
| ldl_d |  | 80 | -0.024 | -0.069 | 0.022 | 0.304 |
| xl_vldl_tg |  | 80 | -0.022 | -0.046 | 0.003 | 0.081 |
| m_vldl_ce |  | 80 | -0.022 | -0.049 | 0.005 | 0.107 |
| l_vldl_pl |  | 80 | -0.021 | -0.045 | 0.002 | 0.078 |
| s_vldl_c |  | 80 | -0.021 | -0.061 | 0.02 | 0.317 |
| m_vldl_fc |  | 80 | -0.02 | -0.044 | 0.003 | 0.093 |
| xs_vldl_1 |  | 80 | -0.02 | -0.091 | 0.051 | 0.577 |
| s_vldl_pl |  | 80 | -0.017 | -0.039 | 0.005 | 0.127 |
| hdl_tg |  | 80 | -0.015 | -0.032 | 0.001 | 0.069 |
| l_vldl_ce |  | 80 | -0.014 | -0.03 | 0.001 | 0.061 |
| 1_vldl_fc |  | 80 | -0.014 | -0.029 | 0.002 | 0.084 |
| bohbut |  | 80 | -0.013 | -0.027 | 0.001 | 0.07 |
| xxl_vldl_1 |  | 80 | -0.013 | -0.028 | 0.001 | 0.072 |
| s_hdl_pl |  | 80 | -0.013 | -0.046 | 0.019 | 0.417 |
| xs_vldl_tg |  | 80 | -0.012 | -0.029 | 0.005 | 0.159 |
| s_vldl_fc |  | 80 | -0.011 | -0.026 | 0.005 | 0.166 |
| gln |  | 80 | -0.011 | -0.037 | 0.015 | 0.41 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit 95\%CI | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_vldl_ce |  | 80 | -0.01 | -0.037 | 0.017 | 0.461 |
| xxl_vldl_tg |  | 80 | -0.009 | -0.019 | 0.001 | 0.073 |
| ldl_tg |  | 80 | -0.009 | -0.032 | 0.014 | 0.454 |
| s_hdl_tg |  | 80 | -0.007 | -0.014 | 0 | 0.051 |
| xl_vldl_c |  | 80 | -0.007 | -0.015 | 0.001 | 0.068 |
| glol |  | 80 | -0.007 | -0.016 | 0.002 | 0.103 |
| idl_tg |  | 80 | -0.007 | -0.021 | 0.007 | 0.316 |
| xs_vldl_ce |  | 80 | -0.007 | -0.032 | 0.018 | 0.595 |
| s_hdl_1 |  | 80 | -0.007 | -0.056 | 0.043 | 0.788 |
| m_hdl_tg |  | 80 | -0.006 | -0.012 | 0 | 0.045 |
| xl_vldl_pl |  | 80 | -0.006 | -0.012 | 0.001 | 0.09 |
| xs_vldl_c |  | 80 | -0.006 | -0.042 | 0.029 | 0.722 |
| freec |  | 80 | -0.006 | -0.173 | 0.161 | 0.944 |
| crea |  | 80 | -0.005 | -0.01 | 0.001 | 0.104 |
| xl_vldl_ce |  | 80 | -0.004 | -0.009 | 0 | 0.06 |
| 1_ldl_tg |  | 80 | -0.004 | -0.017 | 0.008 | 0.481 |
| idl_ce |  | 80 | -0.004 | -0.088 | 0.081 | 0.931 |
| xxl_vldl_c |  | 80 | -0.003 | -0.005 | 0 | 0.064 |
| xl_vldl_fc |  | 80 | -0.003 | -0.007 | 0 | 0.077 |
| s_ldl_tg |  | 80 | -0.003 | -0.008 | 0.002 | 0.295 |
| cit |  | 80 | -0.003 | -0.013 | 0.007 | 0.541 |
| m_ldl_pl |  | 80 | -0.003 | -0.029 | 0.023 | 0.811 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit 95\%CI | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xxl_vldl_ce |  | 80 | -0.002 | -0.003 | 0 | 0.055 |
| xxl_vldl_pl |  | 80 | -0.002 | -0.003 | 0 | 0.088 |
| tyr |  | 80 | -0.002 | -0.008 | 0.003 | 0.381 |
| m_ldl_tg |  | 80 | -0.002 | -0.008 | 0.004 | 0.577 |
| s_ldl_pl |  | 80 | -0.002 | -0.019 | 0.016 | 0.846 |
| xxl_vldl_fc |  | 80 | -0.001 | -0.002 | 0 | 0.088 |
| xl_hdl_tg |  | 80 | -0.001 | -0.005 | 0.003 | 0.695 |
| s_hdl_fc |  | 80 | -0.001 | -0.007 | 0.005 | 0.79 |
| xs_vldl_pl |  | 80 | -0.001 | -0.025 | 0.022 | 0.908 |
| xxl_vldl_p |  | 80 | 0 | 0 | 0 | 0.071 |
| 1_vldl_p |  | 80 | 0 | 0 | 0 | 0.074 |
| xl_vldl_p |  | 80 | 0 | 0 | 0 | 0.079 |
| m_vldl_p |  | 80 | 0 | 0 | 0 | 0.085 |
| s_vldl_p |  | 80 | 0 | 0 | 0 | 0.124 |
| 1_hdl_p |  | 80 | 0 | 0 | 0 | 0.198 |
| xl_hdl_p |  | 80 | 0 | 0 | 0 | 0.317 |
| xs_vldl_p |  | 80 | 0 | 0 | 0 | 0.507 |
| s_hdl_p |  | 80 | 0 | 0 | 0 | 0.681 |
| m_hdl_p |  | 80 | 0 | 0 | 0 | 0.707 |
| m_ldl_p |  | 80 | 0 | 0 | 0 | 0.858 |
| s_ldl_p |  | 80 | 0 | 0 | 0 | 0.873 |
| 1_ldl_p |  | 80 | 0 | 0 | 0 | 0.892 |


| Metabolite | N | Beta | Lower Limit 95\%CI | Upper Limit 95\%CI | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: |
| pyr | 80 | 0 | -0.015 | 0.014 | 0.954 |
| idl_p | 80 | 0 | 0 | 0 | 0.965 |
| xs_vldl_fc | 80 | 0 | -0.011 | 0.011 | 0.966 |
| alb | 80 | 0.001 | -0.001 | 0.003 | 0.452 |
| 1_hdl_tg | 80 | 0.001 | -0.003 | 0.006 | 0.61 |
| his | 80 | 0.001 | -0.004 | 0.006 | 0.647 |
| phe | 80 | 0.001 | -0.005 | 0.007 | 0.725 |
| leu | 80 | 0.001 | -0.007 | 0.009 | 0.778 |
| s_ldl_fc | 80 | 0.001 | -0.01 | 0.012 | 0.822 |
| ile | 80 | 0.001 | -0.008 | 0.009 | 0.891 |
| m_ldl_fc | 80 | 0.002 | -0.016 | 0.02 | 0.813 |
| idl_1 | 80 | 0.002 | -0.167 | 0.17 | 0.986 |
| 1_ldl_pl | 80 | 0.003 | -0.04 | 0.046 | 0.894 |
| totpg | 80 | 0.003 | -0.182 | 0.188 | 0.972 |
| xl_hdl_fc | 80 | 0.004 | -0.005 | 0.012 | 0.387 |
| m_hdl_fc | 80 | 0.004 | -0.005 | 0.013 | 0.408 |
| idl_c | 80 | 0.004 | -0.112 | 0.12 | 0.949 |
| m_hdl_pl | 80 | 0.005 | -0.03 | 0.04 | 0.791 |
| idl_pl | 80 | 0.005 | -0.037 | 0.047 | 0.824 |
| ace | 80 | 0.006 | -0.001 | 0.013 | 0.083 |
| pc | 80 | 0.006 | -0.209 | 0.222 | 0.953 |
| idl_fc | 80 | 0.007 | -0.025 | 0.039 | 0.644 |


| Metabolite | N |  | Beta | Lower Limit 95\%СІ | Upper Limit 95\%CI | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_ldl_1 |  | 80 | 0.008 | -0.07 | 0.086 | 0.839 |
| hdl3_c |  | 80 | 0.009 | -0.005 | 0.023 | 0.199 |
| 1_ldl_fc |  | 80 | 0.009 | -0.028 | 0.045 | 0.632 |
| sm |  | 80 | 0.009 | -0.034 | 0.052 | 0.674 |
| s_ldl_ce |  | 80 | 0.011 | -0.035 | 0.057 | 0.636 |
| 1_ldl_ce |  | 80 | 0.011 | -0.105 | 0.127 | 0.846 |
| val |  | 80 | 0.012 | -0.001 | 0.024 | 0.071 |
| 1_hdl_fc |  | 80 | 0.012 | -0.005 | 0.029 | 0.173 |
| dha |  | 80 | 0.012 | -0.009 | 0.033 | 0.266 |
| xl_hdl_ce |  | 80 | 0.012 | -0.012 | 0.035 | 0.335 |
| s_ldl_c |  | 80 | 0.012 | -0.045 | 0.07 | 0.67 |
| ala |  | 80 | 0.013 | -0.021 | 0.047 | 0.457 |
| m_ldl_1 |  | 80 | 0.013 | -0.111 | 0.137 | 0.836 |
| s_hdl_ce |  | 80 | 0.014 | -0.013 | 0.042 | 0.306 |
| s_hdl_c |  | 80 | 0.014 | -0.016 | 0.043 | 0.356 |
| xl_hdl_c |  | 80 | 0.015 | -0.017 | 0.048 | 0.343 |
| totcho |  | 80 | 0.015 | -0.21 | 0.241 | 0.893 |
| m_hdl_ce |  | 80 | 0.016 | -0.018 | 0.05 | 0.345 |
| m_ldl_ce |  | 80 | 0.016 | -0.061 | 0.092 | 0.685 |
| gly |  | 80 | 0.018 | 0.001 | 0.035 | 0.039 |
| m_ldl_c |  | 80 | 0.018 | -0.076 | 0.112 | 0.708 |
| m_hdl_l |  | 80 | 0.019 | -0.06 | 0.097 | 0.639 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit $95 \% \text { CI }$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1_ldl_1 |  | 80 | 0.019 | -0.187 | 0.224 | 0.856 |
| m_hdl_c |  | 80 | 0.02 | -0.023 | 0.064 | 0.357 |
| 1_ldl_c |  | 80 | 0.02 | -0.132 | 0.172 | 0.793 |
| xl_hdl_pl |  | 80 | 0.023 | -0.021 | 0.067 | 0.297 |
| la |  | 80 | 0.025 | -0.341 | 0.39 | 0.894 |
| faw3 |  | 80 | 0.028 | -0.027 | 0.082 | 0.32 |
| apoa1 |  | 80 | 0.032 | -0.061 | 0.126 | 0.489 |
| 1_hdl_ce |  | 80 | 0.033 | -0.019 | 0.086 | 0.206 |
| xl_hdl_1 |  | 80 | 0.038 | -0.037 | 0.112 | 0.316 |
| serum_c |  | 80 | 0.04 | -0.537 | 0.617 | 0.89 |
| 1_hdl_pl |  | 80 | 0.045 | -0.018 | 0.109 | 0.161 |
| 1_hdl_c |  | 80 | 0.045 | -0.024 | 0.115 | 0.197 |
| estc |  | 80 | 0.046 | -0.365 | 0.458 | 0.824 |
| unsat |  | 80 | 0.049 | 0.019 | 0.079 | 0.002 |
| ldl_c |  | 80 | 0.05 | -0.253 | 0.354 | 0.742 |
| faw6 |  | 80 | 0.05 | -0.361 | 0.461 | 0.809 |
| lac |  | 80 | 0.057 | -0.139 | 0.254 | 0.563 |
| hdl_d |  | 80 | 0.078 | -0.029 | 0.185 | 0.15 |
| pufa |  | 80 | 0.078 | -0.374 | 0.529 | 0.733 |
| hdl2_c |  | 80 | 0.085 | -0.05 | 0.22 | 0.212 |
| 1_hdl_1 |  | 80 | 0.088 | -0.046 | 0.222 | 0.194 |
| hdl_c |  | 80 | 0.094 | -0.052 | 0.241 | 0.204 |


| Metabolite | N | Beta | Lower Limit <br> $\mathbf{9 5 \% C I}$ | Upper Limit <br> 95\%CI | P-value |
| :--- | :---: | :---: | :--- | :--- | :--- |
| dha_fa | 80 | 0.16 | -0.006 | 0.327 | 0.059 |
| faw3_fa | 80 | 0.371 | -0.001 | 0.744 | 0.051 |
| la_fa | 80 | 0.977 | -0.333 | 2.286 | 0.142 |
| faw6_fa | 80 | 1.387 | 0.143 | 2.632 | 0.029 |
| pufa_fa | 80 | 1.757 | 0.384 | 3.13 | 0.013 |

Appendix C Table C 3: Linear regression results of baseline metabolic trait levels in the brisk walking arm, compared to control.

| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit $95 \% \text { CI }$ | P- <br> value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| hdl_d |  | 80 | -0.461 | -0.9 | -0.022 | 0.04 |
| l_hdl_tg |  | 80 | -0.442 | -0.882 | -0.002 | 0.049 |
| xl_hdl_pl |  | 80 | -0.426 | -0.866 | 0.014 | 0.058 |
| la_fa |  | 80 | -0.424 | -0.86 | 0.012 | 0.056 |
| 1_hdl_ce |  | 80 | -0.42 | -0.861 | 0.022 | 0.062 |
| 1_hdl_p |  | 80 | -0.418 | -0.859 | 0.023 | 0.063 |
| 1_hdl_1 |  | 80 | -0.416 | -0.857 | 0.025 | 0.064 |
| 1_hdl_c |  | 80 | -0.415 | -0.856 | 0.027 | 0.065 |
| l_hdl_pl |  | 80 | -0.405 | -0.846 | 0.037 | 0.072 |
| 1_hdl_fc |  | 80 | -0.399 | -0.841 | 0.044 | 0.077 |
| xl_hdl_p |  | 80 | -0.387 | -0.829 | 0.056 | 0.086 |
| xl_hdl_1 |  | 80 | -0.384 | $-0.827$ | 0.059 | 0.088 |
| xl_hdl_fc |  | 80 | -0.382 | -0.825 | 0.062 | 0.091 |
| hdl2_c |  | 80 | -0.359 | -0.803 | 0.086 | 0.112 |
| hdl_c |  | 80 | $-0.353$ | -0.797 | 0.092 | 0.118 |
| faw6_fa |  | 80 | -0.33 | -0.771 | 0.11 | 0.14 |
| xl_hdl_c |  | 80 | -0.307 | -0.754 | 0.14 | 0.176 |
| ldl_d |  | 80 | -0.297 | -0.731 | 0.137 | 0.177 |
| m_hdl_fc |  | 80 | -0.292 | -0.736 | 0.152 | 0.194 |
| m_hdl_pl |  | 80 | -0.292 | -0.736 | 0.152 | 0.194 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit 95\%CI | Pvalue |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| glol |  | 80 | -0.281 | -0.719 | 0.158 | 0.206 |
| m_hdl_l |  | 80 | -0.28 | -0.724 | 0.164 | 0.213 |
| m_hdl_p |  | 80 | -0.278 | -0.721 | 0.166 | 0.216 |
| apoa1 |  | 80 | -0.274 | -0.72 | 0.172 | 0.225 |
| xl_hdl_ce |  | 80 | -0.274 | -0.723 | 0.174 | 0.227 |
| m_hdl_c |  | 80 | -0.27 | -0.714 | 0.175 | 0.231 |
| m_hdl_ce |  | 80 | -0.263 | -0.708 | 0.183 | 0.244 |
| hdl3_c |  | 80 | -0.232 | -0.679 | 0.214 | 0.303 |
| $g \mathrm{ln}$ |  | 80 | -0.222 | -0.67 | 0.226 | 0.326 |
| crea |  | 80 | -0.201 | -0.643 | 0.241 | 0.368 |
| pufa_fa |  | 80 | -0.19 | -0.636 | 0.257 | 0.4 |
| la |  | 80 | -0.183 | -0.633 | 0.268 | 0.422 |
| totpg |  | 80 | -0.152 | -0.604 | 0.3 | 0.506 |
| tyr |  | 80 | -0.143 | -0.586 | 0.3 | 0.523 |
| faw6 |  | 80 | -0.13 | -0.581 | 0.322 | 0.57 |
| pc |  | 80 | -0.126 | -0.579 | 0.327 | 0.581 |
| freec |  | 80 | -0.111 | -0.563 | 0.341 | 0.626 |
| m_ldl_tg |  | 80 | -0.109 | -0.566 | 0.349 | 0.638 |
| totcho |  | 80 | -0.107 | -0.559 | 0.346 | 0.641 |
| idl_fc |  | 80 | -0.1 | -0.55 | 0.35 | 0.661 |
| his |  | 80 | -0.088 | -0.527 | 0.352 | 0.693 |
| serum_c |  | 80 | -0.086 | -0.537 | 0.364 | 0.704 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit $95 \% \text { CI }$ | Pvalue |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| pufa |  | 80 | -0.084 | -0.537 | 0.368 | 0.711 |
| xl_hdl_tg |  | 80 | $-0.082$ | -0.538 | 0.373 | 0.72 |
| 1_ldl_fc |  | 80 | -0.08 | -0.53 | 0.37 | 0.723 |
| idl_pl |  | 80 | -0.08 | -0.53 | 0.371 | 0.727 |
| estc |  | 80 | $-0.076$ | -0.526 | 0.374 | 0.737 |
| 1_ldl_tg |  | 80 | $-0.073$ | -0.531 | 0.385 | 0.752 |
| ldl_tg |  | 80 | -0.067 | -0.525 | 0.391 | 0.771 |
| sm |  | 80 | $-0.063$ | -0.514 | 0.388 | 0.781 |
| hdl_tg |  | 80 | -0.062 | -0.519 | 0.395 | 0.787 |
| idl_1 |  | 80 | -0.059 | -0.511 | 0.393 | 0.795 |
| idl_c |  | 80 | -0.057 | -0.508 | 0.394 | 0.802 |
| idl_p |  | 80 | -0.055 | -0.507 | 0.397 | 0.809 |
| 1_ldl_pl |  | 80 | -0.053 | -0.504 | 0.398 | 0.815 |
| s_hdl_pl |  | 80 | -0.049 | -0.498 | 0.399 | 0.827 |
| 1_1dl_1 |  | 80 | -0.048 | -0.499 | 0.403 | 0.832 |
| 1_ldl_p |  | 80 | -0.047 | -0.498 | 0.405 | 0.837 |
| 1_ldl_c |  | 80 | $-0.044$ | -0.495 | 0.406 | 0.845 |
| idl_ce |  | 80 | -0.04 | -0.492 | 0.411 | 0.86 |
| alb |  | 80 | -0.037 | -0.481 | 0.407 | 0.869 |
| 1_ldl_ce |  | 80 | -0.033 | -0.484 | 0.418 | 0.885 |
| s_hdl_fc |  | 80 | -0.03 | -0.477 | 0.418 | 0.895 |
| ldl_c |  | 80 | $-0.028$ | -0.478 | 0.422 | 0.902 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit $95 \% \text { CI }$ | P- <br> value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xs_vldl_pl |  | 80 | -0.018 | -0.47 | 0.435 | 0.939 |
| m_ldl_p |  | 80 | -0.017 | -0.468 | 0.434 | 0.941 |
| s_ldl_ce |  | 80 | -0.016 | -0.466 | 0.433 | 0.942 |
| m_ldl_1 |  | 80 | -0.015 | -0.465 | 0.436 | 0.948 |
| m_ldl_ce |  | 80 | -0.013 | -0.463 | 0.437 | 0.953 |
| m_ldl_c |  | 80 | -0.012 | -0.462 | 0.438 | 0.958 |
| s_ldl_c |  | 80 | -0.01 | -0.46 | 0.439 | 0.963 |
| s_ldl_p |  | 80 | -0.007 | -0.457 | 0.443 | 0.975 |
| m_ldl_fc |  | 80 | -0.007 | -0.456 | 0.443 | 0.977 |
| s_ldl_1 |  | 80 | -0.006 | -0.456 | 0.444 | 0.978 |
| s_hdl_1 |  | 80 | -0.005 | -0.451 | 0.441 | 0.981 |
| m_ldl_pl |  | 80 | -0.002 | -0.454 | 0.449 | 0.991 |
| idl_tg |  | 80 | 0.001 | -0.458 | 0.46 | 0.998 |
| s_hdl_c |  | 80 | 0.004 | -0.439 | 0.447 | 0.987 |
| unsat |  | 80 | 0.005 | -0.446 | 0.456 | 0.982 |
| m_hdl_tg |  | 80 | 0.006 | -0.449 | 0.462 | 0.978 |
| s_ldl_pl |  | 80 | 0.007 | -0.444 | 0.457 | 0.976 |
| s_hdl_ce |  | 80 | 0.01 | -0.433 | 0.453 | 0.965 |
| xs_vldl_fc |  | 80 | 0.011 | -0.441 | 0.463 | 0.963 |
| totfa |  | 80 | 0.012 | -0.443 | 0.468 | 0.958 |
| sfa |  | 80 | 0.012 | -0.444 | 0.468 | 0.958 |
| s_ldl_fc |  | 80 | 0.014 | -0.435 | 0.464 | 0.949 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit 95\%CI | Pvalue |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xs_vldl_c |  | 80 | 0.014 | -0.438 | 0.467 | 0.95 |
| xs_vldl_ce |  | 80 | 0.016 | -0.437 | 0.468 | 0.945 |
| sfa_fa |  | 80 | 0.019 | -0.433 | 0.47 | 0.935 |
| bohbut |  | 80 | 0.032 | -0.412 | 0.475 | 0.887 |
| xs_vldl_1 |  | 80 | 0.033 | -0.422 | 0.487 | 0.886 |
| xs_vldl_p |  | 80 | 0.041 | -0.414 | 0.496 | 0.857 |
| ace |  | 80 | 0.043 | -0.411 | 0.497 | 0.851 |
| vldl_d |  | 80 | 0.052 | -0.399 | 0.504 | 0.818 |
| remnant_c |  | 80 | 0.059 | -0.395 | 0.514 | 0.795 |
| apob |  | 80 | 0.075 | -0.379 | 0.528 | 0.744 |
| cit |  | 80 | 0.089 | -0.359 | 0.538 | 0.692 |
| gp |  | 80 | 0.091 | -0.361 | 0.544 | 0.689 |
| xxl_vldl_pl |  | 80 | 0.101 | -0.351 | 0.554 | 0.657 |
| s_vldl_ce |  | 80 | 0.102 | -0.351 | 0.556 | 0.655 |
| mufa |  | 80 | 0.102 | -0.353 | 0.557 | 0.656 |
| xxl_vldl_tg |  | 80 | 0.108 | -0.344 | 0.561 | 0.635 |
| ala |  | 80 | 0.111 | -0.339 | 0.561 | 0.625 |
| xxl_vldl_p |  | 80 | 0.111 | -0.341 | 0.564 | 0.626 |
| xxl_vldl_l |  | 80 | 0.114 | -0.338 | 0.567 | 0.616 |
| xxl_vldl_fc |  | 80 | 0.116 | -0.336 | 0.568 | 0.611 |
| s_vldl_c |  | 80 | 0.125 | -0.329 | 0.579 | 0.585 |
| xs_vldl_tg |  | 80 | 0.126 | -0.33 | 0.581 | 0.585 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit $95 \% \text { CI }$ | P- <br> value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| vldl_c |  | 80 | 0.136 | -0.318 | 0.591 | 0.552 |
| xl_vldl_pl |  | 80 | 0.141 | -0.311 | 0.593 | 0.537 |
| xl_vldl_fc |  | 80 | 0.142 | -0.31 | 0.595 | 0.532 |
| xxl_vldl_c |  | 80 | 0.144 | -0.309 | 0.597 | 0.529 |
| gly |  | 80 | 0.147 | -0.304 | 0.597 | 0.519 |
| serum_tg |  | 80 | 0.15 | -0.304 | 0.603 | 0.514 |
| xl_vldl_tg |  | 80 | 0.151 | -0.301 | 0.603 | 0.509 |
| xl_vldl_p |  | 80 | 0.151 | -0.301 | 0.603 | 0.509 |
| xl_vldl_1 |  | 80 | 0.151 | -0.301 | 0.603 | 0.509 |
| s_vldl_fc |  | 80 | 0.152 | -0.303 | 0.606 | 0.508 |
| 1_vldl_fc |  | 80 | 0.156 | -0.297 | 0.608 | 0.495 |
| m_vldl_ce |  | 80 | 0.159 | -0.295 | 0.614 | 0.488 |
| xl_vldl_c |  | 80 | 0.16 | -0.292 | 0.612 | 0.483 |
| s_hdl_tg |  | 80 | 0.161 | -0.292 | 0.614 | 0.482 |
| xxl_vldl_ce |  | 80 | 0.162 | -0.292 | 0.615 | 0.481 |
| m_vldl_c |  | 80 | 0.164 | -0.29 | 0.618 | 0.473 |
| m_vldl_fc |  | 80 | 0.165 | -0.288 | 0.618 | 0.47 |
| s_vldl_pl |  | 80 | 0.165 | -0.289 | 0.618 | 0.472 |
| s_vldl_1 |  | 80 | 0.171 | -0.283 | 0.624 | 0.457 |
| 1_vldl_pl |  | 80 | 0.172 | -0.28 | 0.625 | 0.451 |
| m_vldl_pl |  | 80 | 0.172 | -0.281 | 0.626 | 0.451 |
| xl_vldl_ce |  | 80 | 0.173 | -0.279 | 0.626 | 0.449 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit 95\%CI | Pvalue |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| l_vldl_c |  | 80 | 0.173 | -0.28 | 0.625 | 0.45 |
| s_vldl_p |  | 80 | 0.175 | -0.279 | 0.629 | 0.445 |
| mufa_fa |  | 80 | 0.177 | -0.272 | 0.626 | 0.435 |
| vldl_tg |  | 80 | 0.178 | -0.275 | 0.63 | 0.436 |
| 1_vldl_1 |  | 80 | 0.179 | -0.273 | 0.632 | 0.432 |
| m_vldl_1 |  | 80 | 0.181 | -0.271 | 0.634 | 0.428 |
| 1_vldl_p |  | 80 | 0.181 | -0.271 | 0.633 | 0.428 |
| 1_vldl_tg |  | 80 | 0.183 | -0.269 | 0.635 | 0.423 |
| m_vldl_p |  | 80 | 0.184 | -0.269 | 0.636 | 0.422 |
| s_vldl_tg |  | 80 | 0.189 | -0.263 | 0.642 | 0.408 |
| 1_vldl_ce |  | 80 | 0.19 | -0.263 | 0.642 | 0.407 |
| m_vldl_tg |  | 80 | 0.191 | -0.262 | 0.643 | 0.404 |
| leu |  | 80 | 0.194 | -0.253 | 0.641 | 0.39 |
| pyr |  | 80 | 0.201 | -0.249 | 0.651 | 0.377 |
| glc |  | 80 | 0.216 | -0.243 | 0.676 | 0.352 |
| ile |  | 80 | 0.224 | -0.223 | 0.67 | 0.322 |
| apob_apoa1 |  | 80 | 0.228 | -0.223 | 0.679 | 0.318 |
| tg_pg |  | 80 | 0.234 | -0.217 | 0.685 | 0.304 |
| val |  | 80 | 0.25 | -0.196 | 0.696 | 0.267 |
| faw3 |  | 80 | 0.257 | -0.19 | 0.703 | 0.256 |
| phe |  | 80 | 0.371 | -0.077 | 0.818 | 0.103 |
| dha |  | 80 | 0.385 | -0.057 | 0.826 | 0.087 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \mathrm{CI}$ | Upper Limit $95 \% \mathrm{CI}$ | $\mathbf{P}$ <br> value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| faw3_fa |  | 80 | 0.394 | -0.047 | 0.835 | 0.079 |
| dha_fa |  | 80 | 0.457 | 0.019 | 0.895 | 0.041 |
| lac |  | 80 | 0.497 | 0.059 | 0.934 | 0.027 |

Appendix C Table C 4:Linear regression results of metabolic traits in the lycopene arm compared to control (ITT).

| Metabolite | N |  | Beta |  | Lower 95\%CI | Limit | $\begin{aligned} & \text { Upper Li } \\ & 95 \% \mathrm{CI} \end{aligned}$ |  | p-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| ala |  | 74 |  | -0.614 |  | -1.183 |  | -0.046 | 0.035 |
| pyr |  | 74 |  | -0.514 |  | -1.168 |  | 0.14 | 0.122 |
| s_hdl_l |  | 74 |  | -0.504 |  | -1.147 |  | 0.138 | 0.122 |
| s_hdl_p |  | 74 |  | -0.5 |  | -1.136 |  | 0.135 | 0.121 |
| s_hdl_c |  | 74 |  | -0.44 |  | -1.047 |  | 0.167 | 0.152 |
| s_hdl_ce |  | 74 |  | -0.393 |  | -0.979 |  | 0.193 | 0.186 |
| alb |  | 74 |  | -0.39 |  | -1.022 |  | 0.242 | 0.222 |
| glc |  | 74 |  | -0.387 |  | -1.093 |  | 0.32 | 0.279 |
| sfa_fa |  | 74 |  | -0.369 |  | -0.878 |  | 0.141 | 0.153 |
| ldl_d |  | 74 |  | 0.369 |  | -0.224 |  | 0.962 | 0.218 |
| la_fa |  | 74 |  | 0.345 |  | -0.222 |  | 0.913 | 0.229 |
| lac |  | 74 |  | -0.334 |  | -0.924 |  | 0.257 | 0.263 |
| faw6_fa |  | 74 |  | 0.311 |  | -0.231 |  | 0.853 | 0.257 |
| xl_hdl_fc |  | 74 |  | 0.305 |  | -0.199 |  | 0.809 | 0.231 |
| gln |  | 74 |  | -0.299 |  | -0.799 |  | 0.202 | 0.238 |
| s_hdl_pl |  | 74 |  | -0.294 |  | -0.893 |  | 0.305 | 0.331 |
| 1_hdl_tg |  | 74 |  | 0.286 |  | -0.281 |  | 0.852 | 0.318 |
| hdl_d |  | 74 |  | 0.285 |  | -0.241 |  | 0.811 | 0.283 |
| s_hdl_fc |  | 74 |  | -0.282 |  | -0.896 |  | 0.332 | 0.363 |
| pufa_fa |  | 74 |  | 0.269 |  | -0.266 |  | 0.804 | 0.32 |
| ace |  | 74 |  | 0.265 |  | -0.134 |  | 0.663 | 0.189 |
| xl_hdl_c |  | 74 |  | 0.26 |  | -0.267 |  | 0.788 | 0.328 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit 95\%CI | p-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_ldl_pl |  | 74 | -0.258 | -0.894 | 0.377 | 0.421 |
| xl_hdl_1 |  | 74 | 0.256 | -0.265 | 0.776 | 0.331 |
| xl_hdl_p |  | 74 | 0.254 | -0.269 | 0.776 | 0.336 |
| s_vldl_ce |  | 74 | -0.244 | -0.847 | 0.36 | 0.424 |
| m_hdl_p |  | 74 | -0.244 | -0.885 | 0.398 | 0.451 |
| cit |  | 74 | -0.239 | -0.672 | 0.194 | 0.275 |
| xl_hdl_ce |  | 74 | 0.239 | -0.299 | 0.777 | 0.378 |
| m_hdl_1 |  | 74 | -0.237 | -0.881 | 0.406 | 0.465 |
| s_ldl_fc |  | 74 | -0.236 | -0.869 | 0.396 | 0.459 |
| s_ldl_1 |  | 74 | -0.234 | -0.863 | 0.394 | 0.459 |
| s_ldl_p |  | 74 | -0.233 | -0.862 | 0.395 | 0.461 |
| m_ldl_pl |  | 74 | -0.232 | -0.864 | 0.4 | 0.467 |
| s_ldl_c |  | 74 | -0.229 | -0.85 | 0.392 | 0.464 |
| m_hdl_pl |  | 74 | -0.229 | -0.863 | 0.405 | 0.474 |
| xs_vldl_pl |  | 74 | -0.228 | -0.844 | 0.387 | 0.462 |
| s_ldl_ce |  | 74 | -0.227 | -0.844 | 0.391 | 0.467 |
| xl_hdl_pl |  | 74 | 0.226 | -0.301 | 0.753 | 0.395 |
| m_ldl_ce |  | 74 | -0.225 | -0.844 | 0.394 | 0.471 |
| m_ldl_c |  | 74 | -0.224 | -0.845 | 0.397 | 0.474 |
| m_ldl_1 |  | 74 | -0.221 | -0.846 | 0.404 | 0.484 |
| xs_vldl_fc |  | 74 | -0.218 | -0.832 | 0.395 | 0.48 |
| m_ldl_p |  | 74 | -0.218 | -0.842 | 0.406 | 0.488 |



| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit 95\%CI | p-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1_ldl_fc |  | 74 | -0.175 | -0.791 | 0.441 | 0.573 |
| idl_c |  | 74 | -0.174 | -0.804 | 0.456 | 0.584 |
| s_vldl_pl |  | 74 | -0.171 | -0.739 | 0.397 | 0.55 |
| xs_vldl_1 |  | 74 | -0.169 | -0.785 | 0.447 | 0.587 |
| remnant_c |  | 74 | -0.167 | -0.784 | 0.449 | 0.59 |
| idl_1 |  | 74 | -0.167 | -0.794 | 0.46 | 0.598 |
| xl_hdl_tg |  | 74 | 0.166 | -0.427 | 0.759 | 0.579 |
| idl_p |  | 74 | -0.163 | -0.79 | 0.465 | 0.607 |
| 1_hdl_1 |  | 74 | 0.161 | -0.382 | 0.703 | 0.557 |
| xs_vldl_p |  | 74 | -0.161 | -0.775 | 0.453 | 0.603 |
| idl_fc |  | 74 | -0.161 | -0.778 | 0.455 | 0.603 |
| xs_vldl_c |  | 74 | -0.16 | -0.782 | 0.462 | 0.61 |
| freec |  | 74 | -0.16 | -0.786 | 0.467 | 0.612 |
| 1_hdl_p |  | 74 | 0.158 | -0.385 | 0.701 | 0.563 |
| faw3 |  | 74 | -0.156 | -0.84 | 0.527 | 0.65 |
| phe |  | 74 | 0.154 | -0.466 | 0.774 | 0.622 |
| s_vldl_fc |  | 74 | -0.153 | -0.731 | 0.424 | 0.598 |
| sfa |  | 74 | -0.153 | -0.749 | 0.442 | 0.609 |
| sm |  | 74 | -0.148 | -0.792 | 0.497 | 0.649 |
| s_vldl_l |  | 74 | -0.146 | -0.722 | 0.429 | 0.613 |
| pc |  | 74 | -0.141 | -0.742 | 0.461 | 0.642 |
| m_vldl_ce |  | 74 | -0.138 | -0.721 | 0.444 | 0.637 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit $95 \% \text { CI }$ | p-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| vldl_c |  | 74 | -0.138 | -0.728 | 0.453 | 0.643 |
| s_vldl_p |  | 74 | -0.136 | -0.71 | 0.437 | 0.637 |
| tyr |  | 74 | -0.135 | -0.766 | 0.496 | 0.671 |
| xs_vldl_ce |  | 74 | -0.131 | -0.757 | 0.496 | 0.679 |
| gly |  | 74 | 0.128 | -0.32 | 0.577 | 0.57 |
| apoa1 |  | 74 | -0.113 | -0.71 | 0.483 | 0.706 |
| xxl_vldl_ce |  | 74 | -0.106 | -0.673 | 0.461 | 0.71 |
| his |  | 74 | -0.106 | -0.685 | 0.473 | 0.716 |
| totcho |  | 74 | -0.106 | -0.72 | 0.508 | 0.731 |
| m_vldl_c |  | 74 | -0.102 | -0.674 | 0.471 | 0.725 |
| la |  | 74 | 0.102 | -0.495 | 0.698 | 0.735 |
| 1_hdl_pl |  | 74 | 0.094 | -0.463 | 0.652 | 0.737 |
| dha_fa |  | 74 | 0.094 | -0.501 | 0.69 | 0.753 |
| bohbut |  | 74 | -0.094 | -0.765 | 0.578 | 0.782 |
| dha |  | 74 | -0.092 | -0.769 | 0.585 | 0.787 |
| glol |  | 73 | -0.09 | -0.719 | 0.539 | 0.776 |
| crea |  | 74 | -0.088 | -0.553 | 0.378 | 0.708 |
| s_ldl_tg |  | 74 | -0.088 | -0.674 | 0.497 | 0.765 |
| m_vldl_pl |  | 74 | -0.085 | -0.651 | 0.48 | 0.764 |
| hdl3_c |  | 74 | -0.085 | -0.678 | 0.508 | 0.776 |
| m_vldl_1 |  | 74 | -0.083 | -0.648 | 0.482 | 0.771 |
| m_vldl_p |  | 74 | -0.082 | -0.647 | 0.482 | 0.773 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit 95\%CI | p-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| l_vldl_ce |  | 74 | -0.08 | -0.641 | 0.481 | 0.776 |
| s_vldl_tg |  | 74 | -0.08 | -0.645 | 0.485 | 0.778 |
| m_vldl_tg |  | 74 | -0.073 | -0.635 | 0.489 | 0.797 |
| val |  | 74 | 0.072 | -0.526 | 0.67 | 0.811 |
| gp |  | 74 | 0.07 | -0.485 | 0.625 | 0.802 |
| 1_vldl_pl |  | 74 | -0.07 | -0.629 | 0.489 | 0.804 |
| 1_vldl_tg |  | 74 | -0.069 | -0.626 | 0.488 | 0.806 |
| vldl_tg |  | 74 | -0.069 | -0.631 | 0.493 | 0.808 |
| 1_vldl_p |  | 74 | -0.068 | -0.626 | 0.489 | 0.807 |
| xs_vldl_tg |  | 74 | -0.068 | -0.646 | 0.51 | 0.815 |
| totfa |  | 74 | -0.068 | -0.67 | 0.534 | 0.822 |
| 1_vldl_1 |  | 74 | -0.067 | -0.625 | 0.491 | 0.811 |
| serum_tg |  | 74 | -0.063 | -0.628 | 0.502 | 0.825 |
| m_vldl_fc |  | 74 | -0.062 | -0.627 | 0.503 | 0.827 |
| vldl_d |  | 74 | -0.061 | -0.618 | 0.497 | 0.829 |
| 1_vldl_c |  | 74 | -0.061 | -0.621 | 0.499 | 0.829 |
| xxl_vldl_c |  | 74 | -0.059 | -0.623 | 0.505 | 0.836 |
| xl_vldl_ce |  | 74 | -0.058 | -0.619 | 0.503 | 0.837 |
| faw6 |  | 74 | 0.055 | -0.555 | 0.666 | 0.858 |
| s_hdl_tg |  | 74 | -0.053 | -0.605 | 0.499 | 0.849 |
| mufa |  | 74 | -0.045 | -0.629 | 0.539 | 0.879 |
| xl_vldl_c |  | 74 | -0.044 | -0.606 | 0.517 | 0.875 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit $95 \% \text { CI }$ | p-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_ldl_tg |  | 74 | -0.044 | -0.64 | 0.552 | 0.883 |
| xl_vldl_tg |  | 74 | -0.043 | -0.602 | 0.515 | 0.878 |
| xl_vldl_p |  | 74 | -0.043 | -0.602 | 0.517 | 0.879 |
| xl_vldl_1 |  | 74 | -0.042 | -0.602 | 0.517 | 0.88 |
| l_vldl_fc |  | 74 | -0.042 | -0.602 | 0.518 | 0.882 |
| ldl_tg |  | 74 | -0.042 | -0.64 | 0.555 | 0.888 |
| xxl_vldl_p |  | 74 | -0.041 | -0.602 | 0.52 | 0.885 |
| faw3_fa |  | 74 | -0.041 | -0.647 | 0.566 | 0.894 |
| xxl_vldl_1 |  | 74 | -0.039 | -0.6 | 0.523 | 0.891 |
| mufa_fa |  | 74 | -0.038 | -0.597 | 0.522 | 0.893 |
| xl_vldl_pl |  | 74 | -0.037 | -0.599 | 0.526 | 0.897 |
| xxl_vldl_tg |  | 74 | -0.036 | -0.597 | 0.525 | 0.898 |
| xl_vldl_fc |  | 74 | -0.029 | -0.592 | 0.533 | 0.917 |
| pufa |  | 74 | 0.027 | -0.592 | 0.647 | 0.93 |
| totpg |  | 74 | -0.027 | -0.638 | 0.585 | 0.931 |
| xxl_vldl_pl |  | 74 | -0.024 | -0.586 | 0.538 | 0.932 |
| 1_ldl_tg |  | 74 | -0.022 | -0.623 | 0.58 | 0.943 |
| ile |  | 74 | 0.021 | -0.568 | 0.61 | 0.944 |
| hdl_tg |  | 74 | 0.019 | -0.542 | 0.58 | 0.947 |
| idl_tg |  | 74 | -0.018 | -0.615 | 0.579 | 0.952 |
| hdl2_c |  | 74 | 0.015 | -0.552 | 0.582 | 0.958 |
| tg_pg |  | 74 | -0.007 | -0.575 | 0.561 | 0.98 |


| Metabolite | N |  | Beta |  | Lower Limit 95\%CI |  | Upper Limit $95 \% \text { CI }$ |  | p-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| hdl_c |  | 74 | 0.006 |  | -0.562 |  | 0.574 |  | 0.983 |
| leu |  | 74 |  | 0.004 |  | -0.579 |  | 0.587 | 0.988 |
| xxl_vldl_fc |  | 74 |  | 0 |  | $-0.562$ |  | 0.562 | 1 |

Appendix C Table C 5:Linear regression results of metabolic traits in the plant-based diet arm compared to control (ITT).

| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit 95\%CI | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| la_fa |  | 74 | 0.584 | 0.087 | 1.081 | 0.022 |
| pyr |  | 74 | -0.68 | -1.277 | -0.082 | 0.026 |
| sfa_fa |  | 74 | -0.604 | -1.149 | -0.059 | 0.03 |
| ace |  | 74 | 0.635 | 0.054 | 1.217 | 0.033 |
| faw6_fa |  | 74 | 0.56 | 0.043 | 1.078 | 0.034 |
| lac |  | 74 | -0.583 | -1.124 | -0.042 | 0.035 |
| s_hdl_p |  | 74 | -0.594 | -1.18 | -0.007 | 0.047 |
| s_hdl_1 |  | 74 | -0.581 | -1.168 | 0.006 | 0.052 |
| pufa_fa |  | 74 | 0.537 | -0.01 | 1.084 | 0.054 |
| glc |  | 74 | -0.629 | -1.303 | 0.045 | 0.067 |
| s_hdl_pl |  | 74 | -0.439 | -1.013 | 0.134 | 0.131 |
| Ala |  | 74 | -0.448 | -1.044 | 0.149 | 0.139 |
| sfa |  | 74 | -0.386 | -0.938 | 0.165 | 0.167 |
| m_hdl_tg |  | 74 | -0.392 | -0.96 | 0.175 | 0.173 |
| s_hdl_fc |  | 74 | -0.39 | -0.969 | 0.188 | 0.183 |
| s_hdl_c |  | 74 | -0.355 | -0.888 | 0.177 | 0.188 |
| gln |  | 74 | 0.307 | -0.196 | 0.81 | 0.228 |
| s_vldl_ce |  | 74 | -0.357 | -0.953 | 0.239 | 0.236 |
| s_ldl_pl |  | 74 | -0.342 | -0.914 | 0.23 | 0.237 |
| s_vldl_c |  | 74 | -0.354 | -0.95 | 0.242 | 0.24 |
| xs_vldl_fc |  | 74 | -0.33 | -0.888 | 0.228 | 0.243 |
| xs_vldl_p |  | 74 | -0.339 | -0.914 | 0.237 | 0.244 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit 95\%CI | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xs_vldl_1 |  | 74 | -0.338 | -0.912 | 0.236 | 0.244 |
| xs_vldl_pl |  | 74 | -0.335 | -0.908 | 0.238 | 0.248 |
| cit |  | 74 | 0.353 | -0.272 | 0.979 | 0.264 |
| xs_vldl_c |  | 74 | -0.314 | -0.879 | 0.25 | 0.271 |
| tyr |  | 74 | -0.314 | -0.889 | 0.261 | 0.271 |
| s_hdl_ce |  | 74 | -0.286 | -0.801 | 0.23 | 0.273 |
| m_ldl_pl |  | 74 | -0.316 | -0.893 | 0.26 | 0.277 |
| s_vldl_pl |  | 74 | -0.32 | -0.906 | 0.265 | 0.279 |
| xl_hdl_fc |  | 74 | 0.304 | -0.251 | 0.859 | 0.279 |
| apob |  | 74 | -0.321 | -0.911 | 0.268 | 0.281 |
| s_vldl_fc |  | 74 | -0.314 | -0.896 | 0.267 | 0.285 |
| pc |  | 74 | -0.285 | -0.815 | 0.244 | 0.287 |
| unsat |  | 74 | 0.327 | -0.284 | 0.938 | 0.289 |
| xs_vldl_ce |  | 74 | -0.303 | -0.87 | 0.265 | 0.291 |
| s_ldl_tg |  | 74 | -0.284 | -0.817 | 0.25 | 0.292 |
| remnant_c |  | 74 | -0.309 | -0.895 | 0.276 | 0.295 |
| ldl_tg |  | 74 | -0.265 | -0.767 | 0.237 | 0.296 |
| phe |  | 74 | -0.279 | -0.811 | 0.252 | 0.298 |
| leu |  | 74 | -0.271 | -0.786 | 0.244 | 0.298 |
| 1_ldl_tg |  | 74 | -0.259 | -0.753 | 0.235 | 0.299 |
| hdl_d |  | 74 | 0.296 | -0.269 | 0.861 | 0.3 |
| idl_tg |  | 74 | -0.269 | -0.782 | 0.244 | 0.3 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit $95 \% \text { CI }$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_hdl_c |  | 74 | 0.28 | -0.257 | 0.817 | 0.302 |
| glol |  | 73 | -0.307 | -0.898 | 0.283 | 0.303 |
| xl_hdl_1 |  | 74 | 0.279 | -0.271 | 0.829 | 0.316 |
| totcho |  | 74 | -0.265 | -0.793 | 0.262 | 0.319 |
| xl_hdl_p |  | 74 | 0.276 | -0.273 | 0.825 | 0.32 |
| xl_hdl_ce |  | 74 | 0.267 | -0.264 | 0.798 | 0.32 |
| m_ldl_tg |  | 74 | -0.247 | -0.739 | 0.245 | 0.32 |
| m_hdl_p |  | 74 | -0.32 | -0.959 | 0.318 | 0.321 |
| s_ldl_fc |  | 74 | -0.279 | -0.835 | 0.277 | 0.321 |
| m_hdl_pl |  | 74 | -0.32 | -0.961 | 0.321 | 0.323 |
| idl_ce |  | 74 | -0.283 | -0.85 | 0.285 | 0.324 |
| idl_p |  | 74 | -0.278 | -0.837 | 0.281 | 0.325 |
| xs_vldl_tg |  | 74 | -0.277 | -0.834 | 0.28 | 0.325 |
| s_vldl_l |  | 74 | -0.293 | -0.884 | 0.298 | 0.326 |
| vldl_c |  | 74 | -0.29 | -0.877 | 0.297 | 0.328 |
| s_ldl_p |  | 74 | -0.278 | -0.841 | 0.285 | 0.328 |
| estc |  | 74 | -0.278 | -0.844 | 0.288 | 0.331 |
| idl_1 |  | 74 | -0.271 | -0.83 | 0.287 | 0.336 |
| s_ldl_1 |  | 74 | -0.273 | -0.834 | 0.289 | 0.337 |
| s_vldl_p |  | 74 | -0.282 | -0.871 | 0.308 | 0.344 |
| m_hdl_1 |  | 74 | -0.305 | -0.945 | 0.335 | 0.346 |
| l_ldl_pl |  | 74 | -0.267 | -0.831 | 0.297 | 0.349 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit $95 \% \text { CI }$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_hdl_pl |  | 74 | 0.266 | -0.297 | 0.83 | 0.349 |
| totpg |  | 74 | -0.25 | -0.779 | 0.279 | 0.349 |
| idl_c |  | 74 | -0.265 | -0.825 | 0.296 | 0.35 |
| totfa |  | 74 | -0.259 | -0.81 | 0.291 | 0.351 |
| serum_c |  | 74 | -0.263 | -0.825 | 0.299 | 0.354 |
| m_ldl_1 |  | 74 | -0.26 | -0.822 | 0.301 | 0.359 |
| m_ldl_p |  | 74 | -0.259 | -0.821 | 0.304 | 0.362 |
| 1_ldl_p |  | 74 | -0.257 | -0.817 | 0.304 | 0.364 |
| m_ldl_fc |  | 74 | -0.256 | -0.813 | 0.302 | 0.364 |
| idl_pl |  | 74 | -0.25 | -0.804 | 0.304 | 0.371 |
| m_vldl_ce |  | 74 | -0.263 | -0.851 | 0.325 | 0.376 |
| 1_1dl_1 |  | 74 | -0.25 | -0.809 | 0.309 | 0.376 |
| apob_apoa1 |  | 74 | -0.273 | -0.883 | 0.338 | 0.377 |
| s_hdl_tg |  | 74 | -0.245 | -0.796 | 0.305 | 0.377 |
| gly |  | 74 | 0.262 | -0.327 | 0.852 | 0.378 |
| 1_ldl_ce |  | 74 | -0.248 | -0.81 | 0.315 | 0.383 |
| mufa |  | 74 | -0.234 | -0.775 | 0.307 | 0.392 |
| alb |  | 74 | -0.291 | -0.972 | 0.39 | 0.397 |
| m_ldl_c |  | 74 | -0.237 | -0.794 | 0.32 | 0.399 |
| ldl_c |  | 74 | -0.237 | -0.794 | 0.32 | 0.399 |
| 1_ldl_c |  | 74 | -0.237 | -0.795 | 0.321 | 0.4 |
| s_ldl_c |  | 74 | -0.234 | -0.789 | 0.32 | 0.402 |


| Metabolite | N | Beta | Lower Limit 95\%CI | Upper Limit $95 \% \text { CI }$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 1_hdl_ce | 74 | 0.256 | -0.357 | 0.868 | 0.408 |
| 1_hdl_c | 74 | 0.255 | -0.358 | 0.868 | 0.409 |
| m_ldl_ce | 74 | -0.232 | -0.788 | 0.325 | 0.409 |
| 1_hdl_fc | 74 | 0.254 | -0.36 | 0.868 | 0.412 |
| 1_hdl_tg | 74 | 0.18 | -0.254 | 0.614 | 0.412 |
| m_vldl_c | 74 | -0.237 | -0.814 | 0.34 | 0.415 |
| freec | 74 | -0.225 | -0.774 | 0.325 | 0.418 |
| s_ldl_ce | 74 | -0.223 | -0.777 | 0.331 | 0.425 |
| xxl_vldl_ce | 74 | -0.224 | -0.786 | 0.339 | 0.43 |
| gp | 74 | -0.204 | -0.725 | 0.316 | 0.437 |
| m_hdl_fc | 74 | -0.251 | -0.897 | 0.395 | 0.441 |
| serum_tg | 74 | -0.219 | -0.785 | 0.347 | 0.443 |
| m_vldl_pl | 74 | -0.221 | -0.794 | 0.352 | 0.445 |
| val | 74 | -0.206 | -0.741 | 0.33 | 0.446 |
| s_vldl_tg | 74 | -0.22 | -0.799 | 0.358 | 0.45 |
| idl_fc | 74 | -0.206 | -0.749 | 0.337 | 0.452 |
| m_vldl_1 | 74 | -0.215 | -0.789 | 0.359 | 0.458 |
| m_vldl_p | 74 | -0.213 | -0.787 | 0.361 | 0.462 |
| 1_vldl_pl | 74 | -0.208 | -0.774 | 0.357 | 0.465 |
| 1_vldl_ce | 74 | -0.208 | -0.774 | 0.357 | 0.465 |
| m_vldl_fc | 74 | -0.206 | -0.771 | 0.359 | 0.47 |
| 1_ldl_fc | 74 | -0.197 | -0.743 | 0.348 | 0.473 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit $95 \% \text { CI }$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| vldl_tg |  | 74 | -0.206 | -0.775 | 0.364 | 0.474 |
| 1_vldl_p |  | 74 | -0.203 | -0.767 | 0.362 | 0.477 |
| hdl_tg |  | 74 | -0.181 | -0.686 | 0.324 | 0.477 |
| 1_vldl_tg |  | 74 | -0.202 | -0.767 | 0.364 | 0.479 |
| 1_vldl_1 |  | 74 | -0.201 | -0.766 | 0.363 | 0.479 |
| m_hdl_c |  | 74 | -0.224 | -0.861 | 0.412 | 0.485 |
| m_vldl_tg |  | 74 | -0.201 | -0.773 | 0.371 | 0.486 |
| 1_vldl_c |  | 74 | -0.196 | -0.755 | 0.364 | 0.488 |
| m_hdl_ce |  | 74 | -0.216 | -0.849 | 0.416 | 0.497 |
| 1_hdl_1 |  | 74 | 0.209 | -0.403 | 0.822 | 0.498 |
| xxl_vldl_c |  | 74 | -0.188 | -0.739 | 0.364 | 0.499 |
| dha |  | 74 | -0.218 | -0.867 | 0.431 | 0.506 |
| faw3 |  | 74 | -0.213 | -0.852 | 0.425 | 0.508 |
| 1_hdl_p |  | 74 | 0.204 | -0.407 | 0.816 | 0.508 |
| 1_vldl_fc |  | 74 | -0.183 | -0.736 | 0.371 | 0.512 |
| vldl_d |  | 74 | -0.19 | -0.767 | 0.388 | 0.515 |
| xl_vldl_ce |  | 74 | -0.179 | -0.731 | 0.372 | 0.518 |
| xl_vldl_c |  | 74 | -0.172 | -0.72 | 0.376 | 0.534 |
| sm |  | 74 | -0.179 | -0.752 | 0.395 | 0.536 |
| xxl_vldl_p |  | 74 | -0.17 | -0.716 | 0.376 | 0.536 |
| xxl_vldl_1 |  | 74 | -0.168 | -0.714 | 0.378 | 0.54 |
| xl_vldl_1 |  | 74 | -0.17 | -0.722 | 0.382 | 0.541 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit $95 \% \text { CI }$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_vldl_p |  | 74 | -0.17 | -0.722 | 0.382 | 0.541 |
| xl_vldl_pl |  | 74 | -0.17 | -0.722 | 0.382 | 0.541 |
| xl_vldl_tg |  | 74 | -0.169 | -0.721 | 0.384 | 0.544 |
| xxl_vldl_tg |  | 74 | -0.165 | -0.71 | 0.38 | 0.548 |
| xl_vldl_fc |  | 74 | -0.164 | -0.708 | 0.381 | 0.551 |
| mufa_fa |  | 74 | -0.173 | -0.75 | 0.405 | 0.553 |
| xxl_vldl_pl |  | 74 | -0.16 | -0.702 | 0.382 | 0.558 |
| his |  | 74 | 0.153 | -0.37 | 0.676 | 0.561 |
| ile |  | 74 | -0.153 | -0.694 | 0.388 | 0.574 |
| crea |  | 74 | 0.177 | -0.449 | 0.802 | 0.575 |
| hdl3_c |  | 74 | -0.159 | -0.733 | 0.414 | 0.582 |
| xxl_vldl_fc |  | 74 | -0.141 | -0.678 | 0.396 | 0.602 |
| faw3_fa |  | 74 | 0.168 | -0.489 | 0.824 | 0.612 |
| 1_hdl_pl |  | 74 | 0.147 | -0.47 | 0.765 | 0.636 |
| apoa1 |  | 74 | -0.13 | -0.748 | 0.488 | 0.676 |
| tg_pg |  | 74 | -0.114 | -0.664 | 0.435 | 0.679 |
| dha_fa |  | 74 | 0.12 | -0.483 | 0.723 | 0.693 |
| bohbut |  | 74 | -0.141 | -0.86 | 0.579 | 0.698 |
| ldl_d |  | 74 | 0.077 | -0.382 | 0.537 | 0.739 |
| pufa |  | 74 | -0.078 | -0.622 | 0.466 | 0.776 |
| hdl2_c |  | 74 | 0.078 | -0.557 | 0.714 | 0.806 |
| faw6 |  | 74 | -0.054 | -0.585 | 0.477 | 0.84 |


| Metabolite | N | Beta | Lower Limit <br> 95\%CI | Upper Limit <br> 95\%CI | P-value |
| :--- | :---: | :---: | :---: | :--- | :--- |
| hdl_c | 74 | 0.058 | -0.575 | 0.692 | 0.855 |
| la | 74 | 0.023 | -0.498 | 0.543 | 0.931 |
| xl_hdl_tg | 74 | -0.004 | -0.503 | 0.495 | 0.987 |

Appendix C Table C 6:Linear regression results of metabolic traits in the lycopene arm compared to control, adjusted for baseline metabolic trait levels and smoking status (ITT).

| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit $95 \% \text { CI }$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| ala |  | 70 | -0.86 | -1.345 | -0.374 | 0.001 |
| pyr |  | 70 | -0.549 | -1.153 | 0.055 | 0.074 |
| lac |  | 70 | -0.5 | -1.111 | 0.111 | 0.107 |
| faw3_fa |  | 70 | -0.493 | -0.922 | -0.064 | 0.025 |
| faw3 |  | 70 | -0.432 | -0.892 | 0.028 | 0.065 |
| unsat |  | 70 | -0.432 | -0.953 | 0.089 | 0.102 |
| s_ldl_ce |  | 70 | -0.391 | -0.82 | 0.038 | 0.073 |
| s_hdl_c |  | 70 | -0.389 | -0.892 | 0.113 | 0.127 |
| s_ldl_c |  | 70 | -0.387 | -0.822 | 0.048 | 0.08 |
| s_hdl_ce |  | 70 | -0.381 | -0.863 | 0.102 | 0.12 |
| m_ldl_ce |  | 70 | -0.379 | -0.808 | 0.05 | 0.083 |
| alb |  | 70 | -0.378 | -0.87 | 0.114 | 0.13 |
| m_ldl_c |  | 70 | -0.374 | -0.807 | 0.059 | 0.089 |
| s_ldl_fc |  | 70 | -0.364 | -0.824 | 0.096 | 0.119 |
| ldl_c |  | 70 | -0.359 | -0.795 | 0.077 | 0.105 |
| cit |  | 70 | -0.359 | -0.824 | 0.106 | 0.128 |
| sm |  | 70 | -0.354 | -0.847 | 0.14 | 0.158 |
| 1_ldl_fc |  | 70 | -0.352 | -0.783 | 0.08 | 0.108 |
| m_ldl_fc |  | 70 | -0.352 | -0.801 | 0.098 | 0.123 |
| s_ldl_1 |  | 70 | -0.347 | -0.791 | 0.096 | 0.123 |
| hdl3_c |  | 70 | -0.34 | -0.786 | 0.106 | 0.133 |
| 1_ldl_c |  | 70 | -0.337 | -0.776 | 0.102 | 0.13 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit $95 \% \mathrm{CI}$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| idl_fc |  | 70 | -0.336 | -0.77 | 0.099 | 0.128 |
| m_ldl_1 |  | 70 | -0.336 | -0.773 | 0.102 | 0.13 |
| s_ldl_p |  | 70 | -0.336 | -0.78 | 0.107 | 0.135 |
| estc |  | 70 | -0.334 | -0.793 | 0.124 | 0.15 |
| 1_ldl_ce |  | 70 | -0.33 | -0.771 | 0.11 | 0.139 |
| dha |  | 70 | -0.33 | $-0.827$ | 0.168 | 0.191 |
| m_ldl_p |  | 70 | -0.326 | -0.763 | 0.11 | 0.141 |
| val |  | 70 | -0.317 | -0.873 | 0.239 | 0.259 |
| 1_ldl_pl |  | 70 | -0.313 | -0.759 | 0.134 | 0.167 |
| 1_ldl_1 |  | 70 | -0.31 | -0.751 | 0.131 | 0.165 |
| serum_c |  | 70 | -0.305 | -0.768 | 0.157 | 0.192 |
| idl_pl |  | 70 | -0.3 | -0.741 | 0.141 | 0.179 |
| 1_ldl_p |  | 70 | -0.297 | -0.74 | 0.145 | 0.184 |
| idl_c |  | 70 | -0.29 | -0.752 | 0.172 | 0.214 |
| s_ldl_pl |  | 70 | -0.289 | -0.757 | 0.18 | 0.223 |
| s_hdl_1 |  | 70 | -0.285 | -0.801 | 0.231 | 0.274 |
| leu |  | 70 | -0.283 | -0.756 | 0.19 | 0.236 |
| apoa1 |  | 70 | -0.274 | -0.711 | 0.163 | 0.216 |
| xs_vldl_fc |  | 70 | -0.271 | -0.714 | 0.171 | 0.225 |
| idl_ce |  | 70 | -0.269 | -0.738 | 0.201 | 0.257 |
| s_hdl_p |  | 70 | -0.263 | -0.768 | 0.242 | 0.302 |
| idl_1 |  | 70 | -0.261 | -0.719 | 0.197 | 0.26 |


| Metabolite | N | Beta | Lower Limit 95\%CI | Upper Limit 95\%CI | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: |
| m_ldl_pl | 70 | -0.257 | -0.715 | 0.202 | 0.268 |
| dha_fa | 70 | -0.247 | -0.687 | 0.193 | 0.266 |
| hdl_c | 70 | -0.244 | -0.593 | 0.105 | 0.168 |
| idl_p | 70 | -0.242 | -0.702 | 0.218 | 0.297 |
| sfa_fa | 70 | -0.24 | -0.779 | 0.298 | 0.375 |
| pufa_fa | 70 | -0.234 | -0.631 | 0.163 | 0.243 |
| freec | 70 | -0.228 | -0.701 | 0.245 | 0.339 |
| hdl2_c | 70 | -0.226 | -0.567 | 0.115 | 0.19 |
| xs_vldl_pl | 70 | -0.225 | -0.672 | 0.221 | 0.317 |
| m_hdl_ce | 70 | -0.22 | -0.721 | 0.282 | 0.385 |
| gly | 70 | $-0.217$ | -0.552 | 0.118 | 0.2 |
| m_hdl_c | 70 | -0.216 | -0.711 | 0.279 | 0.386 |
| totcho | 70 | -0.215 | -0.715 | 0.284 | 0.393 |
| his | 70 | -0.215 | -0.758 | 0.327 | 0.431 |
| 1_hdl_pl | 70 | -0.213 | -0.526 | 0.1 | 0.179 |
| pc | 70 | -0.207 | -0.68 | 0.266 | 0.386 |
| m_hdl_fc | 70 | -0.201 | -0.669 | 0.268 | 0.396 |
| xs_vldl_c | 70 | -0.186 | -0.659 | 0.286 | 0.434 |
| tyr | 70 | -0.184 | -0.839 | 0.471 | 0.577 |
| ile | 70 | -0.175 | -0.672 | 0.322 | 0.485 |
| gln | 70 | -0.174 | -0.629 | 0.282 | 0.449 |
| m_hdl_1 | 70 | -0.169 | -0.649 | 0.311 | 0.484 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit $95 \% \text { CI }$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1_hdl_1 |  | 70 | -0.167 | -0.465 | 0.132 | 0.269 |
| 1_hdl_p |  | 70 | -0.167 | -0.466 | 0.133 | 0.271 |
| 1_hdl_fc |  | 70 | -0.163 | -0.453 | 0.128 | 0.268 |
| m_hdl_p |  | 70 | -0.161 | -0.639 | 0.318 | 0.505 |
| xs_vldl_ce |  | 70 | -0.148 | -0.633 | 0.337 | 0.545 |
| 1_hdl_c |  | 70 | -0.143 | -0.438 | 0.153 | 0.338 |
| glc |  | 70 | -0.137 | -0.532 | 0.258 | 0.492 |
| 1_hdl_ce |  | 70 | -0.136 | -0.433 | 0.162 | 0.366 |
| hdl_d |  | 70 | -0.134 | -0.448 | 0.18 | 0.398 |
| xl_hdl_ce |  | 70 | -0.134 | -0.52 | 0.253 | 0.492 |
| s_vldl_ce |  | 70 | -0.125 | -0.576 | 0.327 | 0.583 |
| s_hdl_fc |  | 70 | -0.125 | -0.62 | 0.37 | 0.614 |
| m_hdl_pl |  | 70 | -0.124 | -0.578 | 0.329 | 0.586 |
| pufa |  | 70 | -0.124 | -0.602 | 0.353 | 0.605 |
| bohbut |  | 70 | -0.114 | -0.7 | 0.472 | 0.7 |
| xs_vldl_1 |  | 70 | -0.109 | $-0.575$ | 0.357 | 0.641 |
| totpg |  | 70 | -0.109 | -0.606 | 0.388 | 0.662 |
| faw6_fa |  | 70 | -0.103 | -0.507 | 0.301 | 0.613 |
| apob |  | 70 | -0.1 | -0.558 | 0.358 | 0.665 |
| xl_hdl_c |  | 70 | -0.099 | -0.481 | 0.283 | 0.605 |
| xl_hdl_p |  | 70 | -0.087 | -0.425 | 0.251 | 0.609 |
| xl_hdl_1 |  | 70 | -0.086 | -0.425 | 0.253 | 0.615 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit 95\%CI | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| remnant_c |  | 70 | -0.08 | -0.543 | 0.383 | 0.731 |
| xl_hdl_pl |  | 70 | -0.079 | -0.373 | 0.215 | 0.594 |
| xs_vldl_p |  | 70 | -0.074 | -0.54 | 0.393 | 0.754 |
| faw6 |  | 70 | -0.074 | -0.553 | 0.406 | 0.76 |
| sfa |  | 70 | -0.057 | -0.53 | 0.416 | 0.811 |
| totfa |  | 70 | -0.033 | -0.51 | 0.443 | 0.889 |
| s_vldl_c |  | 70 | -0.025 | -0.471 | 0.422 | 0.913 |
| s_hdl_pl |  | 70 | -0.022 | -0.488 | 0.443 | 0.925 |
| phe |  | 70 | 0.02 | -0.571 | 0.612 | 0.945 |
| 1_hdl_tg |  | 70 | 0.024 | -0.494 | 0.542 | 0.926 |
| xl_hdl_tg |  | 70 | 0.045 | -0.492 | 0.583 | 0.866 |
| apob_apoa1 |  | 70 | 0.062 | -0.345 | 0.468 | 0.763 |
| m_ldl_tg |  | 70 | 0.066 | -0.435 | 0.568 | 0.792 |
| mufa |  | 70 | 0.081 | -0.385 | 0.546 | 0.73 |
| crea |  | 70 | 0.091 | -0.173 | 0.355 | 0.494 |
| la_fa |  | 70 | 0.091 | -0.311 | 0.493 | 0.652 |
| vldl_c |  | 70 | 0.092 | -0.349 | 0.534 | 0.678 |
| ldl_tg |  | 70 | 0.094 | -0.416 | 0.604 | 0.715 |
| 1_ldl_tg |  | 70 | 0.095 | -0.422 | 0.611 | 0.715 |
| s_ldl_tg |  | 70 | 0.111 | -0.374 | 0.596 | 0.649 |
| s_vldl_fc |  | 70 | 0.145 | -0.28 | 0.57 | 0.499 |
| m_vldl_ce |  | 70 | 0.149 | -0.3 | 0.598 | 0.51 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit 95\%CI | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| gp |  | 70 | 0.155 | -0.264 | 0.573 | 0.463 |
| xxl_vldl_ce |  | 70 | 0.158 | -0.289 | 0.606 | 0.482 |
| idl_tg |  | 70 | 0.166 | -0.343 | 0.675 | 0.517 |
| s_vldl_1 |  | 70 | 0.174 | -0.238 | 0.587 | 0.401 |
| xxl_vldl_pl |  | 70 | 0.175 | -0.221 | 0.571 | 0.38 |
| xxl_vldl_p |  | 70 | 0.175 | -0.225 | 0.574 | 0.385 |
| xxl_vldl_tg |  | 70 | 0.176 | -0.219 | 0.571 | 0.377 |
| xxl_vldl_1 |  | 70 | 0.177 | -0.223 | 0.577 | 0.381 |
| xxl_vldl_c |  | 70 | 0.179 | -0.245 | 0.603 | 0.403 |
| m_vldl_c |  | 70 | 0.191 | -0.232 | 0.613 | 0.37 |
| s_vldl_p |  | 70 | 0.194 | -0.213 | 0.602 | 0.345 |
| xl_vldl_fc |  | 70 | 0.195 | -0.209 | 0.598 | 0.339 |
| s_vldl_pl |  | 70 | 0.195 | -0.217 | 0.606 | 0.348 |
| xxl_vldl_fc |  | 70 | 0.196 | -0.203 | 0.594 | 0.331 |
| glol |  | 69 | 0.199 | -0.252 | 0.65 | 0.381 |
| xl_vldl_pl |  | 70 | 0.2 | -0.198 | 0.598 | 0.319 |
| xl_vldl_c |  | 70 | 0.204 | -0.207 | 0.614 | 0.326 |
| xl_vldl_ce |  | 70 | 0.211 | -0.206 | 0.627 | 0.317 |
| xl_vldl_1 |  | 70 | 0.215 | -0.182 | 0.612 | 0.282 |
| xl_vldl_p |  | 70 | 0.217 | -0.179 | 0.613 | 0.277 |
| mufa_fa |  | 70 | 0.219 | -0.2 | 0.638 | 0.3 |
| m_vldl_pl |  | 70 | 0.221 | -0.173 | 0.615 | 0.267 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit $95 \% \text { CI }$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_vldl_tg |  | 70 | 0.222 | -0.17 | 0.615 | 0.262 |
| l_vldl_ff |  | 70 | 0.222 | -0.172 | 0.616 | 0.264 |
| m_vldl_fc |  | 70 | 0.227 | -0.169 | 0.623 | 0.257 |
| m_vldl_1 |  | 70 | 0.229 | -0.164 | 0.623 | 0.248 |
| 1_vldl_c |  | 70 | 0.23 | -0.173 | 0.632 | 0.258 |
| m_vldl_p |  | 70 | 0.233 | -0.159 | 0.624 | 0.239 |
| xs_vldl_tg |  | 70 | 0.233 | -0.214 | 0.68 | 0.303 |
| l_vldl_pl |  | 70 | 0.235 | -0.156 | 0.626 | 0.235 |
| l_vldl_ce |  | 70 | 0.236 | -0.176 | 0.647 | 0.258 |
| 1_vldl_1 |  | 70 | 0.24 | -0.149 | 0.63 | 0.222 |
| 1_vldl_p |  | 70 | 0.242 | -0.147 | 0.63 | 0.218 |
| 1_vldl_tg |  | 70 | 0.246 | -0.139 | 0.63 | 0.206 |
| m_vldl_tg |  | 70 | 0.248 | -0.135 | 0.63 | 0.2 |
| vldl_tg |  | 70 | 0.252 | -0.136 | 0.641 | 0.199 |
| serum_tg |  | 70 | 0.253 | -0.155 | 0.66 | 0.22 |
| ace |  | 70 | 0.256 | -0.21 | 0.722 | 0.276 |
| tg_pg |  | 70 | 0.265 | -0.13 | 0.66 | 0.185 |
| s_vldl_tg |  | 70 | 0.273 | -0.117 | 0.663 | 0.167 |
| hdl_tg |  | 70 | 0.287 | -0.253 | 0.827 | 0.293 |
| m_hdl_tg |  | 70 | 0.293 | -0.17 | 0.756 | 0.211 |
| s_hdl_tg |  | 70 | 0.366 | -0.07 | 0.801 | 0.099 |
| vldl_d |  | 70 | 0.381 | -0.02 | 0.782 | 0.062 |



Appendix C Table C 7:Linear regression results of metabolic traits in the plant-based diet arm compared to control, adjusted for baseline metabolic trait levels (ITT).

| Metabolite | N | Beta | Lower Limit 95\%CI | Upper Limit $95 \% \text { CI }$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: |
| idl_fc | 74 | -0.479 | -0.834 | -0.124 | 0.009 |
| 1_ldl_fc | 74 | -0.461 | -0.822 | -0.099 | 0.013 |
| idl_pl | 74 | -0.433 | -0.777 | -0.089 | 0.014 |
| vldl_d | 74 | 0.539 | 0.114 | 0.964 | 0.014 |
| hdl3_c | 74 | -0.534 | -0.961 | -0.108 | 0.015 |
| s_hdl_tg | 74 | 0.484 | 0.084 | 0.884 | 0.019 |
| idl_c | 74 | -0.429 | -0.79 | -0.069 | 0.02 |
| idl_1 | 74 | -0.406 | -0.753 | -0.059 | 0.023 |
| s_ldl_ce | 74 | -0.428 | -0.8 | -0.056 | 0.025 |
| 1_ldl_c | 74 | -0.413 | -0.772 | -0.053 | 0.025 |
| s_ldl_c | 74 | -0.427 | -0.802 | -0.053 | 0.026 |
| estc | 74 | -0.423 | -0.795 | -0.052 | 0.026 |
| ldl_c | 74 | -0.418 | $-0.783$ | -0.052 | 0.026 |
| xs_vldl_fc | 74 | -0.394 | -0.739 | -0.049 | 0.026 |
| m_ldl_ce | 74 | -0.416 | -0.783 | -0.049 | 0.027 |
| 1_ldl_1 | 74 | -0.398 | -0.751 | -0.046 | 0.027 |
| idl_p | 74 | -0.39 | -0.733 | -0.046 | 0.027 |
| m_ldl_c | 74 | -0.416 | -0.785 | -0.047 | 0.028 |
| glc | 74 | -0.383 | -0.723 | -0.042 | 0.028 |
| 1_ldl_pl | 74 | -0.401 | -0.759 | -0.043 | 0.029 |
| hdl_c | 74 | -0.384 | -0.729 | -0.04 | 0.029 |
| idl_ce | 74 | -0.405 | -0.77 | -0.041 | 0.03 |


| Metabolite | N | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit $95 \% \text { CI }$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 1_ldl_p | 74 | -0.388 | -0.738 | -0.039 | 0.03 |
| xs_vldl_c | 74 | -0.381 | -0.724 | -0.038 | 0.03 |
| pyr | 74 | -0.602 | -1.148 | -0.056 | 0.031 |
| sm | 74 | -0.448 | -0.852 | -0.043 | 0.031 |
| m_ldl_fc | 74 | -0.413 | -0.788 | -0.037 | 0.032 |
| 1_ldl_ce | 74 | -0.395 | -0.754 | -0.036 | 0.032 |
| 1_hdl_pl | 74 | -0.365 | -0.699 | -0.032 | 0.032 |
| xs_vldl_pl | 74 | -0.366 | -0.703 | -0.03 | 0.033 |
| 1_hdl_p | 74 | -0.337 | -0.647 | -0.028 | 0.033 |
| 1_hdl_1 | 74 | -0.336 | -0.645 | -0.027 | 0.033 |
| s_ldl_fc | 74 | -0.418 | -0.804 | -0.033 | 0.034 |
| serum_c | 74 | -0.395 | -0.76 | -0.03 | 0.035 |
| m_ldl_1 | 74 | -0.389 | -0.749 | -0.028 | 0.035 |
| hdl2_c | 74 | -0.359 | -0.695 | -0.023 | 0.036 |
| s_ldl_1 | 74 | -0.391 | -0.76 | -0.023 | 0.038 |
| xs_vldl_ce | 74 | -0.369 | -0.718 | -0.021 | 0.038 |
| m_ldl_p | 74 | -0.377 | -0.735 | -0.02 | 0.039 |
| xl_hdl_pl | 74 | -0.328 | -0.639 | -0.017 | 0.039 |
| $g 1 n$ | 74 | 0.468 | 0.019 | 0.916 | 0.041 |
| s_ldl_p | 74 | -0.38 | -0.746 | -0.015 | 0.042 |
| hdl_d | 74 | -0.316 | -0.62 | -0.012 | 0.042 |
| 1_hdl_fc | 74 | -0.314 | -0.618 | -0.011 | 0.043 |


| Metabolite | N | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit $95 \% \mathrm{CI}$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 1_hdl_c | 74 | -0.31 | -0.612 | -0.009 | 0.044 |
| 1_hdl_ce | 74 | -0.309 | -0.61 | -0.007 | 0.045 |
| lac | 74 | -0.563 | -1.116 | -0.011 | 0.046 |
| totcho | 74 | -0.369 | -0.738 | 0.001 | 0.05 |
| s_vldl_tg | 74 | 0.407 | -0.003 | 0.818 | 0.052 |
| s_hdl_ce | 74 | -0.447 | -0.901 | 0.006 | 0.053 |
| s_hdl_c | 74 | -0.453 | -0.915 | 0.009 | 0.054 |
| sfa_fa | 74 | -0.535 | -1.084 | 0.014 | 0.056 |
| ace | 74 | 0.628 | -0.022 | 1.277 | 0.058 |
| apoa1 | 74 | -0.361 | -0.741 | 0.018 | 0.061 |
| vldl_tg | 74 | 0.387 | -0.034 | 0.808 | 0.071 |
| s_ldl_pl | 74 | -0.345 | -0.722 | 0.033 | 0.073 |
| m_vldl_tg | 74 | 0.387 | -0.038 | 0.813 | 0.074 |
| freec | 74 | -0.32 | -0.674 | 0.034 | 0.076 |
| l_vldl_tg | 74 | 0.384 | -0.049 | 0.818 | 0.081 |
| xl_hdl_p | 74 | -0.283 | -0.603 | 0.036 | 0.082 |
| serum_tg | 74 | 0.355 | -0.049 | 0.76 | 0.084 |
| 1_vldl_1 | 74 | 0.377 | -0.052 | 0.807 | 0.084 |
| l_vldl_p | 74 | 0.379 | -0.052 | 0.809 | 0.084 |
| m_ldl_pl | 74 | -0.317 | -0.678 | 0.045 | 0.085 |
| xl_vldl_tg | 74 | 0.381 | -0.053 | 0.816 | 0.085 |
| tg_pg | 74 | 0.387 | -0.056 | 0.83 | 0.086 |


| Metabolite | N | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit $95 \% \text { CI }$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: |
| xl_hdl_1 | 74 | -0.278 | -0.598 | 0.041 | 0.087 |
| 1_vldl_c | 74 | 0.367 | -0.056 | 0.79 | 0.088 |
| l_vldl_ce | 74 | 0.364 | -0.057 | 0.786 | 0.089 |
| xl_vldl_p | 74 | 0.372 | -0.059 | 0.804 | 0.089 |
| l_vldl_fc | 74 | 0.366 | -0.059 | 0.791 | 0.09 |
| m_vldl_p | 74 | 0.358 | -0.059 | 0.776 | 0.091 |
| l_vldl_pl | 74 | 0.364 | -0.06 | 0.788 | 0.091 |
| xl_vldl_1 | 74 | 0.37 | -0.06 | 0.8 | 0.091 |
| m_vldl_1 | 74 | 0.352 | -0.063 | 0.768 | 0.095 |
| m_vldl_fc | 74 | 0.349 | -0.064 | 0.762 | 0.096 |
| xl_vldl_pl | 74 | 0.352 | -0.071 | 0.774 | 0.101 |
| gly | 74 | 0.388 | -0.078 | 0.853 | 0.101 |
| m_vldl_pl | 74 | 0.34 | -0.071 | 0.75 | 0.103 |
| xxl_vldl_fc | 74 | 0.351 | -0.072 | 0.775 | 0.103 |
| pc | 74 | -0.303 | -0.669 | 0.063 | 0.104 |
| xl_vldl_c | 74 | 0.348 | -0.076 | 0.772 | 0.106 |
| xl_vldl_fc | 74 | 0.346 | -0.077 | 0.769 | 0.107 |
| m_hdl_tg | 74 | 0.328 | -0.074 | 0.73 | 0.108 |
| xl_vldl_ce | 74 | 0.347 | -0.078 | 0.772 | 0.108 |
| s_hdl_1 | 74 | -0.391 | -0.889 | 0.107 | 0.121 |
| s_vldl_pl | 74 | 0.291 | -0.079 | 0.661 | 0.121 |
| m_hdl_fc | 74 | -0.339 | -0.777 | 0.099 | 0.128 |


| Metabolite | N | Beta | Lower Limit 95\%CI | Upper Limit 95\%CI | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: |
| xs_vldl_tg | 74 | 0.27 | -0.08 | 0.62 | 0.128 |
| xs_vldl_1 | 74 | -0.253 | -0.582 | 0.076 | 0.129 |
| m_hdl_c | 74 | -0.33 | -0.76 | 0.099 | 0.13 |
| m_hdl_ce | 74 | -0.327 | -0.755 | 0.101 | 0.132 |
| xxl_vldl_pl | 74 | 0.324 | -0.1 | 0.747 | 0.132 |
| totpg | 74 | -0.281 | -0.651 | 0.089 | 0.134 |
| s_vldl_p | 74 | 0.287 | -0.092 | 0.665 | 0.135 |
| xxl_vldl_tg | 74 | 0.324 | -0.105 | 0.753 | 0.136 |
| xxl_vldl_1 | 74 | 0.32 | -0.107 | 0.746 | 0.14 |
| xxl_vldl_p | 74 | 0.318 | -0.109 | 0.745 | 0.142 |
| s_hdl_p | 74 | -0.367 | -0.862 | 0.128 | 0.143 |
| hdl_tg | 74 | 0.294 | -0.102 | 0.69 | 0.143 |
| cit | 74 | 0.442 | -0.178 | 1.062 | 0.159 |
| xxl_vldl_c | 74 | 0.295 | -0.124 | 0.715 | 0.165 |
| m_vldl_c | 74 | 0.278 | -0.123 | 0.679 | 0.171 |
| m_hdl_1 | 74 | -0.304 | -0.744 | 0.136 | 0.172 |
| xl_hdl_ce | 74 | -0.236 | -0.577 | 0.105 | 0.172 |
| s_vldl_1 | 74 | 0.257 | -0.116 | 0.63 | 0.173 |
| m_hdl_p | 74 | -0.296 | -0.738 | 0.145 | 0.185 |
| mufa_fa | 74 | 0.287 | -0.152 | 0.725 | 0.196 |
| xl_hdl_c | 74 | -0.219 | -0.553 | 0.116 | 0.197 |
| xs_vldl_p | 74 | -0.208 | -0.535 | 0.119 | 0.21 |


| Metabolite | N | Beta | Lower Limit 95\%CI | Upper Limit 95\%CI | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: |
| m_hdl_pl | 74 | -0.278 | -0.722 | 0.167 | 0.217 |
| s_vldl_fc | 74 | 0.212 | -0.151 | 0.575 | 0.247 |
| xxl_vldl_ce | 74 | 0.24 | -0.181 | 0.66 | 0.26 |
| sfa | 74 | -0.208 | -0.592 | 0.176 | 0.283 |
| m_vldl_ce | 74 | 0.204 | -0.191 | 0.599 | 0.307 |
| xl_hdl_fc | 74 | -0.161 | -0.482 | 0.16 | 0.321 |
| la_fa | 74 | 0.184 | -0.202 | 0.569 | 0.345 |
| alb | 74 | -0.312 | -0.975 | 0.352 | 0.352 |
| dha | 74 | -0.219 | -0.687 | 0.249 | 0.354 |
| tyr | 74 | -0.233 | -0.763 | 0.297 | 0.384 |
| s_vldl_ce | 74 | -0.15 | -0.502 | 0.203 | 0.4 |
| unsat | 74 | -0.211 | -0.715 | 0.294 | 0.407 |
| his | 74 | 0.203 | -0.298 | 0.704 | 0.422 |
| remnant_c | 74 | -0.136 | -0.489 | 0.216 | 0.443 |
| faw3 | 74 | -0.174 | -0.629 | 0.282 | 0.45 |
| apob | 74 | -0.129 | -0.49 | 0.231 | 0.477 |
| pufa | 74 | -0.139 | -0.539 | 0.261 | 0.49 |
| val | 74 | -0.171 | -0.665 | 0.323 | 0.492 |
| crea | 74 | 0.131 | -0.256 | 0.519 | 0.501 |
| faw6 | 74 | -0.126 | -0.517 | 0.265 | 0.521 |
| s_hdl_fc | 74 | -0.152 | -0.633 | 0.329 | 0.531 |
| 1_ldl_tg | 74 | -0.091 | -0.38 | 0.198 | 0.533 |


| Metabolite | N | Beta | Lower Limit 95\%CI | Upper Limit 95\%CI | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 1_hdl_tg | 74 | -0.094 | -0.408 | 0.22 | 0.551 |
| vldl_c | 74 | 0.109 | -0.258 | 0.476 | 0.556 |
| totfa | 74 | -0.107 | -0.488 | 0.274 | 0.576 |
| s_ldl_tg | 74 | 0.093 | -0.251 | 0.437 | 0.593 |
| m_ldl_tg | 74 | -0.078 | -0.371 | 0.215 | 0.599 |
| phe | 74 | -0.133 | -0.673 | 0.407 | 0.624 |
| glol | 73 | -0.098 | -0.544 | 0.348 | 0.662 |
| ala | 74 | -0.112 | -0.627 | 0.404 | 0.667 |
| apob_apoa1 | 74 | 0.072 | -0.264 | 0.408 | 0.67 |
| ile | 74 | 0.101 | -0.374 | 0.576 | 0.672 |
| bohbut | 74 | -0.142 | -0.814 | 0.53 | 0.675 |
| pufa_fa | 74 | -0.078 | -0.452 | 0.296 | 0.678 |
| faw6_fa | 74 | -0.074 | -0.466 | 0.317 | 0.706 |
| mufa | 74 | 0.069 | -0.312 | 0.45 | 0.718 |
| leu | 74 | -0.081 | -0.547 | 0.385 | 0.731 |
| faw3_fa | 74 | 0.069 | -0.349 | 0.488 | 0.742 |
| ldl_tg | 74 | -0.049 | -0.348 | 0.25 | 0.745 |
| s_hdl_pl | 74 | -0.049 | -0.476 | 0.379 | 0.821 |
| gp | 74 | -0.033 | -0.394 | 0.328 | 0.855 |
| idl_tg | 74 | 0.026 | -0.277 | 0.33 | 0.863 |
| s_vldl_c | 74 | -0.021 | -0.372 | 0.331 | 0.906 |
| dha_fa | 74 | -0.023 | -0.422 | 0.376 | 0.909 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit 95\%CI | $P$-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| la | 74 |  | -0.013 | -0.406 | 0.379 | 0.946 |
| ldl_d |  | 74 | 0.009 | -0.378 | 0.397 | 0.961 |
| xl_hdl_tg |  | 74 | 0.01 | -0.426 | 0.445 | 0.965 |

Appendix C Table C 8:Linear regression results of metabolic traits in the brisk walking arm compared to control (ITT).

| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit $95 \% \text { CI }$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| la_fa |  | 74 | 0.564 | 0.117 | 1.011 | 0.014 |
| faw6_fa |  | 74 | 0.542 | 0.093 | 0.991 | 0.019 |
| xxl_vldl_ce |  | 74 | -0.529 | -0.984 | -0.073 | 0.024 |
| xl_hdl_tg |  | 74 | -0.517 | -0.974 | -0.06 | 0.027 |
| xxl_vldl_c |  | 74 | -0.507 | -0.964 | -0.05 | 0.03 |
| sfa |  | 74 | -0.5 | -0.957 | -0.044 | 0.032 |
| xl_vldl_ce |  | 74 | -0.493 | -0.951 | -0.035 | 0.035 |
| pufa_fa |  | 74 | 0.487 | 0.034 | 0.939 | 0.035 |
| xl_vldl_c |  | 74 | -0.487 | -0.945 | -0.029 | 0.037 |
| xl_vldl_fc |  | 74 | -0.48 | -0.939 | -0.022 | 0.04 |
| xxl_vldl_fc |  | 74 | -0.471 | -0.93 | -0.012 | 0.045 |
| xxl_vldl_1 |  | 74 | -0.47 | -0.929 | -0.011 | 0.045 |
| m_vldl_ce |  | 74 | -0.468 | -0.926 | -0.009 | 0.046 |
| xxl_vldl_pl |  | 74 | -0.467 | -0.927 | -0.008 | 0.046 |
| xxl_vldl_p |  | 74 | -0.466 | -0.925 | -0.006 | 0.047 |
| xl_vldl_pl |  | 74 | -0.464 | -0.924 | -0.005 | 0.047 |
| 1_vldl_ce |  | 74 | -0.462 | -0.921 | -0.003 | 0.049 |
| xxl_vldl_tg |  | 74 | -0.46 | -0.92 | 0 | 0.05 |
| vldl_c |  | 74 | -0.457 | -0.916 | 0.002 | 0.051 |
| m_vldl_c |  | 74 | -0.455 | -0.915 | 0.004 | 0.052 |
| 1_vldl_c |  | 74 | -0.455 | -0.914 | 0.005 | 0.052 |
| xl_vldl_1 |  | 74 | -0.455 | -0.914 | 0.005 | 0.052 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \mathrm{CI}$ | Upper Limit 95\%CI | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| hdl_tg |  | 74 | -0.454 | -0.914 | 0.006 | 0.053 |
| idl_tg |  | 74 | -0.454 | -0.915 | 0.007 | 0.054 |
| xl_vldl_p |  | 74 | -0.451 | -0.911 | 0.008 | 0.054 |
| xs_vldl_tg |  | 74 | -0.447 | -0.907 | 0.014 | 0.057 |
| 1_vldl_fc |  | 74 | -0.446 | -0.906 | 0.013 | 0.057 |
| xl_vldl_tg |  | 74 | -0.44 | -0.901 | 0.02 | 0.06 |
| sfa_fa |  | 74 | -0.431 | -0.884 | 0.022 | 0.062 |
| serum_tg |  | 74 | -0.437 | -0.897 | 0.024 | 0.063 |
| ile |  | 74 | -0.433 | -0.89 | 0.024 | 0.063 |
| m_vldl_fc |  | 74 | -0.433 | -0.894 | 0.027 | 0.065 |
| 1_vldl_pl |  | 74 | -0.432 | -0.893 | 0.028 | 0.065 |
| s_ldl_tg |  | 74 | -0.433 | -0.895 | 0.029 | 0.066 |
| 1_vldl_1 |  | 74 | -0.428 | -0.888 | 0.033 | 0.068 |
| m_vldl_pl |  | 74 | -0.428 | -0.888 | 0.033 | 0.068 |
| s_hdl_tg |  | 74 | -0.426 | -0.887 | 0.035 | 0.07 |
| m_vldl_1 |  | 74 | -0.426 | -0.886 | 0.035 | 0.07 |
| mufa |  | 74 | -0.426 | -0.887 | 0.036 | 0.07 |
| 1_vldl_p |  | 74 | -0.425 | -0.885 | 0.036 | 0.07 |
| m_vldl_p |  | 74 | -0.422 | -0.883 | 0.038 | 0.072 |
| leu |  | 74 | -0.418 | -0.876 | 0.04 | 0.073 |
| vldl_tg |  | 74 | -0.417 | -0.878 | 0.044 | 0.075 |
| totfa |  | 74 | -0.416 | -0.878 | 0.046 | 0.077 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit $95 \% \text { CI }$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1_vldl_tg |  | 74 | -0.415 | -0.876 | 0.047 | 0.077 |
| 1_ldl_tg |  | 74 | -0.413 | -0.876 | 0.051 | 0.08 |
| m_vldl_tg |  | 74 | -0.408 | -0.87 | 0.053 | 0.082 |
| s_vldl_fc |  | 74 | -0.407 | -0.869 | 0.055 | 0.083 |
| ldl_tg |  | 74 | -0.407 | -0.871 | 0.056 | 0.084 |
| remnant_c |  | 74 | -0.405 | -0.867 | 0.056 | 0.084 |
| s_vldl_1 |  | 74 | -0.403 | -0.865 | 0.059 | 0.086 |
| s_vldl_p |  | 74 | -0.403 | -0.865 | 0.059 | 0.086 |
| s_vldl_tg |  | 74 | -0.397 | -0.859 | 0.065 | 0.091 |
| apob_apoa1 |  | 74 | -0.392 | -0.854 | 0.069 | 0.094 |
| apob |  | 74 | -0.386 | -0.848 | 0.076 | 0.1 |
| s_vldl_pl |  | 74 | -0.383 | -0.846 | 0.08 | 0.103 |
| s_vldl_c |  | 74 | -0.374 | -0.837 | 0.089 | 0.111 |
| xs_vldl_p |  | 74 | -0.372 | -0.834 | 0.091 | 0.114 |
| s_hdl_ce |  | 74 | 0.367 | -0.094 | 0.827 | 0.117 |
| faw3 |  | 74 | -0.365 | -0.827 | 0.096 | 0.119 |
| unsat |  | 74 | 0.356 | -0.102 | 0.813 | 0.126 |
| s_hdl_c |  | 74 | 0.356 | -0.105 | 0.817 | 0.129 |
| xs_vldl_1 |  | 74 | -0.354 | -0.817 | 0.109 | 0.132 |
| m_ldl_tg |  | 74 | -0.354 | -0.82 | 0.112 | 0.134 |
| gp |  | 74 | -0.347 | -0.808 | 0.114 | 0.137 |
| tg_pg |  | 74 | -0.348 | -0.811 | 0.115 | 0.139 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit $95 \% \text { CI }$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| bohbut |  | 74 | 0.338 | -0.114 | 0.789 | 0.141 |
| totcho |  | 74 | -0.344 | -0.808 | 0.12 | 0.144 |
| xs_vldl_ce |  | 74 | -0.336 | -0.799 | 0.127 | 0.152 |
| totpg |  | 74 | -0.334 | -0.798 | 0.131 | 0.157 |
| s_vldl_ce |  | 74 | -0.331 | -0.795 | 0.133 | 0.16 |
| pc |  | 74 | -0.324 | -0.789 | 0.141 | 0.169 |
| xs_vldl_c |  | 74 | -0.316 | -0.78 | 0.148 | 0.178 |
| m_hdl_tg |  | 74 | -0.317 | -0.781 | 0.148 | 0.179 |
| m_hdl_c |  | 74 | 0.277 | -0.187 | 0.74 | 0.238 |
| m_hdl_ce |  | 74 | 0.277 | -0.187 | 0.74 | 0.238 |
| m_hdl_fc |  | 74 | 0.272 | -0.191 | 0.735 | 0.246 |
| ace |  | 74 | 0.265 | -0.193 | 0.724 | 0.252 |
| xs_vldl_fc |  | 74 | -0.265 | -0.731 | 0.2 | 0.26 |
| 1_hdl_pl |  | 74 | 0.263 | -0.202 | 0.728 | 0.264 |
| idl_ce |  | 74 | -0.259 | -0.725 | 0.206 | 0.27 |
| s_ldl_pl |  | 74 | -0.259 | -0.725 | 0.207 | 0.272 |
| hdl2_c |  | 74 | 0.256 | -0.208 | 0.721 | 0.275 |
| vldl_d |  | 74 | -0.256 | -0.72 | 0.209 | 0.276 |
| freec |  | 74 | -0.256 | -0.723 | 0.21 | 0.277 |
| idl_p |  | 74 | -0.256 | -0.721 | 0.21 | 0.278 |
| 1_hdl_tg |  | 74 | -0.255 | -0.724 | 0.214 | 0.282 |
| dha |  | 74 | -0.252 | -0.717 | 0.212 | 0.283 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | $\begin{aligned} & \text { Upper Limit } \\ & 95 \% \text { CI } \end{aligned}$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_ldl_pl |  | 74 | -0.248 | -0.714 | 0.218 | 0.292 |
| hdl_c |  | 74 | 0.243 | -0.222 | 0.708 | 0.301 |
| mufa_fa |  | 74 | -0.242 | -0.706 | 0.223 | 0.303 |
| idl_1 |  | 74 | -0.237 | -0.703 | 0.229 | 0.315 |
| crea |  | 74 | 0.231 | -0.226 | 0.688 | 0.317 |
| 1_hdl_fc |  | 74 | 0.23 | -0.237 | 0.697 | 0.329 |
| xs_vldl_pl |  | 74 | -0.226 | -0.692 | 0.241 | 0.338 |
| idl_c |  | 74 | -0.22 | -0.686 | 0.246 | 0.35 |
| pufa |  | 74 | -0.22 | -0.687 | 0.247 | 0.351 |
| 1_hdl_1 |  | 74 | 0.22 | -0.247 | 0.686 | 0.351 |
| 1_hdl_p |  | 74 | 0.217 | -0.25 | 0.683 | 0.358 |
| sm |  | 74 | -0.214 | -0.681 | 0.252 | 0.363 |
| m_hdl_1 |  | 74 | 0.211 | -0.254 | 0.675 | 0.369 |
| s_ldl_fc |  | 74 | -0.206 | -0.673 | 0.261 | 0.382 |
| 1_hdl_c |  | 74 | 0.206 | -0.262 | 0.673 | 0.383 |
| serum_c |  | 74 | -0.204 | -0.671 | 0.264 | 0.388 |
| val |  | 74 | -0.2 | -0.662 | 0.261 | 0.39 |
| cit |  | 74 | -0.2 | -0.671 | 0.27 | 0.399 |
| phe |  | 74 | -0.198 | -0.663 | 0.268 | 0.4 |
| 1_hdl_ce |  | 74 | 0.198 | -0.27 | 0.665 | 0.402 |
| m_hdl_p |  | 74 | 0.194 | -0.271 | 0.659 | 0.408 |
| 1_ldl_p |  | 74 | -0.194 | -0.661 | 0.273 | 0.41 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit 95\%CI | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| glol |  | 73 | 0.191 | -0.274 | 0.656 | 0.416 |
| faw6 |  | 74 | -0.189 | -0.657 | 0.278 | 0.422 |
| s_ldl_p |  | 74 | -0.188 | -0.655 | 0.279 | 0.425 |
| m_hdl_pl |  | 74 | 0.184 | -0.28 | 0.648 | 0.432 |
| m_ldl_fc |  | 74 | -0.18 | -0.647 | 0.287 | 0.445 |
| estc |  | 74 | -0.18 | -0.647 | 0.287 | 0.445 |
| 1_1dl_1 |  | 74 | -0.179 | -0.646 | 0.288 | 0.447 |
| s_ldl_1 |  | 74 | -0.178 | -0.645 | 0.29 | 0.451 |
| pyr |  | 74 | -0.175 | -0.636 | 0.286 | 0.452 |
| xl_hdl_ce |  | 74 | -0.176 | -0.647 | 0.294 | 0.457 |
| 1_ldl_ce |  | 74 | -0.174 | -0.641 | 0.293 | 0.46 |
| m_ldl_p |  | 74 | -0.174 | -0.641 | 0.293 | 0.46 |
| xl_hdl_pl |  | 74 | 0.174 | -0.295 | 0.643 | 0.462 |
| m_ldl_1 |  | 74 | -0.17 | -0.637 | 0.297 | 0.471 |
| alb |  | 74 | -0.17 | -0.64 | 0.3 | 0.472 |
| 1_ldl_pl |  | 74 | -0.164 | -0.631 | 0.303 | 0.486 |
| s_hdl_1 |  | 74 | 0.159 | -0.308 | 0.626 | 0.5 |
| 1_ldl_c |  | 74 | -0.154 | -0.622 | 0.313 | 0.513 |
| idl_pl |  | 74 | -0.154 | -0.622 | 0.313 | 0.513 |
| xl_hdl_c |  | 74 | -0.152 | $-0.623$ | 0.319 | 0.522 |
| ldl_c |  | 74 | -0.139 | -0.606 | 0.329 | 0.556 |
| hdl_d |  | 74 | 0.138 | -0.331 | 0.607 | 0.559 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | $\begin{aligned} & \text { Upper Limit } \\ & 95 \% \text { CI } \end{aligned}$ | P-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_hdl_p |  | 74 | 0.13 | -0.337 | 0.598 | 0.58 |
| m_ldl_c |  | 74 | -0.128 | -0.595 | 0.34 | 0.587 |
| la |  | 74 | -0.127 | -0.595 | 0.341 | 0.591 |
| ldl_d |  | 74 | -0.121 | -0.585 | 0.344 | 0.606 |
| m_ldl_ce |  | 74 | -0.116 | -0.583 | 0.352 | 0.623 |
| s_ldl_c |  | 74 | -0.116 | -0.583 | 0.352 | 0.624 |
| dha_fa |  | 74 | 0.114 | -0.35 | 0.578 | 0.626 |
| tyr |  | 74 | 0.113 | -0.353 | 0.578 | 0.631 |
| idl_fc |  | 74 | -0.11 | -0.578 | 0.358 | 0.642 |
| glc |  | 74 | -0.108 | -0.578 | 0.362 | 0.648 |
| s_ldl_ce |  | 74 | -0.093 | -0.561 | 0.374 | 0.692 |
| 1_ldl_fc |  | 74 | -0.088 | -0.556 | 0.38 | 0.708 |
| xl_hdl_fc |  | 74 | -0.084 | -0.555 | 0.387 | 0.723 |
| gly |  | 74 | 0.079 | -0.383 | 0.541 | 0.734 |
| gln |  | 74 | 0.062 | -0.408 | 0.531 | 0.794 |
| hdl3_c |  | 74 | 0.054 | -0.416 | 0.525 | 0.818 |
| ala |  | 74 | 0.045 | -0.419 | 0.508 | 0.848 |
| his |  | 74 | -0.04 | -0.505 | 0.425 | 0.865 |
| s_hdl_fc |  | 74 | -0.032 | -0.5 | 0.436 | 0.891 |
| faw3_fa |  | 74 | 0.018 | -0.445 | 0.482 | 0.937 |
| apoa1 |  | 74 | -0.018 | -0.486 | 0.45 | 0.939 |
| xl_hdl_p |  | 74 | 0.009 | -0.462 | 0.48 | 0.97 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit $95 \% \text { CI }$ | P -value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| lac | 74 |  | -0.006 | -0.467 | 0.455 | 0.979 |
| s_hdl_pl |  | 74 | -0.005 | -0.473 | 0.463 | 0.984 |
| xl_hdl_l |  | 74 | 0.004 | -0.467 | 0.475 | 0.987 |

Appendix C Table C 9:Linear regression results of metabolic traits in the brisk walking arm compared to control, adjusted for metabolic trait levels and age (ITT).

| Metabolite | $\mathbf{N}$ | Beta | Lower Limit <br> 95\%CI | Upper Limit <br> 95\%CI | p-value |
| :--- | :--- | :--- | :--- | :--- | :--- |
| ace | 74 | 0.347 | -0.147 | 0.84 | 0.166 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit $95 \% \text { CI }$ | p-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| ala |  | 74 | 0.06 | -0.38 | 0.499 | 0.787 |
| alb |  | 74 | -0.238 | -0.715 | 0.239 | 0.324 |
| apoa1 |  | 74 | -0.318 | -0.652 | 0.016 | 0.062 |
| apob |  | 74 | -0.326 | -0.608 | -0.044 | 0.024 |
| apob_apoa1 |  | 74 | -0.153 | -0.41 | 0.104 | 0.24 |
| bohbut |  | 74 | 0.303 | -0.103 | 0.708 | 0.142 |
| cit |  | 74 | -0.245 | -0.722 | 0.232 | 0.309 |
| crea |  | 74 | 0.056 | -0.233 | 0.344 | 0.702 |
| dha |  | 74 | -0.087 | -0.458 | 0.285 | 0.643 |
| dha_fa |  | 74 | 0.405 | 0.047 | 0.763 | 0.027 |
| estc |  | 74 | $-0.263$ | -0.565 | 0.038 | 0.086 |
| faw3 |  | 74 | -0.259 | -0.604 | 0.085 | 0.138 |
| faw3_fa |  | 74 | 0.268 | -0.093 | 0.628 | 0.144 |
| faw6 |  | 74 | -0.319 | -0.667 | 0.028 | 0.071 |
| faw6_fa |  | 74 | 0.278 | -0.058 | 0.615 | 0.103 |
| freec |  | 74 | -0.387 | -0.677 | -0.098 | 0.01 |
| glc |  | 74 | 0.027 | -0.315 | 0.368 | 0.876 |
| gln |  | 74 | -0.099 | -0.541 | 0.343 | 0.656 |
| glol |  | 73 | -0.026 | -0.521 | 0.468 | 0.916 |
| gly |  | 74 | 0.199 | -0.194 | 0.592 | 0.315 |
| gp |  | 74 | -0.252 | -0.62 | 0.116 | 0.176 |
| hdl_c |  | 74 | -0.158 | -0.444 | 0.127 | 0.273 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit 95\%CI | p-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| hdl_d |  | 74 | -0.283 | -0.534 | -0.032 | 0.028 |
| hdl_tg |  | 74 | -0.506 | -0.827 | -0.184 | 0.003 |
| hdl2_c |  | 74 | -0.149 | -0.427 | 0.128 | 0.287 |
| hdl3_c |  | 74 | -0.227 | -0.603 | 0.148 | 0.232 |
| his |  | 74 | -0.1 | -0.588 | 0.388 | 0.684 |
| idl_c |  | 74 | -0.261 | -0.548 | 0.027 | 0.075 |
| idl_ce |  | 74 | -0.283 | -0.571 | 0.005 | 0.054 |
| idl_fc |  | 74 | -0.192 | -0.481 | 0.097 | 0.189 |
| idl_1 |  | 74 | -0.282 | -0.562 | -0.002 | 0.048 |
| idl_p |  | 74 | -0.298 | -0.575 | -0.02 | 0.036 |
| idl_pl |  | 74 | -0.217 | -0.497 | 0.064 | 0.128 |
| idl_tg |  | 74 | -0.463 | -0.729 | -0.197 | 0.001 |
| ile |  | 74 | -0.325 | -0.687 | 0.037 | 0.077 |
| 1_hdl_c |  | 74 | -0.23 | -0.451 | -0.01 | 0.041 |
| 1_hdl_ce |  | 74 | -0.241 | -0.463 | -0.02 | 0.033 |
| 1_hdl_fc |  | 74 | -0.196 | -0.413 | 0.021 | 0.076 |
| 1_hdl_1 |  | 74 | -0.229 | -0.454 | -0.003 | 0.047 |
| 1_hdl_p |  | 74 | -0.234 | -0.461 | -0.008 | 0.043 |
| 1_hdl_pl |  | 74 | -0.181 | -0.423 | 0.06 | 0.138 |
| 1_hdl_tg |  | 74 | -0.565 | -0.904 | -0.225 | 0.001 |
| 1_ldl_c |  | 74 | -0.199 | -0.491 | 0.092 | 0.177 |
| 1_ldl_ce |  | 74 | -0.21 | -0.501 | 0.081 | 0.154 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit $95 \% \text { CI }$ | p-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1_ldl_fc |  | 74 | -0.161 | -0.455 | 0.133 | 0.278 |
| 1_ldl_1 |  | 74 | -0.23 | -0.516 | 0.057 | 0.114 |
| 1_1dl_p |  | 74 | -0.244 | -0.529 | 0.04 | 0.091 |
| 1_ldl_pl |  | 74 | -0.219 | -0.508 | 0.07 | 0.135 |
| 1_ldl_tg |  | 74 | -0.496 | -0.771 | -0.22 | 0.001 |
| 1_vldl_c |  | 74 | -0.32 | -0.624 | -0.015 | 0.04 |
| 1_vldl_ce |  | 74 | -0.309 | -0.616 | -0.003 | 0.048 |
| 1_vldl_fc |  | 74 | -0.329 | -0.632 | -0.025 | 0.034 |
| 1_vldl_1 |  | 74 | -0.294 | -0.6 | 0.011 | 0.059 |
| 1_vldl_p |  | 74 | -0.291 | -0.597 | 0.015 | 0.062 |
| 1_vldl_pl |  | 74 | -0.303 | -0.607 | 0.001 | 0.051 |
| 1_vldl_tg |  | 74 | -0.281 | -0.588 | 0.025 | 0.071 |
| la |  | 74 | -0.292 | -0.651 | 0.067 | 0.109 |
| la_fa |  | 74 | 0.232 | -0.124 | 0.589 | 0.198 |
| lac |  | 74 | 0.101 | -0.381 | 0.583 | 0.677 |
| ldl_c |  | 74 | -0.177 | -0.475 | 0.122 | 0.241 |
| ldl_d |  | 74 | -0.178 | -0.639 | 0.284 | 0.445 |
| ldl_tg |  | 74 | -0.496 | -0.773 | -0.219 | 0.001 |
| leu |  | 74 | -0.323 | -0.693 | 0.047 | 0.087 |
| m_hdl_c |  | 74 | 0.004 | -0.416 | 0.423 | 0.986 |
| m_hdl_ce |  | 74 | 0.017 | -0.406 | 0.44 | 0.937 |
| m_hdl_fc |  | 74 | -0.049 | -0.45 | 0.352 | 0.808 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit 95\%CI | p-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_hdl_1 |  | 74 | -0.1 | -0.504 | 0.304 | 0.624 |
| m_hdl_p |  | 74 | -0.118 | -0.52 | 0.283 | 0.559 |
| m_hdl_pl |  | 74 | -0.162 | -0.55 | 0.226 | 0.407 |
| m_hdl_tg |  | 74 | -0.354 | -0.669 | -0.039 | 0.028 |
| m_ldl_c |  | 74 | -0.157 | -0.46 | 0.146 | 0.306 |
| m_ldl_ce |  | 74 | -0.144 | -0.446 | 0.158 | 0.346 |
| m_ldl_fc |  | 74 | -0.212 | -0.522 | 0.097 | 0.176 |
| m_ldl_1 |  | 74 | -0.204 | -0.5 | 0.091 | 0.172 |
| m_ldl_p |  | 74 | -0.211 | -0.504 | 0.083 | 0.157 |
| m_ldl_pl |  | 74 | -0.275 | -0.566 | 0.017 | 0.064 |
| m_ldl_tg |  | 74 | -0.485 | -0.771 | -0.2 | 0.001 |
| m_vldl_c |  | 74 | -0.319 | -0.617 | -0.021 | 0.036 |
| m_vldl_ce |  | 74 | -0.324 | -0.625 | -0.023 | 0.035 |
| m_vldl_fc |  | 74 | -0.308 | -0.603 | -0.014 | 0.04 |
| m_vldl_1 |  | 74 | -0.284 | -0.581 | 0.013 | 0.061 |
| m_vldl_p |  | 74 | -0.279 | -0.577 | 0.018 | 0.065 |
| m_vldl_pl |  | 74 | -0.293 | -0.588 | 0.001 | 0.051 |
| m_vldl_tg |  | 74 | -0.263 | -0.56 | 0.034 | 0.082 |
| mufa |  | 74 | -0.354 | -0.647 | -0.06 | 0.019 |
| mufa_fa |  | 74 | -0.108 | -0.41 | 0.195 | 0.481 |
| pc |  | 74 | -0.472 | -0.782 | -0.161 | 0.003 |
| phe |  | 74 | -0.037 | -0.474 | 0.399 | 0.865 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit $95 \% \text { CI }$ | p-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| pufa |  | 74 | -0.323 | -0.667 | 0.02 | 0.065 |
| pufa_fa |  | 74 | 0.307 | -0.01 | 0.624 | 0.057 |
| pyr |  | 74 | -0.137 | -0.591 | 0.317 | 0.549 |
| remnant_c |  | 74 | -0.341 | -0.617 | -0.066 | 0.016 |
| s_hdl_c |  | 74 | 0.253 | -0.144 | 0.649 | 0.208 |
| s_hdl_ce |  | 74 | 0.285 | -0.096 | 0.666 | 0.14 |
| s_hdl_fc |  | 74 | -0.208 | -0.615 | 0.199 | 0.312 |
| s_hdl_1 |  | 74 | -0.001 | -0.42 | 0.417 | 0.996 |
| s_hdl_p |  | 74 | -0.032 | -0.445 | 0.381 | 0.878 |
| s_hdl_pl |  | 74 | -0.2 | -0.561 | 0.161 | 0.272 |
| s_hdl_tg |  | 74 | -0.331 | -0.616 | -0.046 | 0.023 |
| s_ldl_c |  | 74 | -0.149 | -0.458 | 0.16 | 0.34 |
| s_ldl_ce |  | 74 | -0.128 | -0.434 | 0.178 | 0.408 |
| s_ldl_fc |  | 74 | -0.234 | -0.557 | 0.09 | 0.154 |
| s_ldl_1 |  | 74 | -0.215 | -0.519 | 0.09 | 0.164 |
| s_ldl_p |  | 74 | -0.226 | -0.528 | 0.075 | 0.139 |
| s_ldl_pl |  | 74 | -0.298 | -0.61 | 0.014 | 0.061 |
| s_ldl_tg |  | 74 | -0.476 | -0.759 | -0.193 | 0.001 |
| s_vldl_c |  | 74 | -0.265 | -0.541 | 0.012 | 0.06 |
| s_vldl_ce |  | 74 | -0.224 | -0.506 | 0.058 | 0.118 |
| s_vldl_fc |  | 74 | -0.305 | -0.579 | -0.032 | 0.029 |
| s_vldl_1 |  | 74 | -0.272 | -0.55 | 0.006 | 0.055 |


| Metabolite | N |  | Beta | Lower Limit $95 \% \text { CI }$ | Upper Limit 95\%СI | p-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_vldl_p |  | 74 | -0.269 | -0.548 | 0.01 | 0.059 |
| s_vldl_pl |  | 74 | -0.275 | -0.551 | 0.001 | 0.051 |
| s_vldl_tg |  | 74 | -0.258 | -0.544 | 0.029 | 0.077 |
| serum_c |  | 74 | -0.301 | -0.599 | -0.004 | 0.047 |
| serum_tg |  | 74 | -0.328 | -0.621 | -0.035 | 0.029 |
| sfa |  | 74 | -0.512 | -0.814 | -0.211 | 0.001 |
| sfa_fa |  | 74 | -0.359 | -0.813 | 0.094 | 0.118 |
| sm |  | 74 | -0.296 | -0.618 | 0.025 | 0.07 |
| tg_pg |  | 74 | -0.168 | -0.445 | 0.109 | 0.231 |
| totcho |  | 74 | -0.453 | -0.769 | -0.136 | 0.006 |
| totfa |  | 74 | -0.429 | -0.733 | -0.125 | 0.006 |
| totpg |  | 74 | -0.481 | -0.797 | -0.166 | 0.003 |
| tyr |  | 74 | 0.069 | -0.399 | 0.537 | 0.77 |
| unsat |  | 74 | 0.3 | -0.079 | 0.679 | 0.119 |
| val |  | 74 | -0.093 | -0.533 | 0.348 | 0.676 |
| vldl_c |  | 74 | -0.33 | -0.612 | -0.048 | 0.023 |
| vldl_d |  | 74 | -0.25 | -0.555 | 0.054 | 0.105 |
| vldl_tg |  | 74 | -0.284 | -0.58 | 0.012 | 0.06 |
| xl_hdl_c |  | 74 | -0.395 | -0.669 | -0.121 | 0.005 |
| xl_hdl_ce |  | 74 | -0.385 | -0.656 | -0.114 | 0.006 |
| xl_hdl_fc |  | 74 | -0.411 | -0.69 | -0.132 | 0.005 |
| xl_hdl_1 |  | 74 | -0.328 | -0.589 | -0.067 | 0.015 |


| Metabolite | N |  | Beta | Lower Limit 95\%CI | Upper Limit 95\%CI | p-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_hdl_p |  | 74 | -0.325 | -0.586 | -0.064 | 0.015 |
| xl_hdl_pl |  | 74 | -0.217 | -0.455 | 0.02 | 0.073 |
| xl_hdl_tg |  | 74 | -0.523 | -0.845 | -0.202 | 0.002 |
| xl_vldl_c |  | 74 | -0.352 | -0.664 | -0.041 | 0.027 |
| xl_vldl_ce |  | 74 | -0.345 | -0.658 | -0.031 | 0.032 |
| xl_vldl_fc |  | 74 | -0.362 | -0.672 | -0.053 | 0.023 |
| xl_vldl_1 |  | 74 | -0.335 | -0.643 | -0.026 | 0.034 |
| xl_vldl_p |  | 74 | -0.332 | -0.64 | -0.024 | 0.035 |
| xl_vldl_pl |  | 74 | -0.351 | -0.658 | -0.043 | 0.026 |
| xl_vldl_tg |  | 74 | -0.323 | -0.631 | -0.016 | 0.04 |
| xs_vldl_c |  | 74 | -0.268 | -0.551 | 0.015 | 0.063 |
| xs_vldl_ce |  | 74 | -0.272 | -0.562 | 0.017 | 0.065 |
| xs_vldl_fc |  | 74 | -0.256 | -0.542 | 0.03 | 0.078 |
| xs_vldl_1 |  | 74 | -0.309 | -0.578 | -0.041 | 0.025 |
| xs_vldl_p |  | 74 | -0.321 | -0.588 | -0.055 | 0.019 |
| xs_vldl_pl |  | 74 | -0.232 | -0.508 | 0.043 | 0.097 |
| xs_vldl_tg |  | 74 | -0.359 | -0.623 | -0.095 | 0.008 |
| xxl_vldl_c |  | 74 | -0.378 | -0.696 | -0.06 | 0.02 |
| xxl_vldl_ce |  | 74 | -0.38 | -0.707 | -0.054 | 0.023 |
| xxl_vldl_fc |  | 74 | -0.374 | -0.682 | -0.065 | 0.018 |
| xxl_vldl_1 |  | 74 | -0.37 | -0.684 | -0.056 | 0.022 |
| xxl_vldl_p |  | 74 | -0.368 | -0.682 | -0.054 | 0.022 |


| Metabolite | N |  | Beta | Lower Limit95\%CI |  | Upper Limit $95 \% \mathrm{CI}$ |  | p-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xxl_vldl_pl | $74 \quad-0.381$ |  |  | -0.692 |  |  | -0.07 | 0.017 |
| xxl_vldl_tg |  | 74 | -0.365 |  | 0.679 |  | -0.052 | 0.023 |

Appendix C Table C 10: Linear regression results of the instrumented exposure (dairy milk intake) on metabolic traits in the plant-based diet arm compared to control (IV).

| Metabolite | $\mathbf{N}$ | Beta | Standard <br> error | p-value | R2 | F-statistic |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :---: |
| 0 |  |  |  |  |  |  |  |


| ala | 44 | 0.178 | 0.14 | 0.201 | 0.4331945 | 32.0995 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| alb | 44 | 0.172 | 0.159 | 0.278 | 0.4331945 | 32.0995 |
| apoa1 | 44 | 0.11 | 0.146 | 0.452 | 0.4331945 | 32.0995 |
| apob | 44 | 0.158 | 0.135 | 0.241 | 0.4331945 | 32.0995 |
| apob_apoa1 | 44 | 0.117 | 0.137 | 0.394 | 0.4331945 | 32.0995 |
| bohbut | 44 | 0.084 | 0.16 | 0.599 | 0.4331945 | 32.0995 |
| cit | 44 | -0.172 | 0.145 | 0.235 | 0.4331945 | 32.0995 |
| crea | 44 | -0.077 | 0.151 | 0.609 | 0.4331945 | 32.0995 |
| dha | 44 | 0.116 | 0.145 | 0.423 | 0.4331945 | 32.0995 |
| dha_fa | 44 | -0.058 | 0.137 | 0.671 | 0.4331945 | 32.0995 |
| estc | 44 | 0.146 | 0.133 | 0.273 | 0.4331945 | 32.0995 |
| faw3 | 44 | 0.115 | 0.142 | 0.418 | 0.4331945 | 32.0995 |
| faw3_fa | 44 | -0.08 | 0.149 | 0.588 | 0.4331945 | 32.0995 |
| faw6 | 44 | 0.052 | 0.121 | 0.665 | 0.4331945 | 32.0995 |
| faw6_fa | 44 | -0.245 | 0.119 | 0.04 | 0.4331945 | 32.0995 |
| freec | 44 | 0.125 | 0.129 | 0.33 | 0.4331945 | 32.0995 |
| glc | 44 | 0.303 | 0.156 | 0.052 | 0.4331945 | 32.0995 |
| $g 1 n$ | 44 | -0.157 | 0.116 | 0.177 | 0.4331945 | 32.0995 |
| glol | 43 | 0.15 | 0.126 | 0.237 | 0.471923 | 36.64019 |
| gly | 44 | -0.138 | 0.14 | 0.324 | 0.4331945 | 32.0995 |
| gp | 44 | 0.108 | 0.121 | 0.374 | 0.4331945 | 32.0995 |
| hdl_c | 44 | 0.013 | 0.144 | 0.925 | 0.4331945 | 32.0995 |
| hdl_d | 44 | -0.105 | 0.124 | 0.397 | 0.4331945 | 32.0995 |
| hdl_tg | 44 | 0.098 | 0.119 | 0.412 | 0.4331945 | 32.0995 |


| hdl2_c | 44 | 0.004 | 0.144 | 0.98 | 0.4331945 | 32.0995 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| hdl3_c | 44 | 0.112 | 0.138 | 0.415 | 0.4331945 | 32.0995 |
| his | 44 | -0.113 | 0.116 | 0.333 | 0.4331945 | 32.0995 |
| idl_c | 44 | 0.129 | 0.131 | 0.322 | 0.4331945 | 32.0995 |
| idl_ce | 44 | 0.138 | 0.132 | 0.295 | 0.4331945 | 32.0995 |
| idl_fc | 44 | 0.101 | 0.127 | 0.427 | 0.4331945 | 32.0995 |
| idl_1 | 44 | 0.132 | 0.13 | 0.31 | 0.4331945 | 32.0995 |
| idl_p | 44 | 0.136 | 0.131 | 0.299 | 0.4331945 | 32.0995 |
| idl_pl | 44 | 0.121 | 0.13 | 0.35 | 0.4331945 | 32.0995 |
| idl_tg | 44 | 0.133 | 0.12 | 0.269 | 0.4331945 | 32.0995 |
| ile | 44 | 0.082 | 0.122 | 0.5 | 0.4331945 | 32.0995 |
| 1_hdl_c | 44 | -0.087 | 0.136 | 0.524 | 0.4331945 | 32.0995 |
| l_hdl_ce | 44 | -0.087 | 0.136 | 0.524 | 0.4331945 | 32.0995 |
| l_hdl_fc | 44 | -0.087 | 0.136 | 0.522 | 0.4331945 | 32.0995 |
| 1_hdl_1 | 44 | -0.065 | 0.136 | 0.635 | 0.4331945 | 32.0995 |
| 1_hdl_p | 44 | -0.062 | 0.136 | 0.648 | 0.4331945 | 32.0995 |
| l_hdl_pl | 44 | -0.036 | 0.138 | 0.796 | 0.4331945 | 32.0995 |
| l_hdl_tg | 44 | -0.062 | 0.097 | 0.523 | 0.4331945 | 32.0995 |
| 1_ldl_c | 44 | 0.117 | 0.129 | 0.366 | 0.4331945 | 32.0995 |
| 1_ldl_ce | 44 | 0.122 | 0.13 | 0.347 | 0.4331945 | 32.0995 |
| 1_ldl_fc | 44 | 0.098 | 0.127 | 0.442 | 0.4331945 | 32.0995 |
| 1_1dl_1 | 44 | 0.124 | 0.13 | 0.34 | 0.4331945 | 32.0995 |
| 1_ldl_p | 44 | 0.128 | 0.13 | 0.328 | 0.4331945 | 32.0995 |
| 1_ldl_pl | 44 | 0.133 | 0.132 | 0.312 | 0.4331945 | 32.0995 |


| 1_ldl_tg | 44 | 0.134 | 0.118 | 0.257 | 0.4331945 | 32.0995 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| l_vldl_c | 44 | 0.102 | 0.128 | 0.425 | 0.4331945 | 32.0995 |
| 1_vldl_ce | 44 | 0.106 | 0.129 | 0.412 | 0.4331945 | 32.0995 |
| l_vldl_fc | 44 | 0.098 | 0.127 | 0.44 | 0.4331945 | 32.0995 |
| 1_vldl_1 | 44 | 0.105 | 0.129 | 0.415 | 0.4331945 | 32.0995 |
| 1_vldl_p | 44 | 0.106 | 0.129 | 0.413 | 0.4331945 | 32.0995 |
| l_vldl_pl | 44 | 0.109 | 0.13 | 0.402 | 0.4331945 | 32.0995 |
| 1_vldl_tg | 44 | 0.106 | 0.13 | 0.414 | 0.4331945 | 32.0995 |
| la | 44 | 0.017 | 0.118 | 0.884 | 0.4331945 | 32.0995 |
| la_fa | 44 | -0.254 | 0.111 | 0.023 | 0.4331945 | 32.0995 |
| lac | 44 | 0.24 | 0.125 | 0.055 | 0.4331945 | 32.0995 |
| ldl_c | 44 | 0.117 | 0.129 | 0.363 | 0.4331945 | 32.0995 |
| ldl_d | 44 | -0.07 | 0.1 | 0.483 | 0.4331945 | 32.0995 |
| ldl_tg | 44 | 0.139 | 0.12 | 0.248 | 0.4331945 | 32.0995 |
| leu | 44 | 0.133 | 0.117 | 0.254 | 0.4331945 | 32.0995 |
| m_hdl_c | 44 | 0.14 | 0.149 | 0.346 | 0.4331945 | 32.0995 |
| m_hdl_ce | 44 | 0.136 | 0.148 | 0.357 | 0.4331945 | 32.0995 |
| m_hdl_fc | 44 | 0.156 | 0.153 | 0.308 | 0.4331945 | 32.0995 |
| m_hdl_1 | 44 | 0.182 | 0.154 | 0.237 | 0.4331945 | 32.0995 |
| m_hdl_p | 44 | 0.189 | 0.154 | 0.219 | 0.4331945 | 32.0995 |
| m_hdl_pl | 44 | 0.191 | 0.155 | 0.219 | 0.4331945 | 32.0995 |
| m_hdl_tg | 44 | 0.19 | 0.136 | 0.161 | 0.4331945 | 32.0995 |
| m_ldl_c | 44 | 0.117 | 0.129 | 0.364 | 0.4331945 | 32.0995 |
| m_ldl_ce | 44 | 0.113 | 0.129 | 0.378 | 0.4331945 | 32.0995 |


| m_ldl_fc | 44 | 0.131 | 0.129 | 0.311 | 0.4331945 | 32.0995 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_ldl_1 | 44 | 0.13 | 0.13 | 0.318 | 0.4331945 | 32.0995 |
| m_ldl_p | 44 | 0.13 | 0.131 | 0.32 | 0.4331945 | 32.0995 |
| m_ldl_pl | 44 | 0.161 | 0.134 | 0.229 | 0.4331945 | 32.0995 |
| m_ldl_tg | 44 | 0.132 | 0.118 | 0.266 | 0.4331945 | 32.0995 |
| m_vldl_c | 44 | 0.119 | 0.132 | 0.364 | 0.4331945 | 32.0995 |
| m_vldl_ce | 44 | 0.129 | 0.134 | 0.333 | 0.4331945 | 32.0995 |
| m_vldl_fc | 44 | 0.107 | 0.129 | 0.408 | 0.4331945 | 32.0995 |
| m_vldl_1 | 44 | 0.11 | 0.131 | 0.402 | 0.4331945 | 32.0995 |
| m_vldl_p | 44 | 0.109 | 0.131 | 0.406 | 0.4331945 | 32.0995 |
| m_vldl_pl | 44 | 0.113 | 0.131 | 0.388 | 0.4331945 | 32.0995 |
| m_vldl_tg | 44 | 0.103 | 0.13 | 0.428 | 0.4331945 | 32.0995 |
| mufa | 44 | 0.127 | 0.126 | 0.314 | 0.4331945 | 32.0995 |
| mufa_fa | 44 | 0.088 | 0.132 | 0.503 | 0.4331945 | 32.0995 |
| pc | 44 | 0.159 | 0.13 | 0.224 | 0.4331945 | 32.0995 |
| phe | 44 | 0.115 | 0.121 | 0.345 | 0.4331945 | 32.0995 |
| pufa | 44 | 0.063 | 0.123 | 0.611 | 0.4331945 | 32.0995 |
| pufa_fa | 44 | -0.236 | 0.127 | 0.063 | 0.4331945 | 32.0995 |
| pyr | 44 | 0.321 | 0.138 | 0.02 | 0.4331945 | 32.0995 |
| remnant_c | 44 | 0.15 | 0.134 | 0.263 | 0.4331945 | 32.0995 |
| s_hdl_c | 44 | 0.176 | 0.129 | 0.174 | 0.4331945 | 32.0995 |
| s_hdl_ce | 44 | 0.138 | 0.124 | 0.266 | 0.4331945 | 32.0995 |
| s_hdl_fc | 44 | 0.21 | 0.137 | 0.126 | 0.4331945 | 32.0995 |
| s_hdl_1 | 44 | 0.291 | 0.144 | 0.044 | 0.4331945 | 32.0995 |


| s_hdl_p | 44 | 0.297 | 0.145 | 0.04 | 0.4331945 | 32.0995 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_hdl_pl | 44 | 0.223 | 0.137 | 0.104 | 0.4331945 | 32.0995 |
| s_hdl_tg | 44 | 0.119 | 0.127 | 0.347 | 0.4331945 | 32.0995 |
| s_ldl_c | 44 | 0.118 | 0.128 | 0.36 | 0.4331945 | 32.0995 |
| s_ldl_ce | 44 | 0.11 | 0.128 | 0.389 | 0.4331945 | 32.0995 |
| s_ldl_fc | 44 | 0.145 | 0.129 | 0.26 | 0.4331945 | 32.0995 |
| s_ldl_1 | 44 | 0.139 | 0.13 | 0.287 | 0.4331945 | 32.0995 |
| s_ldl_p | 44 | 0.142 | 0.131 | 0.278 | 0.4331945 | 32.0995 |
| s_ldl_pl | 44 | 0.179 | 0.134 | 0.181 | 0.4331945 | 32.0995 |
| s_ldl_tg | 44 | 0.15 | 0.126 | 0.234 | 0.4331945 | 32.0995 |
| s_vldl_c | 44 | 0.163 | 0.136 | 0.232 | 0.4331945 | 32.0995 |
| s_vldl_ce | 44 | 0.159 | 0.136 | 0.244 | 0.4331945 | 32.0995 |
| s_vldl_fc | 44 | 0.153 | 0.133 | 0.25 | 0.4331945 | 32.0995 |
| s_vldl_1 | 44 | 0.142 | 0.134 | 0.292 | 0.4331945 | 32.0995 |
| s_vldl_p | 44 | 0.137 | 0.134 | 0.307 | 0.4331945 | 32.0995 |
| s_vldl_pl | 44 | 0.157 | 0.134 | 0.242 | 0.4331945 | 32.0995 |
| s_vldl_tg | 44 | 0.111 | 0.131 | 0.398 | 0.4331945 | 32.0995 |
| serum_c | 44 | 0.141 | 0.132 | 0.287 | 0.4331945 | 32.0995 |
| serum_tg | 44 | 0.113 | 0.13 | 0.383 | 0.4331945 | 32.0995 |
| sfa | 44 | 0.192 | 0.132 | 0.146 | 0.4331945 | 32.0995 |
| sfa_fa | 44 | 0.248 | 0.132 | 0.06 | 0.4331945 | 32.0995 |
| sm | 44 | 0.104 | 0.132 | 0.429 | 0.4331945 | 32.0995 |
| tg_pg | 44 | 0.062 | 0.125 | 0.617 | 0.4331945 | 32.0995 |
| totcho | 44 | 0.144 | 0.129 | 0.264 | 0.4331945 | 32.0995 |


| totfa | 44 | 0.14 | 0.129 | 0.277 | 0.4331945 | 32.0995 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| totpg | 44 | 0.138 | 0.129 | 0.282 | 0.4331945 | 32.0995 |
| tyr | 44 | 0.138 | 0.13 | 0.289 | 0.4331945 | 32.0995 |
| unsat | 44 | -0.136 | 0.142 | 0.337 | 0.4331945 | 32.0995 |
| val | 44 | 0.075 | 0.121 | 0.534 | 0.4331945 | 32.0995 |
| vldl_c | 44 | 0.14 | 0.134 | 0.295 | 0.4331945 | 32.0995 |
| vldl_d | 44 | 0.096 | 0.131 | 0.464 | 0.4331945 | 32.0995 |
| vldl_tg | 44 | 0.106 | 0.13 | 0.414 | 0.4331945 | 32.0995 |
| xl_hdl_c | 44 | -0.09 | 0.12 | 0.453 | 0.4331945 | 32.0995 |
| xl_hdl_ce | 44 | -0.087 | 0.119 | 0.465 | 0.4331945 | 32.0995 |
| xl_hdl_fc | 44 | -0.096 | 0.124 | 0.439 | 0.4331945 | 32.0995 |
| xl_hdl_1 | 44 | -0.094 | 0.123 | 0.442 | 0.4331945 | 32.0995 |
| xl_hdl_p | 44 | -0.094 | 0.122 | 0.445 | 0.4331945 | 32.0995 |
| xl_hdl_pl | 44 | -0.095 | 0.125 | 0.446 | 0.4331945 | 32.0995 |
| xl_hdl_tg | 44 | 0.016 | 0.114 | 0.885 | 0.4331945 | 32.0995 |
| xl_vldl_c | 44 | 0.092 | 0.126 | 0.463 | 0.4331945 | 32.0995 |
| xl_vldl_ce | 44 | 0.094 | 0.126 | 0.455 | 0.4331945 | 32.0995 |
| xl_vldl_fc | 44 | 0.09 | 0.125 | 0.471 | 0.4331945 | 32.0995 |
| xl_vldl_1 | 44 | 0.093 | 0.126 | 0.464 | 0.4331945 | 32.0995 |
| xl_vldl_p | 44 | 0.093 | 0.127 | 0.464 | 0.4331945 | 32.0995 |
| xl_vldl_pl | 44 | 0.093 | 0.126 | 0.461 | 0.4331945 | 32.0995 |
| xl_vldl_tg | 44 | 0.092 | 0.127 | 0.466 | 0.4331945 | 32.0995 |
| xs_vldl_c | 44 | 0.139 | 0.13 | 0.284 | 0.4331945 | 32.0995 |
| xs_vldl_ce | 44 | 0.132 | 0.13 | 0.307 | 0.4331945 | 32.0995 |


| xs_vldl_fc | 44 | 0.15 | 0.131 | 0.25 | 0.4331945 | 32.0995 |
| :--- | :---: | :---: | :---: | :---: | :--- | :---: |
| xs_vldl_l | 44 | 0.156 | 0.133 | 0.24 | 0.4331945 | 32.0995 |
| xs_vldl_p | 44 | 0.157 | 0.133 | 0.237 | 0.4331945 | 32.0995 |
| xs_vldl_pl | 44 | 0.157 | 0.135 | 0.243 | 0.4331945 | 32.0995 |
| xs_vldl_tg | 44 | 0.134 | 0.128 | 0.294 | 0.4331945 | 32.0995 |
| xxl_vldl_c | 44 | 0.099 | 0.126 | 0.436 | 0.4331945 | 32.0995 |
| xxl_vldl_ce | 44 | 0.113 | 0.129 | 0.381 | 0.4331945 | 32.0995 |
| xxl_vldl_fc | 44 | 0.08 | 0.123 | 0.518 | 0.4331945 | 32.0995 |
| xxl_vldl_1 | 44 | 0.092 | 0.125 | 0.465 | 0.4331945 | 32.0995 |
| xxl_vldl_p | 44 | 0.092 | 0.125 | 0.461 | 0.4331945 | 32.0995 |
| xxl_vldl_pl | 44 | 0.089 | 0.124 | 0.475 | 0.4331945 | 32.0995 |
| xxl_vldl_tg | 44 | 0.09 | 0.125 | 0.471 | 0.4331945 | 32.0995 |

Appendix C Table C 11:Linear regression results of the instrumented exposure (dairy milk) on metabolic traits in the plant-based diet arm compared to control, adjusted for baseline metabolic traits (IV).

| Metabolite | N | Beta | Standard <br> error | p- <br> value | R2 | F-statistic |
| :--- | :--- | :--- | ---: | :--- | :--- | :--- |
| hdl3_c | 44 | 0.277 | 0.139 | 0.047 | 0.3953819 | 26.8114 |
| pyr | 44 | 0.274 | 0.125 | 0.029 | 0.4466715 | 33.09703 |
| idl_fc | 44 | 0.234 | 0.097 | 0.016 | 0.4122619 | 28.75897 |
| lac | 44 | 0.234 | 0.123 | 0.057 | 0.4455559 | 32.94794 |
| sm | 44 | 0.229 | 0.108 | 0.034 | 0.4144611 | 29.02097 |
| l_ldl_fc | 44 | 0.226 | 0.097 | 0.019 | 0.4132081 | 28.87145 |
| sfa_fa | 44 | 0.225 | 0.13 | 0.084 | 0.4257499 | 30.39746 |
| s_hdl_c | 44 | 0.223 | 0.119 | 0.062 | 0.4268661 | 30.53652 |
| s_hdl_ce | 44 | 0.213 | 0.117 | 0.067 | 0.4178801 | 29.43223 |


| Metabolite | N | Beta | Standard error | $\begin{gathered} \mathrm{p}- \\ \text { value } \end{gathered}$ | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_ldl_ce | 44 | 0.211 | 0.094 | 0.025 | 0.4206975 | 29.77477 |
| s_ldl_c | 44 | 0.211 | 0.094 | 0.026 | 0.4224861 | 29.99396 |
| idl_pl | 44 | 0.21 | 0.089 | 0.019 | 0.4260794 | 30.43845 |
| idl_c | 44 | 0.21 | 0.092 | 0.022 | 0.4275242 | 30.61875 |
| estc | 44 | 0.21 | 0.096 | 0.028 | 0.4348331 | 31.54494 |
| s_ldl_fc | 44 | 0.209 | 0.095 | 0.029 | 0.4298179 | 30.90685 |
| hdl_c | 44 | 0.207 | 0.107 | 0.054 | 0.4062469 | 28.05227 |
| ldl_c | 44 | 0.206 | 0.092 | 0.026 | 0.4245986 | 30.2546 |
| m_ldl_ce | 44 | 0.205 | 0.092 | 0.026 | 0.4233517 | 30.10053 |
| m_ldl_c | 44 | 0.205 | 0.093 | 0.027 | 0.4241426 | 30.19818 |
| m_ldl_fc | 44 | 0.205 | 0.094 | 0.028 | 0.4275805 | 30.62579 |
| 1_ldl_c | 44 | 0.203 | 0.091 | 0.026 | 0.4258222 | 30.40645 |
| apoa1 | 44 | 0.202 | 0.112 | 0.072 | 0.4391585 | 32.10443 |
| idl_1 | 44 | 0.197 | 0.088 | 0.024 | 0.4319519 | 31.17699 |
| idl_ce | 44 | 0.197 | 0.09 | 0.029 | 0.4321744 | 31.20527 |
| 1_ldl_1 | 44 | 0.196 | 0.089 | 0.027 | 0.4300038 | 30.9303 |
| 1_ldl_pl | 44 | 0.196 | 0.09 | 0.029 | 0.4321541 | 31.20268 |
| serum_c | 44 | 0.196 | 0.093 | 0.035 | 0.4377052 | 31.91549 |
| 1_ldl_ce | 44 | 0.194 | 0.089 | 0.03 | 0.4292319 | 30.83302 |
| s_ldl_1 | 44 | 0.194 | 0.091 | 0.032 | 0.432736 | 31.27676 |
| hdl2_c | 44 | 0.193 | 0.102 | 0.06 | 0.4081798 | 28.2778 |
| m_ldl_l | 44 | 0.192 | 0.089 | 0.031 | 0.4312154 | 31.08353 |


| Metabolite | N | Beta | Standard error | $\begin{gathered} \text { p- } \\ \text { value } \end{gathered}$ | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1_ldl_p | 44 | 0.191 | 0.087 | 0.029 | 0.4322277 | 31.21205 |
| idl_p | 44 | 0.189 | 0.086 | 0.027 | 0.4346905 | 31.52665 |
| s_ldl_p | 44 | 0.189 | 0.089 | 0.035 | 0.4347679 | 31.53658 |
| m_ldl_p | 44 | 0.186 | 0.088 | 0.034 | 0.432613 | 31.26109 |
| unsat | 44 | 0.186 | 0.13 | 0.151 | 0.3486218 | 21.94346 |
| xs_vldl_fc | 44 | 0.185 | 0.084 | 0.028 | 0.4361833 | 31.71867 |
| s_hdl_l | 44 | 0.185 | 0.115 | 0.107 | 0.4421414 | 32.49532 |
| totcho | 44 | 0.183 | 0.098 | 0.062 | 0.452145 | 33.83733 |
| alb | 44 | 0.181 | 0.16 | 0.256 | 0.4343546 | 31.48357 |
| xs_vldl_c | 44 | 0.18 | 0.081 | 0.026 | 0.4331618 | 31.33104 |
| glc | 44 | 0.179 | 0.083 | 0.031 | 0.4307023 | 31.01856 |
| 1_hdl_pl | 44 | 0.179 | 0.098 | 0.066 | 0.398338 | 27.14458 |
| m_hdl_ce | 44 | 0.177 | 0.107 | 0.098 | 0.4432268 | 32.6386 |
| m_hdl_c | 44 | 0.177 | 0.108 | 0.101 | 0.4449431 | 32.8663 |
| xs_vldl_ce | 44 | 0.176 | 0.081 | 0.03 | 0.4321235 | 31.19879 |
| m_hdl_fc | 44 | 0.176 | 0.111 | 0.113 | 0.4514681 | 33.74496 |
| s_hdl_p | 44 | 0.173 | 0.114 | 0.127 | 0.4392973 | 32.12253 |
| s_ldl_pl | 44 | 0.171 | 0.09 | 0.056 | 0.4455881 | 32.95224 |
| xs_vldl_pl | 44 | 0.17 | 0.081 | 0.036 | 0.4405584 | 32.28736 |
| l_hdl_p | 44 | 0.166 | 0.089 | 0.064 | 0.3956916 | 26.84616 |
| l_hdl_1 | 44 | 0.165 | 0.089 | 0.063 | 0.3951126 | 26.78121 |
| freec | 44 | 0.158 | 0.087 | 0.07 | 0.4450737 | 32.88368 |


| Metabolite | N | Beta | Standard error | $\begin{gathered} \mathrm{p}- \\ \text { value } \end{gathered}$ | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_ldl_pl | 44 | 0.155 | 0.085 | 0.069 | 0.4437824 | 32.71215 |
| 1_hdl_fc | 44 | 0.155 | 0.085 | 0.07 | 0.3912421 | 26.35026 |
| 1_hdl_c | 44 | 0.154 | 0.085 | 0.072 | 0.3918176 | 26.41398 |
| 1_hdl_ce | 44 | 0.153 | 0.085 | 0.073 | 0.3920242 | 26.4369 |
| m_hdl_1 | 44 | 0.148 | 0.105 | 0.159 | 0.463427 | 35.41085 |
| pc | 44 | 0.146 | 0.09 | 0.107 | 0.4732842 | 36.84084 |
| hdl_d | 44 | 0.143 | 0.09 | 0.11 | 0.3791396 | 25.03739 |
| m_hdl_p | 44 | 0.142 | 0.104 | 0.175 | 0.4665844 | 35.86315 |
| totpg | 44 | 0.141 | 0.092 | 0.124 | 0.4679519 | 36.0607 |
| xl_hdl_pl | 44 | 0.136 | 0.086 | 0.111 | 0.3859295 | 25.76758 |
| m_hdl_pl | 44 | 0.131 | 0.105 | 0.21 | 0.469613 | 36.30204 |
| xl_hdl_p | 44 | 0.121 | 0.083 | 0.147 | 0.3890944 | 26.11348 |
| xl_hdl_1 | 44 | 0.12 | 0.083 | 0.151 | 0.389439 | 26.15136 |
| dha | 44 | 0.117 | 0.103 | 0.253 | 0.4493037 | 33.45121 |
| xs_vldl_1 | 44 | 0.111 | 0.074 | 0.131 | 0.4432969 | 32.64787 |
| tyr | 44 | 0.109 | 0.119 | 0.356 | 0.4343787 | 31.48666 |
| xl_hdl_ce | 44 | 0.104 | 0.084 | 0.219 | 0.3916964 | 26.40056 |
| xl_hdl_c | 44 | 0.098 | 0.083 | 0.237 | 0.3943909 | 26.70043 |
| faw3 | 44 | 0.093 | 0.1 | 0.354 | 0.4484496 | 33.33591 |
| sfa | 44 | 0.09 | 0.087 | 0.298 | 0.4580914 | 34.65852 |
| xs_vldl_p | 44 | 0.087 | 0.072 | 0.228 | 0.4438465 | 32.72065 |
| bohbut | 44 | 0.079 | 0.152 | 0.604 | 0.4346111 | 31.51645 |


| Metabolite | N | Beta | Standard error | $\begin{gathered} \mathrm{p}- \\ \text { value } \end{gathered}$ | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_hdl_fc | 44 | 0.078 | 0.081 | 0.335 | 0.4034889 | 27.73301 |
| pufa | 44 | 0.078 | 0.092 | 0.401 | 0.4477955 | 33.24786 |
| phe | 44 | 0.074 | 0.118 | 0.533 | 0.4312194 | 31.08403 |
| faw6 | 44 | 0.071 | 0.091 | 0.436 | 0.4462429 | 33.03968 |
| s_hdl_fc | 44 | 0.068 | 0.104 | 0.517 | 0.4349411 | 31.55881 |
| val | 44 | 0.068 | 0.11 | 0.537 | 0.4358146 | 31.67114 |
| s_vldl_ce | 44 | 0.058 | 0.078 | 0.452 | 0.437852 | 31.93453 |
| remnant_c | 44 | 0.058 | 0.079 | 0.463 | 0.4434764 | 32.67163 |
| apob | 44 | 0.057 | 0.081 | 0.479 | 0.4435565 | 32.68223 |
| totfa | 44 | 0.055 | 0.085 | 0.517 | 0.4575347 | 34.58088 |
| leu | 44 | 0.055 | 0.106 | 0.605 | 0.4173559 | 29.36886 |
| pufa_fa | 44 | 0.045 | 0.096 | 0.635 | 0.36082 | 23.14468 |
| 1_hdl_tg | 44 | 0.037 | 0.079 | 0.64 | 0.4119703 | 28.72437 |
| dha_fa | 44 | 0.035 | 0.094 | 0.709 | 0.4243197 | 30.22008 |
| 1_ldl_tg | 44 | 0.032 | 0.063 | 0.612 | 0.4648437 | 35.61313 |
| m_ldl_tg | 44 | 0.03 | 0.064 | 0.638 | 0.4714417 | 36.5695 |
| ala | 44 | 0.026 | 0.116 | 0.82 | 0.4362438 | 31.72647 |
| faw6_fa | 44 | 0.024 | 0.1 | 0.814 | 0.3502074 | 22.09706 |
| glol | 43 | 0.017 | 0.096 | 0.856 | 0.4631954 | 34.515 |
| la | 44 | 0.016 | 0.089 | 0.857 | 0.4516055 | 33.7637 |
| ldl_tg | 44 | 0.013 | 0.065 | 0.84 | 0.4662082 | 35.80898 |
| s_hdl_pl | 44 | 0.011 | 0.098 | 0.91 | 0.418107 | 29.45969 |


| Metabolite | N | Beta | Standard error | p- <br> value | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| gp | 44 | 0.003 | 0.081 | 0.966 | 0.4317115 | 31.14646 |
| s_vldl_c | 44 | -0.003 | 0.079 | 0.965 | 0.4314 | 31.10693 |
| faw3_fa | 44 | -0.012 | 0.095 | 0.902 | 0.4297984 | 30.9044 |
| mufa | 44 | -0.014 | 0.085 | 0.87 | 0.4412372 | 32.37639 |
| xl_hdl_tg | 44 | -0.018 | 0.095 | 0.851 | 0.4468295 | 33.1182 |
| ile | 44 | -0.026 | 0.11 | 0.814 | 0.4114884 | 28.66728 |
| idl_tg | 44 | -0.029 | 0.064 | 0.651 | 0.4477247 | 33.23834 |
| apob_apoa1 | 44 | -0.043 | 0.079 | 0.583 | 0.4194369 | 29.62109 |
| ldl_d | 44 | -0.045 | 0.083 | 0.586 | 0.4334977 | 31.37393 |
| s_ldl_tg | 44 | -0.049 | 0.076 | 0.523 | 0.4501267 | 33.56263 |
| vldl_c | 44 | -0.059 | 0.087 | 0.499 | 0.4225596 | 30.00299 |
| crea | 44 | -0.073 | 0.095 | 0.442 | 0.4349163 | 31.55562 |
| m_vldl_ce | 44 | -0.099 | 0.097 | 0.308 | 0.4119112 | 28.71737 |
| s_vldl_fc | 44 | -0.104 | 0.088 | 0.238 | 0.4090511 | 28.37994 |
| xxl_vldl_ce | 44 | -0.104 | 0.107 | 0.329 | 0.4052045 | 27.93126 |
| mufa_fa | 44 | -0.107 | 0.107 | 0.316 | 0.3950436 | 26.77348 |
| la_fa | 44 | -0.117 | 0.092 | 0.204 | 0.3933505 | 26.58433 |
| s_vldl_1 | 44 | -0.125 | 0.093 | 0.177 | 0.4023617 | 27.60337 |
| m_vldl_c | 44 | -0.125 | 0.1 | 0.211 | 0.4026999 | 27.64221 |
| xxl_vldl_c | 44 | -0.127 | 0.107 | 0.234 | 0.3994423 | 27.26988 |
| his | 44 | -0.128 | 0.113 | 0.256 | 0.4309428 | 31.04899 |
| xs_vldl_tg | 44 | -0.134 | 0.084 | 0.111 | 0.4051423 | 27.92405 |


| Metabolite | N | Beta | Standard error | $\begin{gathered} \mathrm{p}- \\ \text { value } \end{gathered}$ | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| hdl_tg | 44 | -0.135 | 0.092 | 0.144 | 0.4249783 | 30.30166 |
| s_vldl_p | 44 | -0.137 | 0.095 | 0.149 | 0.398603 | 27.1746 |
| xxl_vldl_p | 44 | -0.138 | 0.099 | 0.164 | 0.3933277 | 26.58179 |
| xxl_vldl_1 | 44 | -0.138 | 0.107 | 0.199 | 0.3938036 | 26.63484 |
| xxl_vldl_pl | 44 | -0.14 | 0.106 | 0.186 | 0.3957091 | 26.84812 |
| xxl_vldl_tg | 44 | -0.14 | 0.108 | 0.193 | 0.3919923 | 26.43335 |
| s_vldl_pl | 44 | -0.143 | 0.093 | 0.125 | 0.3954683 | 26.82109 |
| m_vldl_pl | 44 | -0.147 | 0.103 | 0.151 | 0.3932114 | 26.56884 |
| xl_vldl_fc | 44 | -0.147 | 0.106 | 0.167 | 0.3930947 | 26.55585 |
| xl_vldl_c | 44 | -0.147 | 0.108 | 0.171 | 0.3906596 | 26.28587 |
| xl_vldl_ce | 44 | -0.147 | 0.109 | 0.177 | 0.3887039 | 26.07061 |
| xl_vldl_pl | 44 | -0.148 | 0.105 | 0.159 | 0.3942265 | 26.68206 |
| m_vldl_fc | 44 | -0.149 | 0.102 | 0.146 | 0.3935158 | 26.60275 |
| xxl_vldl_fc | 44 | -0.15 | 0.105 | 0.153 | 0.3931969 | 26.56723 |
| m_vldl_1 | 44 | -0.152 | 0.104 | 0.146 | 0.3915355 | 26.38273 |
| 1_vldl_fc | 44 | -0.153 | 0.105 | 0.145 | 0.3908449 | 26.30634 |
| m_vldl_p | 44 | -0.154 | 0.105 | 0.142 | 0.3905273 | 26.27127 |
| 1_vldl_pl | 44 | -0.154 | 0.106 | 0.146 | 0.3897569 | 26.18633 |
| xl_vldl_p | 44 | -0.155 | 0.107 | 0.148 | 0.3883365 | 26.03032 |
| xl_vldl_1 | 44 | -0.155 | 0.107 | 0.15 | 0.3887317 | 26.07366 |
| 1_vldl_c | 44 | -0.156 | 0.107 | 0.145 | 0.3881364 | 26.00839 |
| serum_tg | 44 | -0.157 | 0.101 | 0.12 | 0.3934363 | 26.59389 |


| Metabolite | N | Beta | Standard <br> error | p- <br> value | R2 | F-statistic |
| :--- | :---: | :--- | :---: | :---: | :---: | :---: |
| l_vldl_p | 44 | -0.158 | 0.107 | 0.14 | 0.3870689 | 25.89169 |
| l_vldl_l | 44 | -0.158 | 0.107 | 0.14 | 0.3873281 | 25.91999 |
| xl_vldl_tg | 44 | -0.158 | 0.108 | 0.141 | 0.3869252 | 25.87601 |
| l_vldl_ce | 44 | -0.158 | 0.109 | 0.148 | 0.3858475 | 25.75866 |
| l_vldl_tg | 44 | -0.16 | 0.107 | 0.137 | 0.3863369 | 25.81191 |
| tg_pg | 44 | -0.162 | 0.101 | 0.106 | 0.3849002 | 25.65585 |
| m_vldl_tg | 44 | -0.164 | 0.106 | 0.124 | 0.3864093 | 25.81979 |
| vldl_tg | 44 | -0.166 | 0.106 | 0.116 | 0.3864867 | 25.82822 |
| m_hdl_tg | 44 | -0.174 | 0.106 | 0.102 | 0.3835804 | 25.51313 |
| gly | 44 | -0.175 | 0.108 | 0.103 | 0.4372215 | 31.85282 |
| s_vldl_tg | 44 | -0.181 | 0.104 | 0.083 | 0.3833653 | 25.48993 |
| cit | 44 | -0.205 | 0.134 | 0.126 | 0.4295515 | 30.87327 |
| gln | 44 | -0.221 | 0.102 | 0.03 | 0.4435737 | 32.68451 |
| s_hdl_tg | 44 | -0.233 | 0.108 | 0.031 | 0.3547233 | 22.53863 |
| vldl_d | 44 | -0.266 | 0.127 | 0.035 | 0.3371741 | 20.85636 |
| ( | 44 | -0.34 | 0.151 | 0.024 | 0.389095 | 26.11354 |

Appendix C Table C 12: Linear regression results of the instrumented exposure (daily fruit and veg portions) on metabolic traits in the plant-based diet arm compared to control (IV)

| Metabolite | N | Standard <br> error |  |  |  |  |  | p-value | R2 | F-statistic |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: | :---: | :---: | :---: | :---: |
| pyr | 44 | -0.263 | 0.15 | 0.08 | 0.1045089 | 4.90164 |  |  |  |  |
| ace | 44 | 0.245 | 0.144 | 0.088 | 0.1045089 | 4.90164 |  |  |  |  |
| s_hdl_p | 44 | -0.243 | 0.153 | 0.113 | 0.1045089 | 4.90164 |  |  |  |  |
| s_hdl_l | 44 | -0.238 | 0.151 | 0.114 | 0.1045089 | 4.90164 |  |  |  |  |
| lac | 44 | -0.196 | 0.125 | 0.115 | 0.1045089 | 4.90164 |  |  |  |  |
| sfa_fa | 44 | -0.203 | 0.129 | 0.116 | 0.1045089 | 4.90164 |  |  |  |  |
| glc | 44 | -0.248 | 0.158 | 0.116 | 0.1045089 | 4.90164 |  |  |  |  |
| la_fa | 44 | 0.208 | 0.136 | 0.127 | 0.1045089 | 4.90164 |  |  |  |  |
| faw6_fa | 44 | 0.2 | 0.136 | 0.14 | 0.1045089 | 4.90164 |  |  |  |  |
| pufa_fa | 44 | 0.193 | 0.14 | 0.167 | 0.1045089 | 4.90164 |  |  |  |  |
| s_hdl_c | 44 | -0.144 | 0.111 | 0.193 | 0.1045089 | 4.90164 |  |  |  |  |
| s_hdl_pl | 44 | -0.182 | 0.141 | 0.196 | 0.1045089 | 4.90164 |  |  |  |  |
| sfa | 44 | -0.157 | 0.123 | 0.2 | 0.1045089 | 4.90164 |  |  |  |  |


| Metabolite | N | Beta | Standard error | p-value | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_hdl_fc | 44 | -0.172 | 0.135 | 0.203 | 0.1045089 | 4.90164 |
| s_ldl_pl | 44 | -0.147 | 0.116 | 0.208 | 0.1045089 | 4.90164 |
| m_hdl_tg | 44 | -0.156 | 0.133 | 0.244 | 0.1045089 | 4.90164 |
| m_ldl_pl | 44 | -0.132 | 0.115 | 0.252 | 0.1045089 | 4.90164 |
| ala | 44 | -0.146 | 0.128 | 0.253 | 0.1045089 | 4.90164 |
| pc | 44 | -0.13 | 0.114 | 0.255 | 0.1045089 | 4.90164 |
| xs_vldl_pl | 44 | -0.128 | 0.115 | 0.265 | 0.1045089 | 4.90164 |
| s_ldl_fc | 44 | -0.119 | 0.107 | 0.266 | 0.1045089 | 4.90164 |
| m_hdl_pl | 44 | -0.156 | 0.141 | 0.267 | 0.1045089 | 4.90164 |
| gln | 44 | 0.128 | 0.116 | 0.267 | 0.1045089 | 4.90164 |
| xs_vldl_fc | 44 | -0.123 | 0.111 | 0.268 | 0.1045089 | 4.90164 |
| m_hdl_p | 44 | -0.155 | 0.14 | 0.269 | 0.1045089 | 4.90164 |
| s_hdl_ce | 44 | -0.113 | 0.102 | 0.269 | 0.1045089 | 4.90164 |
| xs_vldl_p | 44 | -0.129 | 0.117 | 0.271 | 0.1045089 | 4.90164 |
| xs_vldl_l | 44 | -0.127 | 0.116 | 0.272 | 0.1045089 | 4.90164 |
| s_ldl_tg | 44 | -0.122 | 0.112 | 0.276 | 0.1045089 | 4.90164 |
| ldl_tg | 44 | -0.113 | 0.104 | 0.276 | 0.1045089 | 4.90164 |
| s_vldl_c | 44 | -0.133 | 0.123 | 0.278 | 0.1045089 | 4.90164 |
| s_vldl_ce | 44 | -0.13 | 0.12 | 0.281 | 0.1045089 | 4.90164 |
| apob | 44 | -0.13 | 0.12 | 0.281 | 0.1045089 | 4.90164 |
| 1_ldl_tg | 44 | -0.109 | 0.102 | 0.282 | 0.1045089 | 4.90164 |
| estc | 44 | -0.12 | 0.111 | 0.282 | 0.1045089 | 4.90164 |


| Metabolite | N | Beta | Standard error | p-value | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| glol | 43 | -0.122 | 0.114 | 0.282 | 0.1127513 | 5.210268 |
| m_hdl_1 | 44 | -0.149 | 0.139 | 0.283 | 0.1045089 | 4.90164 |
| cit | 44 | 0.141 | 0.131 | 0.283 | 0.1045089 | 4.90164 |
| m_ldl_tg | 44 | -0.108 | 0.101 | 0.285 | 0.1045089 | 4.90164 |
| s_ldl_p | 44 | -0.116 | 0.109 | 0.287 | 0.1045089 | 4.90164 |
| totcho | 44 | -0.118 | 0.111 | 0.288 | 0.1045089 | 4.90164 |
| s_ldl_1 | 44 | -0.114 | 0.108 | 0.293 | 0.1045089 | 4.90164 |
| serum_c | 44 | -0.115 | 0.11 | 0.296 | 0.1045089 | 4.90164 |
| remnant_c | 44 | -0.123 | 0.118 | 0.299 | 0.1045089 | 4.90164 |
| s_vldl_pl | 44 | -0.128 | 0.124 | 0.301 | 0.1045089 | 4.90164 |
| leu | 44 | -0.109 | 0.106 | 0.302 | 0.1045089 | 4.90164 |
| xs_vldl_c | 44 | -0.114 | 0.111 | 0.303 | 0.1045089 | 4.90164 |
| idl_tg | 44 | -0.109 | 0.106 | 0.303 | 0.1045089 | 4.90164 |
| s_vldl_fc | 44 | -0.125 | 0.122 | 0.304 | 0.1045089 | 4.90164 |
| idl_ce | 44 | -0.113 | 0.111 | 0.309 | 0.1045089 | 4.90164 |
| tyr | 44 | -0.113 | 0.111 | 0.309 | 0.1045089 | 4.90164 |
| idl_p | 44 | -0.111 | 0.109 | 0.31 | 0.1045089 | 4.90164 |
| totpg | 44 | -0.113 | 0.112 | 0.311 | 0.1045089 | 4.90164 |
| alb | 44 | -0.141 | 0.139 | 0.312 | 0.1045089 | 4.90164 |
| m_ldl_fc | 44 | -0.107 | 0.106 | 0.313 | 0.1045089 | 4.90164 |
| totfa | 44 | -0.114 | 0.114 | 0.314 | 0.1045089 | 4.90164 |
| idl_1 | 44 | -0.108 | 0.109 | 0.319 | 0.1045089 | 4.90164 |


| Metabolite | N | Beta | Standard <br> error |  | p-value | R2 |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | F-statistic


| Metabolite | N | Beta | Standard error | p-value | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_hdl_c | 44 | -0.115 | 0.129 | 0.374 | 0.1045089 | 4.90164 |
| m_vldl_ce | 44 | -0.106 | 0.12 | 0.376 | 0.1045089 | 4.90164 |
| m_ldl_ce | 44 | -0.093 | 0.105 | 0.377 | 0.1045089 | 4.90164 |
| s_ldl_ce | 44 | -0.09 | 0.104 | 0.385 | 0.1045089 | 4.90164 |
| m_hdl_ce | 44 | -0.111 | 0.128 | 0.385 | 0.1045089 | 4.90164 |
| unsat | 44 | 0.111 | 0.128 | 0.386 | 0.1045089 | 4.90164 |
| phe | 44 | -0.094 | 0.109 | 0.389 | 0.1045089 | 4.90164 |
| s_hdl_tg | 44 | -0.097 | 0.114 | 0.394 | 0.1045089 | 4.90164 |
| hdl3_c | 44 | -0.092 | 0.108 | 0.395 | 0.1045089 | 4.90164 |
| m_vldl_c | 44 | -0.098 | 0.117 | 0.406 | 0.1045089 | 4.90164 |
| xxl_vldl_ce | 44 | -0.092 | 0.114 | 0.419 | 0.1045089 | 4.90164 |
| faw3 | 44 | -0.094 | 0.117 | 0.42 | 0.1045089 | 4.90164 |
| idl_fc | 44 | -0.083 | 0.102 | 0.421 | 0.1045089 | 4.90164 |
| serum_tg | 44 | -0.093 | 0.116 | 0.423 | 0.1045089 | 4.90164 |
| apob_apoa1 | 44 | -0.095 | 0.119 | 0.423 | 0.1045089 | 4.90164 |
| gp | 44 | -0.088 | 0.11 | 0.423 | 0.1045089 | 4.90164 |
| sm | 44 | -0.085 | 0.107 | 0.427 | 0.1045089 | 4.90164 |
| m_vldl_pl | 44 | -0.092 | 0.117 | 0.428 | 0.1045089 | 4.90164 |
| dha | 44 | -0.095 | 0.121 | 0.432 | 0.1045089 | 4.90164 |
| 1_ldl_fc | 44 | -0.08 | 0.102 | 0.435 | 0.1045089 | 4.90164 |
| s_vldl_tg | 44 | -0.091 | 0.117 | 0.438 | 0.1045089 | 4.90164 |
| m_vldl_1 | 44 | -0.09 | 0.116 | 0.441 | 0.1045089 | 4.90164 |


| Metabolite | N | Beta | Standard error | p-value | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| l_vldl_pl | 44 | -0.089 | 0.116 | 0.442 | 0.1045089 | 4.90164 |
| hdl_tg | 44 | -0.08 | 0.104 | 0.442 | 0.1045089 | 4.90164 |
| m_vldl_p | 44 | -0.089 | 0.116 | 0.445 | 0.1045089 | 4.90164 |
| m_vldl_fc | 44 | -0.087 | 0.114 | 0.446 | 0.1045089 | 4.90164 |
| hdl_d | 44 | 0.086 | 0.113 | 0.446 | 0.1045089 | 4.90164 |
| l_vldl_ce | 44 | -0.087 | 0.115 | 0.449 | 0.1045089 | 4.90164 |
| 1_vldl_p | 44 | -0.087 | 0.115 | 0.452 | 0.1045089 | 4.90164 |
| vldl_tg | 44 | -0.087 | 0.115 | 0.452 | 0.1045089 | 4.90164 |
| apoa1 | 44 | -0.09 | 0.12 | 0.452 | 0.1045089 | 4.90164 |
| l_vldl_1 | 44 | -0.086 | 0.115 | 0.453 | 0.1045089 | 4.90164 |
| l_vldl_tg | 44 | -0.087 | 0.115 | 0.453 | 0.1045089 | 4.90164 |
| l_vldl_c | 44 | -0.084 | 0.113 | 0.461 | 0.1045089 | 4.90164 |
| m_vldl_tg | 44 | -0.085 | 0.116 | 0.465 | 0.1045089 | 4.90164 |
| xxl_vldl_c | 44 | -0.081 | 0.111 | 0.468 | 0.1045089 | 4.90164 |
| 1_vldl_fc | 44 | -0.08 | 0.112 | 0.474 | 0.1045089 | 4.90164 |
| xl_hdl_fc | 44 | 0.078 | 0.11 | 0.478 | 0.1045089 | 4.90164 |
| ldl_d | 44 | 0.057 | 0.081 | 0.48 | 0.1045089 | 4.90164 |
| xl_hdl_1 | 44 | 0.077 | 0.11 | 0.483 | 0.1045089 | 4.90164 |
| xl_hdl_pl | 44 | 0.078 | 0.112 | 0.484 | 0.1045089 | 4.90164 |
| xl_hdl_p | 44 | 0.077 | 0.11 | 0.485 | 0.1045089 | 4.90164 |
| xl_vldl_ce | 44 | -0.077 | 0.111 | 0.486 | 0.1045089 | 4.90164 |
| xxl_vldl_p | 44 | -0.076 | 0.11 | 0.491 | 0.1045089 | 4.90164 |


| Metabolite | N | Beta | Standard error | p-value | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_vldl_pl | 44 | -0.076 | 0.111 | 0.492 | 0.1045089 | 4.90164 |
| xl_hdl_c | 44 | 0.074 | 0.107 | 0.492 | 0.1045089 | 4.90164 |
| xl_vldl_c | 44 | -0.075 | 0.11 | 0.493 | 0.1045089 | 4.90164 |
| xxl_vldl_1 | 44 | -0.075 | 0.109 | 0.494 | 0.1045089 | 4.90164 |
| xl_vldl_p | 44 | -0.076 | 0.111 | 0.495 | 0.1045089 | 4.90164 |
| xl_vldl_1 | 44 | -0.076 | 0.111 | 0.495 | 0.1045089 | 4.90164 |
| xl_vldl_tg | 44 | -0.076 | 0.111 | 0.498 | 0.1045089 | 4.90164 |
| xxl_vldl_tg | 44 | -0.074 | 0.109 | 0.499 | 0.1045089 | 4.90164 |
| vldl_d | 44 | -0.078 | 0.116 | 0.499 | 0.1045089 | 4.90164 |
| xl_vldl_fc | 44 | -0.074 | 0.109 | 0.501 | 0.1045089 | 4.90164 |
| xxl_vldl_pl | 44 | -0.073 | 0.108 | 0.503 | 0.1045089 | 4.90164 |
| xl_hdl_ce | 44 | 0.071 | 0.106 | 0.503 | 0.1045089 | 4.90164 |
| ile | 44 | -0.067 | 0.105 | 0.523 | 0.1045089 | 4.90164 |
| xxl_vldl_fc | 44 | -0.065 | 0.107 | 0.541 | 0.1045089 | 4.90164 |
| mufa_fa | 44 | -0.072 | 0.119 | 0.542 | 0.1045089 | 4.90164 |
| 1_hdl_ce | 44 | 0.071 | 0.118 | 0.549 | 0.1045089 | 4.90164 |
| 1_hdl_fc | 44 | 0.071 | 0.119 | 0.549 | 0.1045089 | 4.90164 |
| 1_hdl_c | 44 | 0.071 | 0.118 | 0.55 | 0.1045089 | 4.90164 |
| 1_hdl_tg | 44 | 0.051 | 0.085 | 0.55 | 0.1045089 | 4.90164 |
| val | 44 | -0.062 | 0.105 | 0.555 | 0.1045089 | 4.90164 |
| bohbut | 44 | -0.069 | 0.132 | 0.601 | 0.1045089 | 4.90164 |
| faw3_fa | 44 | 0.066 | 0.128 | 0.608 | 0.1045089 | 4.90164 |


| Metabolite | N | Beta | Standard <br> error | p-value | R2 | F-statistic |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| pufa | 44 | -0.051 | 0.101 | 0.61 | 0.1045089 | 4.90164 |
| crea | 44 | 0.063 | 0.126 | 0.617 | 0.1045089 | 4.90164 |
| tg_pg | 44 | -0.051 | 0.107 | 0.632 | 0.1045089 | 4.90164 |
| l_hdl_l | 44 | 0.053 | 0.117 | 0.65 | 0.1045089 | 4.90164 |
| l_hdl_p | 44 | 0.051 | 0.116 | 0.661 | 0.1045089 | 4.90164 |
| faw6 | 44 | -0.043 | 0.098 | 0.664 | 0.1045089 | 4.90164 |
| dha_fa | 44 | 0.048 | 0.115 | 0.68 | 0.1045089 | 4.90164 |
| l_hdl_pl | 44 | 0.029 | 0.116 | 0.8 | 0.1045089 | 4.90164 |
| la | 44 | -0.014 | 0.096 | 0.883 | 0.1045089 | 4.90164 |
| xl_hdl_tg | 44 | -0.013 | 0.093 | 0.885 | 0.1045089 | 4.90164 |
| hdl_c | 44 | -0.011 | 0.117 | 0.925 | 0.1045089 | 4.90164 |
| hdl2_c | 44 | -0.003 | 0.118 | 0.98 | 0.1045089 | 4.90164 |

Appendix C Table C 13: Linear regression results of the instrumented exposure (daily fruit and veg portions) on metabolic traits in the plant-based diet arm compared to control (IV), adjusted for baseline metabolic traits.

| Metabolite | N | Beta | Standard error | P-value | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| idl_fc | 44 | -0.17 | 0.09 | 0.07 | 0.11 | 5.17 |
| pyr | 44 | -0.21 | 0.12 | 0.07 | 0.12 | 5.83 |
| 1_ldl_fc | 44 | -0.17 | 0.09 | 0.07 | 0.11 | 5.22 |
| idl_pl | 44 | -0.16 | 0.09 | 0.07 | 0.11 | 4.97 |
| s_ldl_c | 44 | -0.16 | 0.09 | 0.08 | 0.11 | 5.08 |
| s_ldl_fc | 44 | -0.16 | 0.09 | 0.08 | 0.11 | 4.92 |
| s_ldl_ce | 44 | -0.16 | 0.09 | 0.08 | 0.11 | 5.13 |
| idl_c | 44 | -0.16 | 0.09 | 0.08 | 0.11 | 4.86 |
| estc | 44 | -0.17 | 0.10 | 0.08 | 0.10 | 4.75 |
| hdl3_c | 44 | -0.20 | 0.11 | 0.08 | 0.11 | 5.03 |
| ldl_c | 44 | -0.16 | 0.09 | 0.08 | 0.11 | 5.00 |
| m_ldl_fc | 44 | -0.16 | 0.09 | 0.08 | 0.11 | 4.98 |
| m_ldl_c | 44 | -0.16 | 0.09 | 0.08 | 0.11 | 5.03 |
| m_ldl_ce | 44 | -0.16 | 0.09 | 0.08 | 0.11 | 5.05 |
| 1_ldl_c | 44 | -0.16 | 0.09 | 0.08 | 0.11 | 4.95 |


| idl_1 | 44 | -0.16 | 0.09 | 0.08 | 0.11 | 4.85 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1_1dl_1 | 44 | -0.15 | 0.09 | 0.08 | 0.11 | 4.90 |
| glc | 44 | -0.15 | 0.08 | 0.08 | 0.10 | 4.76 |
| sm | 44 | -0.18 | 0.10 | 0.08 | 0.10 | 4.52 |
| s_ldl_1 | 44 | -0.15 | 0.09 | 0.08 | 0.11 | 4.89 |
| idl_p | 44 | -0.15 | 0.09 | 0.09 | 0.11 | 4.83 |
| 1_ldl_pl | 44 | -0.16 | 0.09 | 0.09 | 0.11 | 4.87 |
| 1_ldl_ce | 44 | -0.15 | 0.09 | 0.09 | 0.11 | 4.89 |
| 1_1dl_p | 44 | -0.15 | 0.09 | 0.09 | 0.11 | 4.88 |
| m_ldl_1 | 44 | -0.15 | 0.09 | 0.09 | 0.11 | 4.92 |
| idl_ce | 44 | -0.16 | 0.09 | 0.09 | 0.10 | 4.80 |
| serum_c | 44 | -0.16 | 0.09 | 0.09 | 0.10 | 4.75 |
| s_ldl_p | 44 | -0.15 | 0.09 | 0.09 | 0.11 | 4.86 |
| vldl_d | 44 | 0.13 | 0.08 | 0.09 | 0.19 | 9.78 |
| xs_vldl_fc | 44 | -0.15 | 0.09 | 0.09 | 0.11 | 4.88 |
| m_ldl_p | 44 | -0.15 | 0.09 | 0.09 | 0.11 | 4.90 |
| xs_vldl_c | 44 | -0.14 | 0.09 | 0.09 | 0.11 | 4.89 |
| s_hdl_tg | 44 | 0.14 | 0.08 | 0.09 | 0.14 | 6.74 |
| xs_vldl_pl | 44 | -0.14 | 0.08 | 0.10 | 0.10 | 4.81 |
| xs_vldl_ce | 44 | -0.14 | 0.09 | 0.10 | 0.11 | 4.89 |
| ace | 44 | 0.28 | 0.17 | 0.10 | 0.09 | 3.87 |
| s_ldl_pl | 44 | -0.14 | 0.09 | 0.11 | 0.10 | 4.79 |
| s_hdl_ce | 44 | -0.15 | 0.09 | 0.11 | 0.14 | 6.44 |
| s_hdl_c | 44 | -0.17 | 0.11 | 0.11 | 0.12 | 5.75 |


| totcho | 44 | -0.15 | 0.09 | 0.11 | 0.10 | 4.62 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| hdl_c | 44 | -0.16 | 0.10 | 0.12 | 0.10 | 4.54 |
| m_ldl_pl | 44 | -0.13 | 0.08 | 0.12 | 0.10 | 4.78 |
| apoa1 | 44 | -0.16 | 0.11 | 0.12 | 0.10 | 4.46 |
| freec | 44 | -0.13 | 0.08 | 0.12 | 0.10 | 4.75 |
| lac | 44 | -0.20 | 0.13 | 0.12 | 0.10 | 4.57 |
| 1_hdl_1 | 44 | -0.13 | 0.08 | 0.13 | 0.09 | 4.24 |
| 1_hdl_p | 44 | -0.13 | 0.08 | 0.13 | 0.09 | 4.23 |
| hdl2_c | 44 | -0.15 | 0.10 | 0.13 | 0.10 | 4.51 |
| gln | 44 | 0.20 | 0.13 | 0.13 | 0.09 | 4.13 |
| 1_hdl_pl | 44 | -0.14 | 0.09 | 0.13 | 0.10 | 4.42 |
| 1_hdl_fc | 44 | -0.12 | 0.08 | 0.13 | 0.09 | 4.27 |
| 1_hdl_c | 44 | -0.12 | 0.08 | 0.13 | 0.09 | 4.17 |
| l_hdl_ce | 44 | -0.12 | 0.08 | 0.13 | 0.09 | 4.14 |
| sfa_fa | 44 | -0.19 | 0.13 | 0.14 | 0.10 | 4.52 |
| gly | 44 | 0.14 | 0.10 | 0.14 | 0.10 | 4.75 |
| s_vldl_tg | 44 | 0.11 | 0.08 | 0.15 | 0.16 | 7.54 |
| hdl_d | 44 | -0.11 | 0.08 | 0.15 | 0.09 | 3.93 |
| pc | 44 | -0.12 | 0.08 | 0.15 | 0.11 | 4.91 |
| m_hdl_tg | 44 | 0.10 | 0.07 | 0.16 | 0.18 | 8.72 |
| totpg | 44 | -0.12 | 0.08 | 0.16 | 0.11 | 4.96 |
| xl_hdl_pl | 44 | -0.10 | 0.08 | 0.17 | 0.09 | 4.19 |
| tg_pg | 44 | 0.11 | 0.08 | 0.17 | 0.14 | 6.47 |
| m_hdl_ce | 44 | -0.14 | 0.11 | 0.18 | 0.10 | 4.69 |


| m_hdl_c | 44 | -0.14 | 0.11 | 0.18 | 0.10 | 4.71 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| vldl_tg | 44 | 0.10 | 0.07 | 0.18 | 0.16 | 8.00 |
| cit | 44 | 0.17 | 0.13 | 0.18 | 0.10 | 4.51 |
| xs_vldl_1 | 44 | -0.09 | 0.07 | 0.18 | 0.10 | 4.77 |
| m_hdl_fc | 44 | -0.14 | 0.11 | 0.18 | 0.10 | 4.76 |
| m_vldl_tg | 44 | 0.10 | 0.07 | 0.19 | 0.16 | 7.95 |
| xl_hdl_p | 44 | -0.09 | 0.07 | 0.19 | 0.09 | 4.07 |
| serum_tg | 44 | 0.10 | 0.07 | 0.19 | 0.16 | 7.86 |
| s_hdl_1 | 44 | -0.16 | 0.12 | 0.19 | 0.10 | 4.44 |
| xl_hdl_1 | 44 | -0.09 | 0.07 | 0.19 | 0.09 | 4.07 |
| xs_vldl_tg | 44 | 0.09 | 0.07 | 0.19 | 0.13 | 6.27 |
| 1_vldl_tg | 44 | 0.09 | 0.07 | 0.20 | 0.17 | 8.44 |
| xl_vldl_tg | 44 | 0.09 | 0.07 | 0.20 | 0.18 | 8.72 |
| 1_vldl_p | 44 | 0.09 | 0.07 | 0.20 | 0.17 | 8.38 |
| 1_vldl_1 | 44 | 0.09 | 0.07 | 0.20 | 0.17 | 8.37 |
| s_vldl_pl | 44 | 0.10 | 0.07 | 0.20 | 0.14 | 6.46 |
| xl_vldl_p | 44 | 0.09 | 0.07 | 0.20 | 0.17 | 8.64 |
| l_vldl_fc | 44 | 0.09 | 0.07 | 0.20 | 0.17 | 8.40 |
| m_vldl_p | 44 | 0.09 | 0.07 | 0.20 | 0.16 | 7.80 |
| 1_vldl_c | 44 | 0.09 | 0.07 | 0.20 | 0.17 | 8.21 |
| s_hdl_p | 44 | -0.15 | 0.12 | 0.20 | 0.10 | 4.48 |
| hdl_tg | 44 | 0.08 | 0.07 | 0.20 | 0.18 | 9.20 |
| xl_vldl_1 | 44 | 0.09 | 0.07 | 0.21 | 0.17 | 8.61 |
| 1_vldl_pl | 44 | 0.09 | 0.07 | 0.21 | 0.17 | 8.28 |


| l_vldl_ce | 44 | 0.09 | 0.08 | 0.21 | 0.16 | 7.98 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_vldl_1 | 44 | 0.09 | 0.07 | 0.21 | 0.16 | 7.77 |
| m_vldl_fc | 44 | 0.09 | 0.07 | 0.21 | 0.16 | 7.85 |
| xxl_vldl_fc | 44 | 0.09 | 0.07 | 0.21 | 0.17 | 8.64 |
| m_vldl_pl | 44 | 0.09 | 0.07 | 0.21 | 0.16 | 7.69 |
| xl_vldl_pl | 44 | 0.09 | 0.07 | 0.21 | 0.17 | 8.50 |
| unsat | 44 | -0.12 | 0.09 | 0.22 | 0.12 | 5.85 |
| s_vldl_p | 44 | 0.09 | 0.07 | 0.22 | 0.14 | 6.67 |
| m_hdl_l | 44 | -0.12 | 0.10 | 0.22 | 0.11 | 4.95 |
| xl_vldl_fc | 44 | 0.09 | 0.07 | 0.22 | 0.17 | 8.41 |
| xxl_vldl_p | 44 | 0.08 | 0.07 | 0.22 | 0.17 | 8.54 |
| xl_vldl_c | 44 | 0.09 | 0.07 | 0.22 | 0.17 | 8.32 |
| xl_vldl_ce | 44 | 0.09 | 0.07 | 0.23 | 0.17 | 8.22 |
| m_hdl_p | 44 | -0.12 | 0.10 | 0.23 | 0.11 | 5.03 |
| xxl_vldl_pl | 44 | 0.08 | 0.07 | 0.24 | 0.17 | 8.59 |
| xl_hdl_ce | 44 | -0.08 | 0.07 | 0.24 | 0.09 | 4.17 |
| la_fa | 44 | 0.08 | 0.07 | 0.25 | 0.13 | 6.27 |
| xxl_vldl_tg | 44 | 0.08 | 0.07 | 0.25 | 0.17 | 8.61 |
| s_vldl_l | 44 | 0.08 | 0.07 | 0.25 | 0.14 | 6.49 |
| xxl_vldl_1 | 44 | 0.08 | 0.07 | 0.25 | 0.17 | 8.51 |
| xl_hdl_c | 44 | -0.08 | 0.07 | 0.26 | 0.09 | 4.10 |
| m_hdl_pl | 44 | -0.11 | 0.10 | 0.26 | 0.11 | 5.05 |
| xs_vldl_p | 44 | -0.07 | 0.06 | 0.26 | 0.10 | 4.80 |
| m_vldl_c | 44 | 0.08 | 0.07 | 0.27 | 0.15 | 7.34 |


| his | 44 | 0.10 | 0.09 | 0.28 | 0.11 | 5.01 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xxl_vldl_c | 44 | 0.08 | 0.07 | 0.28 | 0.16 | 8.04 |
| dha | 44 | -0.10 | 0.09 | 0.29 | 0.10 | 4.78 |
| sfa | 44 | -0.07 | 0.07 | 0.29 | 0.13 | 5.92 |
| alb | 44 | -0.14 | 0.14 | 0.29 | 0.12 | 5.33 |
| s_vldl_fc | 44 | 0.07 | 0.07 | 0.30 | 0.13 | 6.25 |
| mufa_fa | 44 | 0.07 | 0.07 | 0.33 | 0.14 | 6.78 |
| xl_hdl_fc | 44 | -0.06 | 0.07 | 0.34 | 0.09 | 4.03 |
| m_vldl_ce | 44 | 0.07 | 0.07 | 0.36 | 0.14 | 6.80 |
| faw3 | 44 | -0.08 | 0.08 | 0.36 | 0.11 | 4.84 |
| xxl_vldl_ce | 44 | 0.07 | 0.07 | 0.37 | 0.15 | 7.49 |
| tyr | 44 | -0.09 | 0.10 | 0.37 | 0.10 | 4.76 |
| pufa | 44 | -0.06 | 0.08 | 0.40 | 0.10 | 4.76 |
| faw6 | 44 | -0.06 | 0.08 | 0.44 | 0.10 | 4.73 |
| s_vldl_ce | 44 | -0.05 | 0.06 | 0.46 | 0.11 | 4.85 |
| remnant_c | 44 | -0.05 | 0.06 | 0.46 | 0.11 | 5.09 |
| apob | 44 | -0.05 | 0.06 | 0.48 | 0.11 | 5.18 |
| crea | 44 | 0.06 | 0.09 | 0.49 | 0.11 | 4.83 |
| totfa | 44 | -0.04 | 0.07 | 0.51 | 0.12 | 5.78 |
| vldl_c | 44 | 0.04 | 0.07 | 0.53 | 0.13 | 6.02 |
| s_hdl_fc | 44 | -0.05 | 0.08 | 0.53 | 0.12 | 5.52 |
| phe | 44 | -0.06 | 0.10 | 0.55 | 0.12 | 5.36 |
| s_ldl_tg | 44 | 0.04 | 0.06 | 0.55 | 0.13 | 6.16 |
| val | 44 | -0.06 | 0.09 | 0.55 | 0.11 | 5.08 |


| ldl_d | 44 | 0.04 | 0.07 | 0.58 | 0.10 | 4.70 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| apob_apoa1 | 44 | 0.03 | 0.06 | 0.60 | 0.11 | 5.02 |
| leu | 44 | -0.04 | 0.08 | 0.60 | 0.13 | 6.21 |
| bohbut | 44 | -0.07 | 0.13 | 0.61 | 0.10 | 4.78 |
| 1_ldl_tg | 44 | -0.03 | 0.05 | 0.61 | 0.11 | 4.84 |
| 1_hdl_tg | 44 | -0.03 | 0.06 | 0.63 | 0.09 | 4.03 |
| m_ldl_tg | 44 | -0.03 | 0.06 | 0.64 | 0.10 | 4.77 |
| pufa_fa | 44 | -0.03 | 0.06 | 0.64 | 0.14 | 6.92 |
| idl_tg | 44 | 0.02 | 0.05 | 0.66 | 0.11 | 5.14 |
| dha_fa | 44 | -0.03 | 0.07 | 0.72 | 0.12 | 5.59 |
| faw6_fa | 44 | -0.01 | 0.06 | 0.82 | 0.14 | 6.61 |
| ile | 44 | 0.02 | 0.08 | 0.82 | 0.14 | 6.94 |
| ala | 44 | -0.02 | 0.10 | 0.82 | 0.10 | 4.30 |
| ldl_tg | 44 | -0.01 | 0.05 | 0.84 | 0.11 | 4.99 |
| xl_hdl_tg | 44 | 0.01 | 0.08 | 0.85 | 0.12 | 5.38 |
| la | 44 | -0.01 | 0.07 | 0.86 | 0.11 | 4.84 |
| glol | 43 | -0.01 | 0.08 | 0.86 | 0.12 | 5.23 |
| mufa | 44 | 0.01 | 0.06 | 0.87 | 0.14 | 6.76 |
| faw3_fa | 44 | 0.01 | 0.07 | 0.90 | 0.12 | 5.34 |
| s_hdl_pl | 44 | -0.01 | 0.07 | 0.91 | 0.13 | 6.14 |
| s_vldl_c | 44 | 0.00 | 0.06 | 0.97 | 0.11 | 5.18 |
| gp | 44 | 0.00 | 0.06 | 0.97 | 0.13 | 6.12 |

Appendix C Table C 14: Linear regression results of the instrumented exposure (percentage of days with 12,500 steps or more) on metabolic traits in the brisk walking arm compared to control (IV).

| Metabolite | N | Beta | Standard <br> error | p- <br> value | R2 | F-statistic |  |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| la_fa | 73 | 6.27 | 3.27 | 0.06 | 0.08 | 5.76 |  |
| faw6_fa | 73 | 5.81 | 3.04 | 0.06 | 0.08 | 5.76 |  |
| pufa_fa | 73 | 5.07 | 2.90 | 0.08 | 0.08 | 5.76 |  |
| xxl_vldl_ce | 73 | -5.58 | 3.20 | 0.08 | 0.08 | 5.76 |  |
| ile | 73 | -5.18 | 2.97 | 0.08 | 0.08 | 5.76 |  |
| xxl_vldl_c | 73 | -5.33 | 3.11 | 0.09 | 0.08 | 5.76 |  |
| xl_hdl_tg | 73 | -5.51 | 3.26 | 0.09 | 0.08 | 5.76 |  |
| xl_vldl_ce | 73 | -5.18 | 3.07 | 0.09 | 0.08 | 5.76 |  |
| leu | 73 | -4.93 | 2.94 | 0.09 | 0.08 | 5.76 |  |
| xl_vldl_c | 73 | -5.11 | 3.05 | 0.09 | 0.08 | 5.76 |  |
| xl_vldl_fc | 73 | -5.02 | 3.02 | 0.10 | 0.08 | 5.76 |  |
| sfa | 73 | -5.26 | 3.18 | 0.10 | 0.08 | 5.76 |  |
| xxl_vldl_fc | 73 | -4.92 | 3.00 | 0.10 | 0.08 | 5.76 |  |
| xxl_vldl_1 | 73 | -4.92 | 3.01 | 0.10 | 0.08 | 5.76 |  |
| xl_vldl_pl | 73 | -4.84 | 2.97 | 0.10 | 0.08 | 5.76 |  |
| xxl_vldl_pl | 73 | -4.88 | 3.00 | 0.10 | 0.08 | 5.76 |  |
| l_vldl_ce | 73 | -4.84 | 2.98 | 0.10 | 0.08 | 5.76 |  |
| xxl_vldl_p | 73 | -4.87 | 3.00 | 0.10 | 0.08 | 5.76 |  |
| l_vldl_c | 73 | -4.75 | 2.94 | 0.11 | 0.08 | 5.76 |  |
| xxl_vldl_tg | 73 | -4.81 | 2.98 | 0.11 | 0.08 | 5.76 |  |
|  | 73 |  |  |  |  |  |  |


| Metabolite | N | Beta | Standard error | $\begin{gathered} \text { p- } \\ \text { value } \end{gathered}$ | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_vldl_1 | 73 | -4.74 | 2.94 | 0.11 | 0.08 | 5.76 |
| xl_vldl_p | 73 | -4.70 | 2.93 | 0.11 | 0.08 | 5.76 |
| xs_vldl_tg | 73 | -4.67 | 2.92 | 0.11 | 0.08 | 5.76 |
| 1_vldl_fc | 73 | -4.65 | 2.91 | 0.11 | 0.08 | 5.76 |
| m_vldl_c | 73 | -4.76 | 2.98 | 0.11 | 0.08 | 5.76 |
| m_vldl_ce | 73 | -4.91 | 3.07 | 0.11 | 0.08 | 5.76 |
| s_hdl_tg | 73 | -4.44 | 2.80 | 0.11 | 0.08 | 5.76 |
| vldl_c | 73 | -4.81 | 3.04 | 0.11 | 0.08 | 5.76 |
| xl_vldl_tg | 73 | -4.58 | 2.90 | 0.11 | 0.08 | 5.76 |
| serum_tg | 73 | -4.54 | 2.88 | 0.12 | 0.08 | 5.76 |
| idl_tg | 73 | -4.77 | 3.03 | 0.12 | 0.08 | 5.76 |
| m_vldl_fc | 73 | -4.52 | 2.88 | 0.12 | 0.08 | 5.76 |
| 1_vldl_pl | 73 | -4.49 | 2.87 | 0.12 | 0.08 | 5.76 |
| sfa_fa | 73 | -4.54 | 2.91 | 0.12 | 0.08 | 5.76 |
| 1_vldl_1 | 73 | -4.45 | 2.86 | 0.12 | 0.08 | 5.76 |
| m_vldl_pl | 73 | -4.46 | 2.87 | 0.12 | 0.08 | 5.76 |
| m_vldl_1 | 73 | -4.44 | 2.86 | 0.12 | 0.08 | 5.76 |
| 1_vldl_p | 73 | -4.41 | 2.85 | 0.12 | 0.08 | 5.76 |
| hdl_tg | 73 | -4.67 | 3.02 | 0.12 | 0.08 | 5.76 |
| m_vldl_p | 73 | -4.40 | 2.85 | 0.12 | 0.08 | 5.76 |
| vldl_tg | 73 | -4.34 | 2.83 | 0.13 | 0.08 | 5.76 |
| l_vldl_tg | 73 | -4.30 | 2.82 | 0.13 | 0.08 | 5.76 |


| Metabolite | N | Beta | Standard error | $\begin{gathered} \mathrm{p}- \\ \text { value } \end{gathered}$ | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_vldl_tg | 73 | -4.25 | 2.81 | 0.13 | 0.08 | 5.76 |
| s_ldl_tg | 73 | -4.50 | 2.99 | 0.13 | 0.08 | 5.76 |
| mufa | 73 | -4.46 | 2.98 | 0.14 | 0.08 | 5.76 |
| apob_apoa1 | 73 | -4.23 | 2.83 | 0.14 | 0.08 | 5.76 |
| s_vldl_tg | 73 | -4.13 | 2.78 | 0.14 | 0.08 | 5.76 |
| s_vldl_fc | 73 | -4.24 | 2.86 | 0.14 | 0.08 | 5.76 |
| s_vldl_p | 73 | -4.20 | 2.84 | 0.14 | 0.08 | 5.76 |
| s_vldl_1 | 73 | -4.20 | 2.85 | 0.14 | 0.08 | 5.76 |
| totfa | 73 | -4.36 | 3.03 | 0.15 | 0.08 | 5.76 |
| 1_ldl_tg | 73 | -4.34 | 3.01 | 0.15 | 0.08 | 5.76 |
| ldl_tg | 73 | -4.26 | 2.99 | 0.15 | 0.08 | 5.76 |
| remnant_c | 73 | -4.30 | 3.02 | 0.16 | 0.08 | 5.76 |
| s_vldl_pl | 73 | -3.96 | 2.81 | 0.16 | 0.08 | 5.76 |
| s_hdl_ce | 73 | 3.89 | 2.81 | 0.17 | 0.08 | 5.76 |
| faw3 | 73 | -4.23 | 3.05 | 0.17 | 0.08 | 5.76 |
| apob | 73 | -4.08 | 2.96 | 0.17 | 0.08 | 5.76 |
| s_hdl_c | 73 | 3.83 | 2.80 | 0.17 | 0.08 | 5.76 |
| unsat | 73 | 3.57 | 2.61 | 0.17 | 0.08 | 5.76 |
| s_vldl_c | 73 | -3.95 | 2.91 | 0.17 | 0.08 | 5.76 |
| tg_pg | 73 | -3.64 | 2.70 | 0.18 | 0.08 | 5.76 |
| xs_vldl_p | 73 | -3.95 | 2.96 | 0.18 | 0.08 | 5.76 |
| m_hdl_c | 73 | 3.32 | 2.54 | 0.19 | 0.08 | 5.76 |


| Metabolite | N | Beta | Standard error | $\begin{gathered} \mathrm{p}- \\ \text { value } \end{gathered}$ | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_hdl_ce | 73 | 3.31 | 2.55 | 0.19 | 0.08 | 5.76 |
| m_hdl_fc | 73 | 3.28 | 2.53 | 0.20 | 0.08 | 5.76 |
| gp | 73 | -3.54 | 2.75 | 0.20 | 0.08 | 5.76 |
| xs_vldl_1 | 73 | $-3.78$ | 2.94 | 0.20 | 0.08 | 5.76 |
| m_ldl_tg | 73 | -3.69 | 2.91 | 0.20 | 0.08 | 5.76 |
| xs_vldl_ce | 73 | -3.64 | 2.94 | 0.22 | 0.08 | 5.76 |
| s_vldl_ce | 73 | -3.52 | 2.89 | 0.22 | 0.08 | 5.76 |
| 1_hdl_pl | 73 | 2.97 | 2.49 | 0.23 | 0.08 | 5.76 |
| bohbut | 73 | 3.77 | 3.17 | 0.24 | 0.08 | 5.76 |
| hdl2_c | 73 | 2.92 | 2.48 | 0.24 | 0.08 | 5.76 |
| xs_vldl_c | 73 | -3.43 | 2.91 | 0.24 | 0.08 | 5.76 |
| totcho | 73 | -3.58 | 3.06 | 0.24 | 0.08 | 5.76 |
| ace | 73 | 3.15 | 2.71 | 0.25 | 0.08 | 5.76 |
| totpg | 73 | -3.42 | 2.97 | 0.25 | 0.08 | 5.76 |
| m_hdl_tg | 73 | -3.08 | 2.75 | 0.26 | 0.08 | 5.76 |
| hdl_c | 73 | 2.77 | 2.48 | 0.26 | 0.08 | 5.76 |
| dha | 73 | -3.09 | 2.79 | 0.27 | 0.08 | 5.76 |
| pc | 73 | -3.27 | 2.97 | 0.27 | 0.08 | 5.76 |
| m_hdl_1 | 73 | 2.68 | 2.48 | 0.28 | 0.08 | 5.76 |
| val | 73 | -2.75 | 2.60 | 0.29 | 0.08 | 5.76 |
| 1_hdl_fc | 73 | 2.55 | 2.49 | 0.30 | 0.08 | 5.76 |
| xs_vldl_fc | 73 | -2.89 | 2.84 | 0.31 | 0.08 | 5.76 |


| Metabolite | N | Beta | Standard error | $\begin{gathered} \mathrm{p}- \\ \text { value } \end{gathered}$ | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_hdl_p | 73 | 2.51 | 2.47 | 0.31 | 0.08 | 5.76 |
| 1_hdl_1 | 73 | 2.48 | 2.47 | 0.32 | 0.08 | 5.76 |
| vldl_d | 73 | $-2.55$ | 2.56 | 0.32 | 0.08 | 5.76 |
| 1_hdl_p | 73 | 2.46 | 2.47 | 0.32 | 0.08 | 5.76 |
| idl_ce | 73 | -2.79 | 2.83 | 0.32 | 0.08 | 5.76 |
| s_ldl_pl | 73 | -2.76 | 2.80 | 0.33 | 0.08 | 5.76 |
| idl_p | 73 | -2.75 | 2.82 | 0.33 | 0.08 | 5.76 |
| m_hdl_pl | 73 | 2.40 | 2.47 | 0.33 | 0.08 | 5.76 |
| mufa_fa | 73 | -2.48 | 2.60 | 0.34 | 0.08 | 5.76 |
| m_ldl_pl | 73 | -2.65 | 2.77 | 0.34 | 0.08 | 5.76 |
| freec | 73 | -2.71 | 2.85 | 0.34 | 0.08 | 5.76 |
| 1_hdl_c | 73 | 2.31 | 2.47 | 0.35 | 0.08 | 5.76 |
| 1_hdl_tg | 73 | -2.61 | 2.84 | 0.36 | 0.08 | 5.76 |
| idl_1 | 73 | -2.55 | 2.80 | 0.36 | 0.08 | 5.76 |
| phe | 73 | -2.34 | 2.56 | 0.36 | 0.08 | 5.76 |
| 1_hdl_ce | 73 | 2.23 | 2.47 | 0.37 | 0.08 | 5.76 |
| xs_vldl_pl | 73 | -2.41 | 2.78 | 0.39 | 0.08 | 5.76 |
| idl_c | 73 | -2.39 | 2.78 | 0.39 | 0.08 | 5.76 |
| s_ldl_fc | 73 | -2.29 | 2.74 | 0.40 | 0.08 | 5.76 |
| pufa | 73 | -2.30 | 2.76 | 0.40 | 0.08 | 5.76 |
| sm | 73 | -2.38 | 2.86 | 0.41 | 0.08 | 5.76 |
| serum_c | 73 | -2.16 | 2.76 | 0.44 | 0.08 | 5.76 |


| Metabolite | N | Beta | Standard error | $\begin{gathered} \mathrm{p}- \\ \text { value } \end{gathered}$ | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_hdl_pl | 73 | 1.93 | 2.48 | 0.44 | 0.08 | 5.76 |
| 1_ldl_p | 73 | -2.11 | 2.72 | 0.44 | 0.08 | 5.76 |
| s_ldl_p | 73 | -2.07 | 2.71 | 0.45 | 0.08 | 5.76 |
| s_hdl_1 | 73 | 1.98 | 2.60 | 0.45 | 0.08 | 5.76 |
| m_ldl_fc | 73 | -2.01 | 2.70 | 0.46 | 0.08 | 5.76 |
| glol | 73 | 2.07 | 2.79 | 0.46 | 0.08 | 5.76 |
| s_ldl_1 | 73 | -1.96 | 2.69 | 0.47 | 0.08 | 5.76 |
| 1_1dl_1 | 73 | -1.96 | 2.70 | 0.47 | 0.08 | 5.76 |
| cit | 73 | -1.91 | 2.66 | 0.47 | 0.08 | 5.76 |
| m_ldl_p | 73 | -1.92 | 2.69 | 0.48 | 0.08 | 5.76 |
| xl_hdl_ce | 73 | -1.96 | 2.75 | 0.48 | 0.08 | 5.76 |
| 1_ldl_ce | 73 | -1.92 | 2.70 | 0.48 | 0.08 | 5.76 |
| faw6 | 73 | -1.92 | 2.70 | 0.48 | 0.08 | 5.76 |
| pyr | 73 | -1.73 | 2.44 | 0.48 | 0.08 | 5.76 |
| m_ldl_1 | 73 | -1.88 | 2.68 | 0.48 | 0.08 | 5.76 |
| estc | 73 | -1.91 | 2.72 | 0.48 | 0.08 | 5.76 |
| crea | 73 | 1.80 | 2.65 | 0.50 | 0.08 | 5.76 |
| 1_ldl_pl | 73 | -1.77 | 2.69 | 0.51 | 0.08 | 5.76 |
| s_hdl_p | 73 | 1.69 | 2.58 | 0.51 | 0.08 | 5.76 |
| 1_ldl_c | 73 | -1.71 | 2.67 | 0.52 | 0.08 | 5.76 |
| hdl_d | 73 | 1.56 | 2.45 | 0.52 | 0.08 | 5.76 |
| idl_pl | 73 | -1.68 | 2.69 | 0.53 | 0.08 | 5.76 |


| Metabolite | N | Beta | Standard error | pvalue | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| alb | 73 | -1.58 | 2.57 | 0.54 | 0.08 | 5.76 |
| xl_hdl_c | 73 | -1.66 | 2.71 | 0.54 | 0.08 | 5.76 |
| ldl_c | 73 | -1.57 | 2.65 | 0.56 | 0.08 | 5.76 |
| m_ldl_c | 73 | -1.46 | 2.64 | 0.58 | 0.08 | 5.76 |
| s_ldl_c | 73 | -1.34 | 2.63 | 0.61 | 0.08 | 5.76 |
| m_ldl_ce | 73 | -1.34 | 2.62 | 0.61 | 0.08 | 5.76 |
| gly | 73 | 1.28 | 2.64 | 0.63 | 0.08 | 5.76 |
| idl_fc | 73 | -1.24 | 2.64 | 0.64 | 0.08 | 5.76 |
| la | 73 | -1.16 | 2.60 | 0.66 | 0.08 | 5.76 |
| s_ldl_ce | 73 | -1.11 | 2.60 | 0.67 | 0.08 | 5.76 |
| ldl_d | 73 | -1.10 | 2.61 | 0.67 | 0.08 | 5.76 |
| 1_ldl_fc | 73 | -1.02 | 2.61 | 0.70 | 0.08 | 5.76 |
| glc | 73 | -0.94 | 2.47 | 0.70 | 0.08 | 5.76 |
| gln | 73 | 0.96 | 2.56 | 0.71 | 0.08 | 5.76 |
| tyr | 73 | 0.92 | 2.60 | 0.72 | 0.08 | 5.76 |
| xl_hdl_f | 73 | -0.84 | 2.61 | 0.75 | 0.08 | 5.76 |
| his | 73 | -0.77 | 2.51 | 0.76 | 0.08 | 5.76 |
| ala | 73 | 0.74 | 2.58 | 0.77 | 0.08 | 5.76 |
| dha_fa | 73 | 0.61 | 2.46 | 0.80 | 0.08 | 5.76 |
| hdl3_c | 73 | 0.55 | 2.52 | 0.83 | 0.08 | 5.76 |
| faw3_fa | 73 | -0.47 | 2.45 | 0.85 | 0.08 | 5.76 |
| s_hdl_pl | 73 | 0.29 | 2.51 | 0.91 | 0.08 | 5.76 |


| Metabolite | $\mathbf{N}$ | Beta | Standard <br> error | p- <br> value | R2 | F-statistic |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| lac | 73 | 0.22 | 2.53 | 0.93 | 0.08 | 5.76 |
| xl_hdl_p | 73 | 0.13 | 2.53 | 0.96 | 0.08 | 5.76 |
| xl_hdl_l | 73 | 0.07 | 2.53 | 0.98 | 0.08 | 5.76 |
| s_hdl_fc | 73 | -0.06 | 2.52 | 0.98 | 0.08 | 5.76 |
| apoa1 | 73 | 0.04 | 2.52 | 0.99 | 0.08 | 5.76 |

Appendix C Table C 15: Linear regression results of the instrumented exposure (percentage of days with 12,500 steps or more) on metabolic traits in the brisk walking arm compared to control (IV), adjusted for baseline metabolic traits and age

| Metabolite | N | BetaStandard <br> error | P-value | R2 | F-statistic |  |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| sfa | 73 | -4.25 | 1.77 | 0.02 | 0.12 | 9.56 |
| s_ldl_tg | 73 | -3.98 | 1.67 | 0.02 | 0.12 | 9.57 |
| ldl_tg | 73 | -4.15 | 1.75 | 0.02 | 0.12 | 9.51 |
| xl_hdl_tg | 73 | -4.31 | 1.84 | 0.02 | 0.13 | 10.15 |
| l_ldl_tg | 73 | -4.15 | 1.77 | 0.02 | 0.12 | 9.52 |
| m_ldl_tg | 73 | -4.09 | 1.76 | 0.02 | 0.12 | 9.38 |
| idl_tg | 73 | -3.87 | 1.67 | 0.02 | 0.12 | 9.57 |
| hdl_tg | 73 | -4.17 | 1.81 | 0.02 | 0.12 | 9.52 |
| totfa | 73 | -3.62 | 1.68 | 0.03 | 0.12 | 9.58 |
| l_hdl_tg | 73 | -5.66 | 2.66 | 0.03 | 0.09 | 6.81 |
| xl_hdl_fc | 73 | -3.75 | 1.77 | 0.03 | 0.10 | 7.53 |
| xl_hdl_c | 73 | -3.52 | 1.67 | 0.04 | 0.11 | 8.36 |
| totpg | 73 | -4.15 | 1.98 | 0.04 | 0.11 | 8.90 |
| xs_vldl_tg | 73 | -3.07 | 1.47 | 0.04 | 0.12 | 9.35 |
| xl_hdl_ce | 73 | -3.40 | 1.63 | 0.04 | 0.11 | 8.70 |
| pc | 73 | -4.02 | 1.94 | 0.04 | 0.12 | 9.10 |
| totcho | 73 | -3.89 | 1.95 | 0.05 | 0.12 | 9.18 |
| xxl_vldl_pl | 73 | -3.15 | 1.60 | 0.05 | 0.12 | 9.57 |
| xl_hdl_1 | 73 | -3.02 | 1.54 | 0.05 | 0.10 | 7.57 |
| xl_hdl_p | 73 | -3.00 | 1.54 | 0.05 | 0.10 | 7.54 |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |


| Metabolite | N | Beta | Standard error | $P$-value | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xxl_vldl_fc | 73 | -3.09 | 1.59 | 0.05 | 0.12 | 9.54 |
| xxl_vldl_c | 73 | -3.14 | 1.62 | 0.05 | 0.12 | 9.51 |
| xxl_vldl_ce | 73 | -3.16 | 1.64 | 0.06 | 0.12 | 9.51 |
| mufa | 73 | $-2.98$ | 1.56 | 0.06 | 0.12 | 9.59 |
| xxl_vldl_l | 73 | -3.06 | 1.60 | 0.06 | 0.12 | 9.54 |
| xxl_vldl_p | 73 | $-3.05$ | 1.60 | 0.06 | 0.12 | 9.54 |
| s_hdl_tg | 73 | $-2.79$ | 1.47 | 0.06 | 0.12 | 9.28 |
| freec | 73 | $-3.32$ | 1.75 | 0.06 | 0.12 | 9.35 |
| xl_vldl_fc | 73 | -3.01 | 1.59 | 0.06 | 0.12 | 9.50 |
| remnant_c | 73 | $-2.90$ | 1.53 | 0.06 | 0.12 | 9.54 |
| xxl_vldl_tg | 73 | -3.02 | 1.60 | 0.06 | 0.12 | 9.54 |
| vldl_c | 73 | $-2.81$ | 1.49 | 0.06 | 0.12 | 9.37 |
| xl_vldl_pl | 73 | $-2.91$ | 1.55 | 0.06 | 0.12 | 9.48 |
| m_hdl_tg | 73 | $-2.85$ | 1.52 | 0.06 | 0.12 | 9.54 |
| serum_tg | 73 | $-2.76$ | 1.48 | 0.06 | 0.12 | 9.35 |
| s_vldl_fc | 73 | $-2.63$ | 1.42 | 0.06 | 0.12 | 9.32 |
| ile | 73 | -3.05 | 1.64 | 0.06 | 0.12 | 9.73 |
| xl_vldl_c | 73 | $-2.93$ | 1.59 | 0.07 | 0.12 | 9.45 |
| apob | 73 | $-2.78$ | 1.52 | 0.07 | 0.12 | 9.52 |
| pufa_fa | 73 | 2.48 | 1.37 | 0.07 | 0.12 | 9.10 |
| l_vldl_fc | 73 | $-2.75$ | 1.52 | 0.07 | 0.12 | 9.41 |
| xs_vldl_p | 73 | $-2.75$ | 1.52 | 0.07 | 0.12 | 9.52 |


| Metabolite | N | Beta | Standard error | P-value | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_vldl_ce | 73 | -2.87 | 1.59 | 0.07 | 0.12 | 9.40 |
| xl_vldl_1 | 73 | -2.78 | 1.54 | 0.07 | 0.12 | 9.43 |
| xl_vldl_p | 73 | -2.76 | 1.54 | 0.07 | 0.12 | 9.43 |
| dha_fa | 73 | 2.87 | 1.60 | 0.07 | 0.11 | 8.68 |
| m_vldl_c | 73 | -2.69 | 1.51 | 0.07 | 0.12 | 9.33 |
| m_vldl_ce | 73 | -2.73 | 1.53 | 0.08 | 0.12 | 9.39 |
| m_vldl_fc | 73 | -2.61 | 1.48 | 0.08 | 0.12 | 9.31 |
| xl_vldl_tg | 73 | -2.69 | 1.52 | 20.08 | 0.12 | 9.42 |
| l_vldl_c | 73 | -2.68 | 1.52 | - 0.08 | 0.12 | 9.35 |
| xs_vldl_1 | 73 | -2.65 | 1.52 | 0.08 | 0.12 | 9.54 |
| 1_vldl_pl | 73 | -2.55 | 1.49 | 0.09 | 0.12 | 9.32 |
| s_vldl_pl | 73 | -2.36 | 1.38 | 0.09 | 0.12 | 9.31 |
| m_vldl_pl | 73 | -2.49 | 1.46 | 0.09 | 0.12 | 9.26 |
| l_vldl_ce | 73 | -2.60 | 1.53 | 0.09 | 0.12 | 9.29 |
| s_vldl_1 | 73 | -2.36 | 1.39 | 0.09 | 0.12 | 9.22 |
| hdl_d | 73 | -2.75 | 1.63 | 0.09 | 0.09 | 6.50 |
| l_hdl_ce | 73 | -2.34 | 1.39 | 0.09 | 0.09 | 6.42 |
| leu | 73 | -2.89 | 1.73 | 0.09 | 0.12 | 9.73 |
| s_vldl_p | 73 | -2.33 | 1.39 | 0.10 | 0.12 | 9.19 |
| 1_vldl_1 | 73 | -2.48 | 1.49 | 0.10 | 0.12 | 9.29 |
| pufa | 73 | -2.84 | 1.71 | 0.10 | 0.12 | 9.38 |
| vldl_tg | 73 | -2.41 | 1.45 | 0.10 | 0.12 | 9.24 |


| Metabolite | N | Beta | Standard error | P-value | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_vldl_1 | 73 | -2.42 | 1.46 | - 0.10 | 0.12 | 9.22 |
| 1_vldl_p | 73 | -2.45 | 1.48 | - 0.10 | 0.12 | 9.28 |
| idl_p | 73 | -2.56 | 1.56 | 0.10 | 0.12 | 9.57 |
| s_vldl_c | 73 | -2.31 | 1.41 | 0.10 | 0.12 | 9.38 |
| 1_hdl_c | 73 | -2.22 | 1.36 | 0.10 | 0.09 | 6.48 |
| m_vldl_p | 73 | -2.38 | 1.46 | 0.10 | 0.12 | 9.21 |
| faw6_fa | 73 | 2.38 | 1.46 | 0.10 | 0.11 | 8.57 |
| s_ldl_pl | 73 | -2.56 | 1.59 | 0.11 | 0.12 | 9.57 |
| serum_c | 73 | -2.60 | 1.61 | 0.11 | 0.12 | 9.42 |
| 1_vldl_tg | 73 | -2.37 | 1.47 | 0.11 | 0.12 | 9.27 |
| faw6 | 73 | -2.77 | 1.73 | 0.11 | 0.12 | 9.25 |
| faw3 | 73 | -2.40 | 1.51 | 0.11 | 0.13 | 10.09 |
| 1_hdl_p | 73 | -2.30 | 1.46 | 0.11 | 0.08 | 6.27 |
| s_vldl_tg | 73 | -2.22 | 1.41 | 0.11 | 0.12 | 9.13 |
| idl_1 | 73 | -2.43 | 1.55 | 0.12 | 0.12 | 9.57 |
| m_ldl_pl | 73 | -2.37 | 1.52 | - 0.12 | 0.12 | 9.58 |
| 1_hdl_1 | 73 | -2.24 | 1.43 | 0.12 | 0.08 | 6.31 |
| m_vldl_tg | 73 | -2.25 | 1.44 | 0.12 | 0.12 | 9.17 |
| idl_ce | 73 | -2.43 | 1.56 | 0.12 | 0.12 | 9.58 |
| vldl_d | 73 | -2.12 | 1.38 | 0.13 | 0.12 | 9.56 |
| unsat | 73 | 2.26 | 1.48 | 0.13 | 0.12 | 9.63 |
| sm | 73 | -2.66 | 1.74 | 0.13 | 0.12 | 9.40 |


| Metabolite | N | Beta | Standard error | P-value | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_hdl_pl | 73 | -1.99 | 1.32 | 2.13 | 0.09 | 7.09 |
| xs_vldl_fc | 73 | -2.25 | 1.51 | 0.14 | 0.12 | 9.60 |
| xs_vldl_c | 73 | -2.32 | 1.56 | - 0.14 | 0.12 | 9.54 |
| ace | 73 | 3.08 | 2.10 | 0.14 | 0.13 | 10.25 |
| xs_vldl_ce | 73 | -2.33 | 1.59 | 0.14 | 0.12 | 9.49 |
| idl_c | 73 | -2.26 | 1.55 | 0.15 | 0.12 | 9.57 |
| estc | 73 | $-2.28$ | 1.57 | 0.15 | 0.12 | 9.45 |
| 1_hdl_fc | 73 | -1.85 | 1.27 | 0.15 | 0.09 | 6.65 |
| sfa_fa | 73 | -2.72 | 1.88 | 0.15 | 0.12 | 9.70 |
| la | 73 | -2.51 | 1.73 | 0.15 | 0.12 | 9.11 |
| apoa1 | 73 | -2.94 | 2.04 | 0.15 | 0.10 | 7.32 |
| 1_ldl_p | 73 | -2.14 | 1.50 | 0.15 | 0.12 | 9.57 |
| s_vldl_ce | 73 | -1.97 | 1.42 | 20.16 | 0.12 | 9.45 |
| xs_vldl_pl | 73 | -2.01 | 1.46 | 0.17 | 0.12 | 9.60 |
| 1_1dl_1 | 73 | -2.02 | 1.49 | 0.17 | 0.12 | 9.57 |
| s_ldl_p | 73 | -2.01 | 1.49 | 0.18 | 0.12 | 9.58 |
| s_ldl_fc | 73 | -2.09 | 1.56 | - 0.18 | 0.12 | 9.58 |
| la_fa | 73 | 2.08 | 1.55 | 0.18 | 0.11 | 8.28 |
| s_hdl_ce | 73 | 2.24 | 1.72 | 0.19 | 0.12 | 9.63 |
| 1_ldl_pl | 73 | -1.92 | 1.48 | 0.20 | 0.12 | 9.56 |
| s_ldl_1 | 73 | -1.92 | 1.49 | 0.20 | 0.12 | 9.58 |
| m_ldl_p | 73 | -1.88 | 1.46 | - 0.20 | 0.12 | 9.58 |


| Metabolite | N | Beta | Standard error | P-value | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| l_ldl_ce | 73 | -1.87 | 1.47 | 0.20 | 0.12 | 9.58 |
| idl_pl | 73 | -1.90 | 1.49 | 0.20 | 0.12 | 9.54 |
| m_ldl_fc | 73 | -1.90 | 1.51 | 0.21 | 0.12 | 9.58 |
| bohbut | 73 | 2.67 | 2.12 | 0.21 | 0.13 | 10.07 |
| gp | 73 | -2.12 | 1.69 | 0.21 | 0.12 | 9.31 |
| m_ldl_1 | 73 | -1.84 | 1.46 | 0.21 | 0.12 | 9.58 |
| 1_1dl_c | 73 | -1.78 | 1.47 | 0.22 | 0.12 | 9.57 |
| faw3_fa | 73 | 1.64 | 1.35 | 0.22 | 0.12 | 9.65 |
| l_hdl_pl | 73 | -1.75 | 1.49 | 0.24 | 0.08 | 6.27 |
| idl_fc | 73 | -1.71 | 1.50 | 0.25 | 0.12 | 9.50 |
| s_hdl_c | 73 | 2.01 | 1.77 | 0.26 | 0.12 | 9.54 |
| gly | 73 | 1.98 | 1.75 | 0.26 | 0.12 | 9.04 |
| apob_apoa1 | 73 | -1.48 | 1.31 | 0.26 | 0.11 | 8.43 |
| hdl3_c | 73 | -2.13 | 1.89 | 0.26 | 0.10 | 8.07 |
| ldl_c | 73 | -1.61 | 1.45 | 0.27 | 0.12 | 9.58 |
| tg_pg | 73 | -1.50 | 1.39 | 0.28 | 0.11 | 8.91 |
| m_ldl_c | 73 | -1.46 | 1.44 | 0.31 | 0.12 | 9.58 |
| 1_ldl_fc | 73 | -1.47 | 1.46 | 0.31 | 0.12 | 9.54 |
| s_ldl_c | 73 | -1.40 | 1.44 | 0.33 | 0.12 | 9.58 |
| m_ldl_ce | 73 | -1.36 | 1.43 | 0.34 | 0.12 | 9.58 |
| s_hdl_pl | 73 | -1.58 | 1.70 | 0.35 | 0.11 | 8.47 |
| hdl_c | 73 | -1.46 | 1.61 | 0.36 | 0.09 | 6.64 |


| Metabolite | N | Beta | Standard error | $P$-value | R2 | F-statistic |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_hdl_fc | 73 | -1.67 | 1.85 | 0.37 | 0.11 | 8.48 |
| hdl2_c | 73 | -1.36 | 1.56 | 0.38 | 0.09 | 6.57 |
| cit | 73 | -1.80 | 2.06 | 0.38 | 0.12 | 9.27 |
| s_ldl_ce | 73 | -1.23 | 1.42 | 0.39 | 0.12 | 9.58 |
| alb | 73 | -1.75 | 2.06 | 0.40 | 0.12 | 9.41 |
| val | 73 | -1.28 | 1.69 | 0.45 | 0.12 | 9.65 |
| dha | 73 | -1.09 | 1.48 | 0.46 | 0.12 | 9.60 |
| ldl_d | 73 | -1.28 | 1.90 | 0.50 | 0.12 | 9.15 |
| mufa_fa | 73 | -0.96 | 1.48 | 0.52 | 0.12 | 9.14 |
| m_hdl_pl | 73 | -1.35 | 2.12 | 0.53 | 0.08 | 6.10 |
| his | 73 | -1.10 | 1.90 | 0.56 | 0.13 | 10.36 |
| lac | 73 | 1.10 | 2.13 | 0.61 | 0.11 | 8.79 |
| pyr | 73 | -0.95 | 1.89 | 0.62 | 0.12 | 9.46 |
| glc | 73 | 0.66 | 1.33 | 0.62 | 0.12 | 9.29 |
| m_hdl_p | 73 | -0.88 | 2.12 | 0.68 | 0.08 | 6.19 |
| $g \ln$ | 73 | -0.60 | 1.78 | 0.74 | 0.12 | 9.42 |
| m_hdl_l | 73 | -0.70 | 2.11 | 0.74 | 0.08 | 6.18 |
| ala | 73 | 0.55 | 1.83 | 0.76 | 0.12 | 9.76 |
| phe | 73 | -0.52 | 2.03 | 0.80 | 0.10 | 7.98 |
| tyr | 73 | 0.43 | 1.88 | 0.82 | 0.12 | 9.54 |
| m_hdl_ce | 73 | 0.43 | 2.07 | 0.84 | 0.09 | 6.49 |
| crea | 73 | -0.15 | 1.00 | - 0.88 | 0.13 | 10.53 |


| Metabolite | N | Beta | Standard <br> error | P-value | R2 | F-statistic |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_hdl_c | 73 | 0.30 | 2.07 | 0.88 | 0.09 | 6.41 |
| glol | 73 | -0.19 | 1.56 | 0.90 | 0.16 | 12.94 |
| m_hdl_fc | 73 | -0.24 | 2.06 | 0.91 | 0.08 | 6.21 |
| S_hdl_p | 73 | -0.16 | 1.76 | 0.93 | 0.12 | 9.03 |
| s_hdl_1 | 73 | 0.10 | 1.78 | 0.96 | 0.12 | 8.99 |

Appendix C Table C 16: List of metabolic traits taken forward to MR stage and instrument availability which had a p-value $<0.05$ in the unadjusted ITT

| Metabolite | Intervention arm |
| :--- | :--- |
| Acetate | Dietary advice |
| Alanine | Lycopene |
| Ratio of omega-6 fatty acids to total fatty acids | Dietary advice, Brisk walking |
| Cholesteryl esters in large VLDL | Brisk Walking |
| Ratio of linoleic acid to total fatty acids | Brisk Walking |
| Lactate | Dietary advice |
| Cholesteryl esters in medium VLDL | Brisk Walking |
| Ratio of polyunsaturated fatty acids to total fatty acids | Brisk Walking |
| Pyruvate | Dietary advice |
| Concentration of small HDL particles | Dietary advice |
| Saturated fatty acids | Brisk Walking |
| Ratio of saturated fatty acids to total fatty acids | Dietary advice |
| Triglycerides in very large HDL | Brisk Walking |
| Cholesterol in very large VLDL | Brisk Walking |
| Cholesteryl esters in very large VLDL | Brisk Walking |
| Free cholesterol in very large VLDL | Brisk Walking Walking |
| Phospholipids in very large VLDL | Brisk Walking Walking |
| Cholesterol in chylomicrons and extremely large VLDL | Bripids in chylomicrons and extremely large VLDL |
| Cholesteryl esters in chylomicrons and extremely large VLDL | Bresterol in chylomicrons and extremely large VLDL |
| Total |  |

Appendix C Table C 17: F-statistic and r-square measures for instruments of individual metabolites (metabolites with p-value $<0.05$ in the unadjusted ITT)

| Metabolite | Total r2 | Number of SNPs | Sample <br> Size | F statistic |
| :--- | :---: | :---: | :---: | :---: |
| Acetate | 0.004 | 6 | 115,046 | 68.2 |
| Alanine | 0.012 | 22 | 115,074 | 65.0 |
| Ratio of omega-6 fatty acids to total fatty acids | 0.028 | 36 | 114,999 | 91.8 |
| Cholesteryl esters in large VLDL | 0.039 | 36 | 115,078 | 129.1 |
| Ratio of linoleic acid to total fatty acids | 0.013 | 24 | 114,999 | 63.8 |
| Lactate | 0.003 | 6 | 114,802 | 51.9 |
| Cholesteryl esters in medium VLDL | 0.039 | 38 | 115,078 | 124.2 |
| Ratio of polyunsaturated fatty acids to total fatty acids | 0.031 | 35 | 114,999 | 103.9 |
| Pyruvate | 0.014 | 15 | 114,748 | 109.0 |
| Concentration of small HDL particles | 0.036 | 31 | 115,078 | 138.3 |
| Saturated fatty acids | 0.038 | 40 | 114,999 | 114.7 |
| Ratio of saturated fatty acids to total fatty acids | 0.013 | 20 | 114,999 | 77.2 |


| Triglycerides in very large HDL | 0.079 | 46 | 115,078 | 214.8 |
| :--- | :--- | :--- | :--- | :--- |
| Cholesterol in very large VLDL | 0.040 | 37 | 115,078 | 128.8 |
| Cholesteryl esters in very large VLDL | 0.042 | 42 | 115,078 | 119.6 |
| Free cholesterol in very large VLDL | 0.039 | 34 | 115,078 | 136.7 |
| Phospholipids in very large VLDL | 0.042 | 39 | 115,078 | 128.9 |
| Cholesterol in chylomicrons and extremely large VLDL | 0.047 | 44 | 115,078 | 128.5 |
| Cholesteryl esters in chylomicrons and extremely large VLDL | 0.049 | 44 | 115,078 | 136.0 |
| Free cholesterol in chylomicrons and extremely large VLDL | 0.045 | 45 | 115,078 | 119.4 |
| Total lipids in chylomicrons and extremely large VLDL | 0.046 | 45 | 115,078 | 122.7 |
| Concentration of chylomicrons and extremely large VLDL particles | 0.043 | 42 | 115,078 | 123.2 |
| Phospholipids in chylomicrons and extremely large VLDL | 0.045 | 45 | 115,078 | 121.4 |

Appendix C Table C 18: Mendelian Randomisation results for Prostate Cancer survival in the PRACTICAL consortium, using altered metabolites in the PrEvENT trial ( $\mathrm{p}<0.05$ )

| Exposure (metabolic trait) | Method | Number of snps | p-val | $\begin{aligned} & \text { Odds } \\ & \text { ratio } \end{aligned}$ | Odds ratio lower $95 \% \text { CI }$ | Odds ratio upper 95\% CI | Intervention |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Acetate | MR Egger | 6 | 0.790 | 1.199 | 0.345 | 4.169 | Dietary advice |
| Acetate | Simple mode | 6 | 0.711 | 0.841 | 0.355 | 1.993 | Dietary advice |
| Acetate | Weighted median | 6 | 0.927 | 1.026 | 0.596 | 1.765 | Dietary advice |
| Acetate | Weighted mode | 6 | 0.929 | 1.030 | 0.555 | 1.912 | Dietary advice |
| LA_pct | Inverse variance weighted | 24 | 0.133 | 0.846 | 0.680 | 1.053 | Brisk Walking |


| Exposure <br> (metabolic trait) | Method |
| :--- | :--- |
| Ala | MR Egger |
| Ala | Simple mode |
| Ala | Weighted mode |
| Ala | Weighted median |
| Omega_6_pct | Inverse variance weighted |
| Pyruvate | Inverse variance weighted |
| L_VLDL_CE | MR Egger |
| L_VLDL_CE | Weighted mode |
| L_VLDL_CE | Simple mode |
| L_VLDL_CE | Weighted median |
| PUFA_pct | Inverse variance weighted |
| LA_pct | Simple mode |
| LA_pct | Weighted mode |
| LA_pct | Weighted median |


| Number of snps | p-val | $\begin{aligned} & \text { Odds } \\ & \text { ratio } \end{aligned}$ | Odds ratio lower 95\% CI | Odds ratio upper 95\% CI | Intervention |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 22 | 0.053 | 2.036 | 1.034 | 4.010 | Dietary advice |
| 22 | 0.662 | 1.121 | 0.677 | 1.854 | Dietary advice |
| 22 | 0.513 | 1.138 | 0.777 | 1.666 | Dietary advice |
| 22 | 0.339 | 1.165 | 0.852 | 1.594 | Dietary advice |
| 36 | 0.183 | 0.898 | 0.768 | 1.052 | Dietary advice, brisk walking |
| 15 | 0.695 | 0.948 | 0.728 | 1.236 | Dietary advice |
| 36 | 0.768 | 0.959 | 0.727 | 1.264 | Brisk Walking |
| 36 | 0.346 | 1.091 | 0.912 | 1.306 | Brisk Walking |
| 36 | 0.429 | 1.112 | 0.857 | 1.444 | Brisk Walking |
| 36 | 0.175 | 1.129 | 0.947 | 1.346 | Brisk Walking |
| 35 | 0.649 | 0.967 | 0.839 | 1.115 | Brisk Walking |
| 24 | 0.442 | 0.799 | 0.455 | 1.402 | Brisk Walking |
| 24 | 0.492 | 0.875 | 0.602 | 1.272 | Brisk Walking |
| 24 | 0.503 | 0.893 | 0.641 | 1.243 | Brisk Walking |


| Exposure <br> (metabolic trait) | Method |
| :--- | :--- |
| LA_pct | MR Egger |
| S_HDL_P | Inverse variance weighted |
| Lactate | MR Egger |
| Lactate | Simple mode |
| Lactate | Weighted median |
| Lactate | Weighted mode |
| M_VLDL_CE | MR Egger |
| SFA_pct | Inverse variance weighted |
| M_VLDL_CE | Weighted median |
| M_VLDL_CE | Weighted mode |
| M_VLDL_CE | Simple mode |
| Omega_6_pct | Simple mode |
| Acetate | Inverse variance weighted |
| Omega_6_pct | Weighted mode |


| Number of snps | p-val | $\begin{aligned} & \text { Odds } \\ & \text { ratio } \end{aligned}$ | Odds ratio lower 95\% CI | Odds ratio upper 95\% CI | Intervention |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 24 | 0.996 | 0.999 | 0.610 | 1.635 | Brisk Walking |
| 31 | 0.869 | 0.989 | 0.866 | 1.129 | Dietary advice |
| 6 | 0.794 | 1.375 | 0.147 | 12.850 | Dietary advice |
| 6 | 0.341 | 1.504 | 0.704 | 3.214 | Dietary advice |
| 6 | 0.132 | 1.520 | 0.881 | 2.624 | Dietary advice |
| 6 | 0.237 | 1.629 | 0.799 | 3.320 | Dietary advice |
| 38 | 0.035 | 1.305 | 1.028 | 1.657 | Dietary advice |
| 20 | 0.796 | 1.038 | 0.780 | 1.382 | Brisk Walking |
| 38 | 0.059 | 1.208 | 0.993 | 1.471 | Brisk Walking |
| 38 | 0.023 | 1.344 | 1.052 | 1.717 | Brisk Walking |
| 38 | 0.128 | 1.379 | 0.920 | 2.065 | Brisk Walking |
| 36 | 0.386 | 0.864 | 0.624 | 1.197 | Dietary advice, brisk wlking |
| 6 | 0.807 | 1.056 | 0.683 | 1.634 | Dietary advice |
| 36 | 0.427 | 0.907 | 0.716 | 1.150 | Dietary advice, brisk wlking |


| Exposure |  |
| :--- | :--- |
| (metabolic trait) | Method |
| Omega_6_pct | Weighted median |
| Omega_6_pct | MR Egger |
| XL_VLDL_CE | Inverse variance weighted |
| PUFA_pct | Weighted median |
| PUFA_pct | Simple mode |
| PUFA_pct | Weighted mode |
| PUFA_pct | MR Egger |
| XL_VLDL_C | Inverse variance weighted |
| Pyruvate | MR Egger |
| Pyruvate | Weighted median |
| Pyruvate | Weighted mode |
| Pyruvate | Simple mode |
| XL_HDL_TG | Inverse variance weighted |
| Weighted median |  |


| Number of snps | p-val | Odds ratio | Odds ratio lower $95 \% \text { CI }$ |
| :---: | :---: | :---: | :---: |
| 36 | 0.461 | 0.916 | 0.725 |
| 36 | 0.246 | 0.842 | 0.633 |
| 42 | 0.369 | 1.067 | 0.926 |
| 35 | 0.730 | 0.962 | 0.774 |
| 35 | 0.945 | 0.987 | 0.675 |
| 35 | 0.699 | 1.046 | 0.833 |
| 35 | 0.562 | 1.078 | 0.838 |
| 37 | 0.372 | 1.073 | 0.920 |
| 15 | 0.470 | 1.212 | 0.730 |
| 15 | 0.477 | 1.107 | 0.837 |
| 15 | 0.312 | 1.154 | 0.883 |
| 15 | 0.465 | 1.217 | 0.729 |
| 46 | 0.110 | 1.081 | 0.983 |
| 31 | 0.770 | 1.026 | 0.863 |

Odds ratio
upper 95\% CI Intervention
1.157 Dietary advice, brisk wlking
1.120 Dietary advice, brisk wlking
1.229 Brisk Walking
1.197 Brisk Walking
1.442 Brisk Walking
1.314 Brisk Walking
1.387 Brisk Walking
1.251 Brisk Walking
2.013 Dietary advice
1.464 Dietary advice
1.509 Dietary advice
2.030 Dietary advice
1.189 Brisk Walking
1.221 Dietary advice

| Exposure <br> (metabolic trait) | Method |
| :--- | :--- |
| S_HDL_P | Simple mode |
| S_HDL_P | Weighted mode |
| S_HDL_P | MR Egger |
| XXL_VLDL_CE | Inverse variance weighted |
| SFA | Weighted median |
| SFA | Weighted mode |
| SFA | MR Egger |
| SFA | Simple mode |
| XL_VLDL_C | Weighted mode |
| XL_VLDL_C | Weighted median |
| XL_VLDL_C | MR Egger |
| XL_VLDL_C | Simple mode |
| XL_VLDL_PL | Inverse variance weighted |
| XL_VLDL_CE | Weighted mode |


| Number of snps | p-val | $\begin{aligned} & \text { Odds } \\ & \text { ratio } \end{aligned}$ | Odds ratio lower 95\% CI | Odds ratio upper 95\% CI | Intervention |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 31 | 0.730 | 1.049 | 0.800 | 1.377 | Dietary advice |
| 31 | 0.522 | 1.058 | 0.893 | 1.253 | Dietary advice |
| 31 | 0.556 | 1.075 | 0.848 | 1.362 | Dietary advice |
| 43 | 0.164 | 1.085 | 0.967 | 1.217 | Brisk Walking |
| 40 | 0.366 | 1.087 | 0.907 | 1.303 | Brisk Walking |
| 40 | 0.351 | 1.094 | 0.908 | 1.318 | Brisk Walking |
| 40 | 0.618 | 1.066 | 0.830 | 1.369 | Brisk Walking |
| 40 | 0.348 | 1.155 | 0.858 | 1.553 | Brisk Walking |
| 37 | 0.218 | 1.130 | 0.933 | 1.369 | Brisk Walking |
| 37 | 0.148 | 1.136 | 0.956 | 1.351 | Brisk Walking |
| 37 | 0.222 | 1.195 | 0.902 | 1.584 | Brisk Walking |
| 37 | 0.248 | 1.168 | 0.901 | 1.514 | Brisk Walking |
| 39 | 0.159 | 1.093 | 0.966 | 1.236 | Brisk Walking |
| 42 | 0.389 | 1.087 | 0.901 | 1.311 | Brisk Walking |


| Exposure <br> (metabolic trait) | Method |
| :--- | :--- |
| XL_VLDL_CE | Weighted median |
| XL_VLDL_CE | Simple mode |
| XL_VLDL_CE | MR Egger |
| XXL_VLDL_P | Inverse variance weighted |
| XL_VLDL_FC | Weighted mode |
| XL_VLDL_FC | Weighted median |
| XL_VLDL_FC | Simple mode |
| XL_VLDL_FC | MR Egger |
| SFA | Inverse variance weighted |
| XXL_VLDL_C | Inverse variance weighted |
| XL_VLDL_PL | Weighted median |
| XL_VLDL_PL | Simple mode |
| XL_VLDL_PL | Weighted mode |
| XL_VLDL_PL | MR Egger |


| Number <br> of snps | p-val | Odds <br> ratio | Odds ratio lower <br> $\mathbf{9 5 \%} \mathbf{C I}$ | Odds ratio <br> upper 95\% CI | Intervention |
| :---: | :---: | :---: | :---: | ---: | :--- |
| 42 | 0.311 | 1.094 | 0.920 | 1.301 | Brisk Walking |


| Exposure <br> (metabolic trait) | Method |
| :--- | :--- |
| XL_HDL_TG | Weighted median |
| XL_HDL_TG | Weighted mode |
| XXL_VLDL_FC | Inverse variance weighted |
| XL_HDL_TG | MR Egger |
| XL_HDL_TG | Simple mode |
| XXL_VLDL_C | Weighted median |
| XXL_VLDL_C | Weighted mode |
| XXL_VLDL_C | Simple mode |
| XXL_VLDL_C | MR Egger |
| XXL_VLDL_L | Inverse variance weighted |
| XXL_VLDL_CE | Weighted median |
| XXL_VLDL_CE | Weighted mode |
| XXL_VLDL_CE | Simple mode |
| XXL_VLDL_CE | MR Egger |


| Number of snps | p-val | $\begin{aligned} & \text { Odds } \\ & \text { ratio } \end{aligned}$ | Odds ratio lower $95 \% \text { CI }$ | Odds ratio upper 95\% CI | Intervention |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 46 | 0.573 | 1.034 | 0.920 | 1.164 | Brisk Walking |
| 46 | 0.515 | 1.036 | 0.932 | 1.151 | Brisk Walking |
| 44 | 0.090 | 1.109 | 0.984 | 1.251 | Brisk Walking |
| 46 | 0.873 | 0.989 | 0.860 | 1.136 | Brisk Walking |
| 46 | 0.128 | 1.212 | 0.950 | 1.545 | Brisk Walking |
| 43 | 0.390 | 1.075 | 0.911 | 1.270 | Brisk Walking |
| 43 | 0.275 | 1.109 | 0.923 | 1.334 | Brisk Walking |
| 43 | 0.277 | 1.170 | 0.885 | 1.548 | Brisk Walking |
| 43 | 0.183 | 1.141 | 0.942 | 1.382 | Brisk Walking |
| 45 | 0.080 | 1.111 | 0.987 | 1.251 | Brisk Walking |
| 43 | 0.404 | 1.073 | 0.909 | 1.266 | Brisk Walking |
| 43 | 0.437 | 1.081 | 0.890 | 1.313 | Brisk Walking |
| 43 | 0.333 | 1.154 | 0.866 | 1.538 | Brisk Walking |
| 43 | 0.351 | 1.094 | 0.907 | 1.320 | Brisk Walking |


| Exposure <br> (metabolic trait) | Method |
| :--- | :--- |
| XL_VLDL_FC | Inverse variance weighted |
| XXL_VLDL_FC | Weighted median |
| XXL_VLDL_FC | Weighted mode |
| XXL_VLDL_FC | Simple mode |
| XXL_VLDL_FC | MR Egger |
| L_VLDL_CE | Inverse variance weighted |
| XXL_VLDL_PL | Inverse variance weighted |
| XXL_VLDL_L | Weighted median |
| XXL_VLDL_L | Weighted mode |
| XXL_VLDL_L | Simple mode |
| XXL_VLDL_L | MR Egger |
| XXL_VLDL_P | Weighted median |
| XXL_VLDL_P | Weighted mode |
| XXL_VLDL_P | Simple mode |


| Number <br> of snps | p-val | Odds <br> ratio | Odds ratio lower <br> $\mathbf{9 5 \%} \mathbf{C I}$ | Odds ratio <br> upper 95\% CI | Intervention |
| :---: | :---: | :---: | :---: | ---: | :--- |
| 34 | 0.079 | 1.121 | 0.987 | 1.273 | Brisk Walking |


| Exposure <br> (metabolic trait) | Method |
| :--- | :--- |
| XXL_VLDL_P | MR Egger |
| M_VLDL_CE | Inverse variance weighted |
| XXL_VLDL_PL | Weighted median |
| XXL_VLDL_PL | Weighted mode |
| XXL_VLDL_PL | Simple mode |
| XXL_VLDL_PL | MR Egger |
| Ala | Inverse variance weighted |
| SFA_pct | Simple mode |
| SFA_pct | Weighted mode |
| SFA_pct | Weighted median |
| Lactate | Inverse variance weighted |
| SFA_pct | MR Egger |


| Number <br> of snps | p-val | Odds <br> ratio | Odds ratio lower <br> $\mathbf{9 5 \%} \mathbf{C I}$ | Odds ratio <br> upper 95\% CI | Intervention |
| ---: | :---: | :---: | :---: | ---: | :--- |
| 42 | 0.077 | 1.193 | 0.986 | 1.443 | Brisk Walking |
| 38 | 0.033 | 1.163 | 1.013 | 1.336 | Brisk Walking |
| 45 | 0.285 | 1.104 | 0.921 | 1.325 | Brisk Walking |
| 45 | 0.297 | 1.118 | 0.908 | 1.377 | Brisk Walking |
| 45 | 0.266 | 1.205 | 0.871 | 1.666 | Brisk Walking |
| 45 | 0.209 | 1.138 | 0.933 | 1.389 | Brisk Walking |
| 22 | 0.057 | 1.240 | 0.994 | 1.547 | Lycopene |
| 20 | 0.717 | 0.916 | 0.573 | 1.464 | Brisk Walking |
| 20 | 0.880 | 0.968 | 0.637 | 1.470 | Brisk Walking |
| 20 | 0.840 | 1.034 | 0.748 | 1.429 | Brisk Walking |
| 6 | 0.086 | 1.488 | 0.945 | 2.343 | Dietary advice |
| 20 | 0.728 | 0.852 | 0.351 | 2.067 | Brisk Walking |

Appendix C Table C 19: F-statistic and r-square measures for instruments of individual 6 extra metabolites from the dietary intervention arm (top $10 \%$ metabolites in the unadjusted ITT)

| Metabolite | Total r2 | Number of <br> SNPs | Sample <br> Size | F statistic |
| :--- | :---: | :---: | :---: | ---: |
| Total lipids in small HDL | 0.0421592 |  | 37 | 115078 |
| Glucose | 0.01125766 | 16 | 114867 | 136.9 |
| Phospholipids in small HDL | 0.04218063 | 43 | 115078 | 117.8 |
| Triglycerides in medium HDL | 0.04937673 | 41 | 115078 | 145.7 |
| Free cholesterol in small HDL | 0.03781267 | 37 | 115078 | 122.2 |
| Cholesterol in small HDL | 0.03808213 | 36 | 115078 | 126.5 |

Appendix C Table C 20: Mendelian Randomization results for Prostate Cancer survival in the PRACTICAL consortium of select metabolites (sensitivity analysis, metabolites taken forward from PrEvENT trial, top 10\% effect sizes)

| Exposure (metabolic trait) | Outcome | Method | Number of snps | p-value | Odds ratio | Odds ratio lower 95 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Glucose | PCa mortality | MR Egger | 16 | 0.362 | 1.270 |  |
| Glucose | PCa mortality | Simple mode | 16 | 0.830 | 0.943 |  |
| Glucose | PCa mortality | Inverse variance weighted | 16 | 0.849 | 0.978 |  |
| Glucose | PCa mortality | Weighted median | 16 | 0.731 | 1.054 |  |
| Glucose | PCa mortality | Weighted mode | 16 | 0.544 | 1.103 |  |
| S_HDL_PL | PCa mortality | Inverse variance weighted | 43 | 0.996 | 1.000 |  |
| S_HDL_PL | PCa mortality | Simple mode | 43 | 0.821 | 1.033 |  |
| S_HDL_PL | PCa mortality | Weighted median | 43 | 0.559 | 1.051 |  |
| S_HDL_PL | PCa mortality | Weighted mode | 43 | 0.501 | 1.054 |  |
| S_HDL_PL | PCa mortality | MR Egger | 43 | 0.536 | 1.088 |  |
| S_HDL_L | PCa mortality | Inverse variance weighted | 37 | 0.675 | 1.028 |  |
| S_HDL_L | PCa mortality | Weighted median | 37 | 0.501 | 1.059 |  |


| S_HDL_L | PCa mortality | Weighted mode | 37 | 0.461 | 1.063 |
| :---: | :---: | :---: | :---: | :---: | :---: |
| S_HDL_L | PCa mortality | MR Egger | 37 | 0.986 | 1.002 |
| S_HDL_L | PCa mortality | Simple mode | 37 | 0.482 | 1.100 |
| M_HDL_TG | PCa mortality | Inverse variance weighted | 40 | 0.783 | 1.017 |
| M_HDL_TG | PCa mortality | MR Egger | 40 | 0.817 | 1.024 |
| M_HDL_TG | PCa mortality | Weighted median | 40 | 0.537 | 1.052 |
| M_HDL_TG | PCa mortality | Weighted mode | 40 | 0.456 | 1.058 |
| M_HDL_TG | PCa mortality | Simple mode | 40 | 0.463 | 1.120 |
| S_HDL_FC | PCa mortality | Inverse variance weighted | 37 | 0.453 | 1.057 |
| S_HDL_FC | PCa mortality | Simple mode | 37 | 0.604 | 1.070 |
| S_HDL_FC | PCa mortality | Weighted mode | 37 | 0.496 | 1.070 |
| S_HDL_FC | PCa mortality | Weighted median | 37 | 0.440 | 1.075 |
| S_HDL_FC | PCa mortality | MR Egger | 37 | 0.819 | 1.034 |
| S_HDL_C | PCa mortality | Weighted median | 36 | 0.729 | 1.033 |
| S_HDL_C | PCa mortality | Inverse variance weighted | 36 | 0.548 | 1.043 |
| S_HDL_C | PCa mortality | Simple mode | 36 | 0.706 | 1.051 |
| S_HDL_C | PCa mortality | Weighted mode | 36 | 0.565 | 1.051 |
| S_HDL_C | PCa mortality | MR Egger | 36 | 0.959 | 1.007 |

### 9.4. Appendix D

Appendix D Table D 1: Cox regression results for individual metabolites and PCa death in the minimally and fully adjusted model, cases only.

| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xxl_vldl_p | 2093 | 1.381 | 1.064 | 1.793 | 0.015 | Main | PCAdeath |
| xxl_vldl_1 | 2085 | 1.446 | 1.098 | 1.904 | 0.009 | Main | PCAdeath |
| xxl_vldl_pl | 2085 | 1.456 | 1.111 | 1.908 | 0.006 | Main | PCAdeath |
| xxl_vldl_c | 2087 | 1.383 | 1.038 | 1.844 | 0.027 | Main | PCAdeath |
| xxl_vldl_ce | 2088 | 1.302 | 0.967 | 1.754 | 0.082 | Main | PCAdeath |
| xxl_vldl_fc | 2086 | 1.446 | 1.101 | 1.898 | 0.008 | Main | PCAdeath |
| xxl_vldl_tg | 2086 | 1.445 | 1.101 | 1.896 | 0.008 | Main | PCAdeath |
| xl_vldl_p | 2089 | 1.401 | 1.062 | 1.849 | 0.017 | Main | PCAdeath |
| xl_vldl_l | 2088 | 1.412 | 1.068 | 1.867 | 0.015 | Main | PCAdeath |
| xl_vldl_pl | 2087 | 1.413 | 1.069 | 1.869 | 0.015 | Main | PCAdeath |
| xl_vldl_c | 2089 | 1.373 | 1.034 | 1.823 | 0.028 | Main | PCAdeath |
| xl_vldl_ce | 2088 | 1.357 | 1.016 | 1.813 | 0.039 | Main | PCAdeath |
| xl_vldl_fc | 2088 | 1.402 | 1.061 | 1.851 | 0.017 | Main | PCAdeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| l_vldl_p | 2089 | 1.391 | 1.041 | 1.859 | 0.026 | Main | PCAdeath |
| l_vldl_1 | 2089 | 1.392 | 1.042 | 1.859 | 0.025 | Main | PCAdeath |
| l_vldl_pl | 2089 | 1.386 | 1.037 | 1.851 | 0.027 | Main | PCAdeath |
| l_vldl_c | 2088 | 1.376 | 1.025 | 1.846 | 0.033 | Main | PCAdeath |
| l_vldl_ce | 2088 | 1.339 | 0.989 | 1.812 | 0.059 | Main | PCAdeath |
| l_vldl_fc | 2089 | 1.396 | 1.050 | 1.857 | 0.02 | Main | PCAdeath |
| l_vldl_tg | 2089 | 1.399 | 1.048 | 1.866 | 0.022 | Main | PCAdeath |
| m_vldl_p | 2088 | 1.346 | 0.991 | 1.828 | 0.058 | Main | PCAdeath |
| m_vldl_l | 2088 | 1.340 | 0.985 | 1.821 | 0.062 | Main | PCAdeath |
| m_vldl_pl | 2089 | 1.324 | 0.976 | 1.798 | 0.071 | Main | PCAdeath |
| m_vldl_c | 2089 | 1.277 | 0.934 | 1.747 | 0.126 | Main | PCAdeath |
| m_vldl_ce | 2091 | 1.208 | 0.874 | 1.670 | 0.251 | Main | PCAdeath |
| m_vldl_fc | 2088 | 1.343 | 0.989 | 1.824 | 0.059 | Main | PCAdeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| s_vldl_p | 2091 | 1.268 | 0.918 | 1.752 | 0.149 | Main | PCAdeath |
| s_vldl_1 | 2092 | 1.258 | 0.910 | 1.739 | 0.164 | Main | PCAdeath |
| s_vldl_pl | 2092 | 1.295 | 0.941 | 1.782 | 0.113 | Main | PCAdeath |
| s_vldl_c | 2092 | 1.131 | 0.812 | 1.575 | 0.468 | Main | PCAdeath |
| s_vldl_ce | 2092 | 1.043 | 0.744 | 1.463 | 0.806 | Main | PCAdeath |
| s_vldl_fc | 2092 | 1.266 | 0.919 | 1.745 | 0.149 | Main | PCAdeath |
| s_vldl_tg | 2089 | 1.297 | 0.947 | 1.778 | 0.105 | Main | PCAdeath |
| xs_vldl_p | 2092 | 1.047 | 0.748 | 1.463 | 0.790 | Main | PCAdeath |
| xs_vldl_1 | 2092 | 1.029 | 0.734 | 1.441 | 0.870 | Main | PCAdeath |
| xs_vldl_pl | 2093 | 0.898 | 0.634 | 1.273 | 0.546 | Main | PCAdeath |
| xs_vldl_c | 2093 | 1.015 | 0.721 | 1.429 | 0.932 | Main | PCAdeath |
| xs_vldl_ce | 2093 | 1.053 | 0.750 | 1.480 | 0.765 | Main | PCAdeath |
| xs_vldl_fc | 2092 | 0.935 | 0.662 | 1.321 | 0.705 | Main | PCAdeath |
| xs_vldl_tg | 2092 | 1.204 | 0.875 | 1.657 | 0.255 | Main | PCAdeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | Outcome


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| m_ldl_p | 2093 | 0.879 | 0.619 | 1.247 | 0.470 | Main | PCAdeath |
| m_ldl_1 | 2093 | 0.874 | 0.616 | 1.241 | 0.453 | Main | PCAdeath |
| m_ldl_pl | 2093 | 1.000 | 0.710 | 1.409 | 0.999 | Main | PCAdeath |
| m_ldl_c | 2093 | 0.836 | 0.588 | 1.189 | 0.318 | Main | PCAdeath |
| m_ldl_ce | 2093 | 0.820 | 0.576 | 1.165 | $0.268 ~ M a i n$ | PCAdeath |  |
| m_ldl_fc | 2093 | 0.914 | 0.645 | 1.294 | 0.611 | Main | PCAdeath |
| m_ldl_tg | 2092 | 1.060 | 0.763 | 1.472 | 0.729 | Main | PCAdeath |
| s_ldl_p | 2093 | 0.917 | 0.648 | 1.298 | 0.626 | Main | PCAdeath |
| s_ldl_1 | 2093 | 0.910 | 0.643 | 1.290 | 0.598 | Main | PCAdeath |
| s_ldl_pl | 2093 | 1.095 | 0.781 | 1.534 | 0.600 | Main | PCAdeath |
| s_ldl_c | 2093 | 0.840 | 0.591 | 1.193 | 0.329 | Main | PCAdeath |
| s_ldl_ce | 2093 | 0.814 | 0.573 | 1.157 | 0.252 | Main | PCAdeath |
| s_ldl_fc | 2093 | 0.962 | 0.681 | 1.358 | 0.825 | Main | PCAdeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_hdl_p | 2091 | 1.037 | 0.746 | 1.441 | 0.829 | Main | PCAdeath |
| xl_hdl_1 | 2091 | 1.045 | 0.753 | 1.452 | 0.791 | Main | PCAdeath |
| xl_hdl_pl | 2091 | 0.926 | 0.657 | 1.306 | 0.661 | Main | PCAdeath |
| xl_hdl_c | 2091 | 1.168 | 0.847 | 1.611 | 0.342 | Main | PCAdeath |
| xl_hdl_ce | 2091 | 1.187 | 0.861 | 1.637 | 0.295 | Main | PCAdeath |
| xl_hdl_fc | 2091 | 1.128 | 0.817 | 1.557 | 0.464 | Main | PCAdeath |
| xl_hdl_tg | 2088 | 1.268 | 0.932 | 1.724 | 0.131 | Main | PCAdeath |
| l_hdl_p | 2092 | 0.928 | 0.654 | 1.316 | 0.675 | Main | PCAdeath |
| l_hdl_1 | 2092 | 0.925 | 0.652 | 1.312 | 0.661 | Main | PCAdeath |
| l_hdl_pl | 2093 | 0.918 | 0.647 | 1.302 | 0.631 | Main | PCAdeath |
| l_hdl_c | 2092 | 0.929 | 0.655 | 1.318 | 0.680 | Main | PCAdeath |
| l_hdl_ce | 2092 | 0.939 | 0.663 | 1.331 | 0.725 | Main | PCAdeath |
| l_hdl_fc | 2092 | 0.898 | 0.631 | 1.278 | 0.549 | Main | PCAdeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_hdl_p | 2093 | 1.141 | 0.819 | 1.590 | 0.436 | Main | PCAdeath |
| m_hdl_1 | 2093 | 1.124 | 0.805 | 1.569 | 0.493 | Main | PCAdeath |
| m_hdl_pl | 2092 | 1.167 | 0.838 | 1.623 | 0.361 | Main | PCAdeath |
| m_hdl_c | 2093 | 1.053 | 0.749 | 1.481 | 0.765 | Main | PCAdeath |
| m_hdl_ce | 2093 | 1.056 | 0.751 | 1.484 | 0.754 | Main | PCAdeath |
| m_hdl_fc | 2093 | 1.046 | 0.743 | 1.472 | 0.798 | Main | PCAdeath |
| m_hdl_tg | 2092 | 1.306 | 0.945 | 1.805 | 0.106 | Main | PCAdeath |
| s_hdl_p | 2093 | 1.306 | 0.947 | 1.802 | 0.104 | Main | PCAdeath |
| s_hdl_l | 2093 | 1.268 | 0.916 | 1.755 | 0.152 | Main | PCAdeath |
| s_hdl_pl | 2093 | 1.454 | 1.071 | 1.974 | 0.016 | Main | PCAdeath |
| s_hdl_c | 2087 | 0.848 | 0.606 | 1.187 | 0.337 | Main | PCAdeath |
| s_hdl_ce | 2087 | 0.784 | 0.568 | 1.084 | 0.141 | Main | PCAdeath |
| s_hdl_fc | 2093 | 1.466 | 1.073 | 2.005 | 0.016 | Main | PCAdeath |
| s_hdl_tg | 2091 | 1.362 | 0.997 | 1.860 | 0.052 | Main | PCAdeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| vldl_d | 2093 | 1.440 | 1.023 | 2.028 | 0.037 | Main | PCAdeath |
| ldl_d | 2093 | 0.886 | 0.631 | 1.245 | 0.485 | Main | PCAdeath |
| hdl_d | 2093 | 0.902 | 0.639 | 1.274 | 0.559 | Main | PCAdeath |
| serum_c | 2093 | 0.948 | 0.672 | 1.339 | 0.763 | Main | PCAdeath |
| vldl_c | 2093 | 1.248 | 0.904 | 1.721 | 0.178 | Main | PCAdeath |
| remnant_c | 2093 | 1.098 | 0.785 | 1.536 | 0.585 | Main | PCAdeath |
| ldl_c | 2093 | 0.843 | 0.593 | 1.199 | 0.342 | Main | PCAdeath |
| hdl_c | 2093 | 0.989 | 0.703 | 1.391 | 0.948 | Main | PCAdeath |
| hdl2_c | 2093 | 0.980 | 0.696 | 1.381 | 0.909 | Main | PCAdeath |
| hdl3_c | 2093 | 1.037 | 0.741 | 1.452 | 0.831 | Main | PCAdeath |
| estc | 2090 | 0.947 | 0.670 | 1.337 | 0.757 | Main | PCAdeath |
| freec | 2090 | 0.950 | 0.673 | 1.340 | 0.769 | Main | PCAdeath |
| serum_tg | 2093 | 1.334 | 0.990 | 1.796 | 0.058 | Main | PCAdeath |


| Metabolite | $\mathbf{N}$ | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| ldl_tg | 2093 | 1.092 | 0.789 | 1.513 | 0.596 | Main | PCAdeath |
| hdl_tg | 2093 | 1.273 | 0.933 | 1.736 | 0.127 | Main | PCAdeath |
| totpg | 2090 | 1.197 | 0.859 | 1.667 | 0.288 | Main | PCAdeath |
| tg_pg | 2090 | 1.316 | 0.961 | 1.802 | 0.087 | Main | PCAdeath |
| pc | 2090 | 1.127 | 0.807 | 1.573 | 0.483 | Main | PCAdeath |
| sm | 2090 | 1.131 | 0.802 | 1.595 | 0.484 | Main | PCAdeath |
| totcho | 2090 | 1.121 | 0.803 | 1.566 | 0.503 | Main | PCAdeath |
| apoa1 | 2093 | 1.078 | 0.772 | 1.505 | 0.659 | Main | PCAdeath |
| apob | 2092 | 1.091 | 0.780 | 1.527 | 0.610 | Main | PCAdeath |
| apob_apoa1 | 2092 | 1.060 | 0.755 | 1.489 | 0.736 | Main | PCAdeath |
| totfa | 2089 | 1.279 | 0.931 | 1.757 | 0.129 | Main | PCAdeath |
| unsat | 2089 | 0.863 | 0.614 | 1.213 | 0.397 | Main | PCAdeath |
| dha | 2089 | 1.239 | 0.942 | 1.628 | 0.125 | Main | PCAdeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| faw3 | 2089 | 1.298 | 0.983 | 1.714 | 0.066 | Main | PCAdeath |
| faw6 | 2089 | 1.036 | 0.738 | 1.454 | 0.839 | Main | PCAdeath |
| pufa | 2089 | 1.099 | 0.786 | 1.536 | 0.583 | Main | PCAdeath |
| mufa | 2089 | 1.304 | 0.962 | 1.766 | 0.087 | Main | PCAdeath |
| sfa | 2089 | 1.337 | 0.986 | 1.814 | 0.062 | Main | PCAdeath |
| dha_fa | 2089 | 1.072 | 0.775 | 1.484 | 0.673 | Main | PCAdeath |
| la_fa | 2089 | 0.675 | 0.518 | 0.879 | 0.003 | Main | PCAdeath |
| faw3_fa | 2089 | 1.169 | 0.866 | 1.579 | 0.308 | Main | PCAdeath |
| faw6_fa | 2089 | 0.669 | 0.494 | 0.905 | 0.009 | Main | PCAdeath |
| pufa_fa | 2089 | 0.721 | 0.528 | 0.986 | 0.040 | Main | PCAdeath |
| mufa_fa | 2089 | 1.245 | 0.910 | 1.704 | 0.171 | Main | PCAdeath |
| sfa_fa | 2089 | 1.379 | 1.010 | 1.882 | 0.043 | Main | PCAdeath |
| glc | 2087 | 1.228 | 0.979 | 1.539 | 0.076 | Main | PCAdeath |


| Metabolite | $\mathbf{N}$ | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| pyr | 2091 | 1.029 | 0.823 | 1.286 | 0.803 | Main | PCAdeath |
| cit | 2092 | 0.967 | 0.684 | 1.367 | 0.848 | Main | PCAdeath |
| ala | 2093 | 0.870 | 0.616 | 1.229 | 0.430 | Main | PCAdeath |
| gln | 2092 | 0.908 | 0.648 | 1.270 | 0.572 | Main | PCAdeath |
| gly | 2084 | 0.565 | 0.374 | 0.854 | 0.007 | Main | PCAdeath |
| his | 2088 | 1.028 | 0.745 | 1.417 | 0.867 | Main | PCAdeath |
| ile | 2092 | 1.563 | 1.177 | 2.075 | 0.002 | Main | PCAdeath |
| leu | 2093 | 1.532 | 1.157 | 2.028 | 0.003 | Main | PCAdeath |
| val | 2092 | 1.423 | 1.055 | 1.919 | 0.021 | Main | PCAdeath |
| phe | 2093 | 0.909 | 0.649 | 1.273 | 0.580 | Main | PCAdeath |
| tyr | 2088 | 1.259 | 0.933 | 1.698 | 0.131 | Main | PCAdeath |
| ace | 2093 | 1.171 | 1.044 | 1.313 | 0.007 | Main | PCAdeath |
| acace | 2093 | 1.274 | 1.061 | 1.529 | 0.009 | Main | PCAdeath |
| bohbut | 2041 | 1.183 | 0.955 | 1.465 | 0.125 | Main | PCAdeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| crea | 2087 | 1.064 | 0.773 | 1.463 | 0.704 | Main | PCAdeath |
| alb | 2093 | 1.146 | 0.807 | 1.628 | 0.446 | Main | PCAdeath |
| gp | 2093 | 1.285 | 0.937 | 1.761 | 0.119 | Main | PCAdeath |
| xxl_vldl_p | 2039 | 1.207 | 0.885 | 1.648 | 0.235 | FullyAdjusted | PCAdeath |
| xxl_vldl_1 | 2031 | 1.231 | 0.904 | 1.678 | 0.187 | FullyAdjusted | PCAdeath |
| xxl_vldl_pl | 2031 | 1.242 | 0.915 | 1.686 | 0.164 | FullyAdjusted | PCAdeath |
| xxl_vldl_c | 2033 | 1.194 | 0.858 | 1.663 | 0.294 | FullyAdjusted | PCAdeath |
| xxl_vldl_ce | 2034 | 1.145 | 0.808 | 1.622 | 0.446 | FullyAdjusted | PCAdeath |
| xxl_vldl_fc | 2032 | 1.230 | 0.903 | 1.675 | 0.189 | FullyAdjusted | PCAdeath |
| xxl_vldl_tg | 2032 | 1.233 | 0.908 | 1.675 | 0.179 | FullyAdjusted | PCAdeath |
| xl_vldl_p | 2035 | 1.215 | 0.882 | 1.674 | 0.234 | FullyAdjusted | PCAdeath |
| xl_vldl_l | 2034 | 1.217 | 0.883 | 1.677 | 0.231 | FullyAdjusted | PCAdeath |
| xl_vldl_pl | 2033 | 1.214 | 0.883 | 1.671 | 0.233 | FullyAdjusted | PCAdeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_vldl_ce | 2034 | 1.174 | 0.836 | 1.649 | 0.354 | FullyAdjusted | PCAdeath |
| xl_vldl_fc | 2034 | 1.209 | 0.877 | 1.667 | 0.246 | FullyAdjusted | PCAdeath |
| xl_vldl_tg | 2034 | 1.222 | 0.888 | 1.681 | 0.218 | FullyAdjusted | PCAdeath |
| l_vldl_p | 2035 | 1.212 | 0.867 | 1.695 | 0.260 | FullyAdjusted | PCAdeath |
| l_vldl_1 | 2035 | 1.212 | 0.867 | 1.695 | 0.261 | FullyAdjusted | PCAdeath |
| l_vldl_pl | 2035 | 1.210 | 0.865 | 1.692 | 0.266 | FullyAdjusted | PCAdeath |
| l_vldl_c | 2034 | 1.191 | 0.846 | 1.677 | 0.315 | FullyAdjusted | PCAdeath |
| l_vldl_ce | 2034 | 1.158 | 0.812 | 1.651 | 0.418 | FullyAdjusted | PCAdeath |
| l_vldl_fc | 2035 | 1.214 | 0.873 | 1.690 | 0.250 | FullyAdjusted | PCAdeath |
| l_vldl_tg | 2035 | 1.219 | 0.874 | 1.700 | 0.243 | FullyAdjusted | PCAdeath |
| m_vldl_p | 2034 | 1.174 | 0.825 | 1.670 | 0.372 | FullyAdjusted | PCAdeath |
| m_vldl_l | 2034 | 1.168 | 0.821 | 1.664 | 0.388 | FullyAdjusted | PCAdeath |
| m_vldl_pl | 2035 | 1.160 | 0.815 | 1.650 | 0.410 | FullyAdjusted | PCAdeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :---: | ---: | :---: | :---: | :---: | :---: | :---: |
| m_vldl_ce | 2037 | 1.058 | 0.726 | 1.541 | 0.770 | FullyAdjusted | PCAdeath |
| m_vldl_fc | 2034 | 1.173 | 0.825 | 1.669 | 0.375 | FullyAdjusted | PCAdeath |
| m_vldl_tg | 2035 | 1.190 | 0.840 | 1.687 | 0.328 | FullyAdjusted | PCAdeath |
| s_vldl_p | 2037 | 1.123 | 0.774 | 1.629 | 0.540 | FullyAdjusted | PCAdeath |
| s_vldl_1 | 2038 | 1.120 | 0.771 | 1.629 | 0.552 | FullyAdjusted | PCAdeath |
| s_vldl_pl | 2038 | 1.147 | 0.790 | 1.664 | 0.471 | FullyAdjusted | PCAdeath |
| s_vldl_c | 2038 | 1.052 | 0.719 | 1.537 | 0.795 | FullyAdjusted | PCAdeath |
| s_vldl_ce | 2038 | 0.999 | 0.683 | 1.461 | 0.995 | FullyAdjusted | PCAdeath |
| s_vldl_fc | 2038 | 1.136 | 0.782 | 1.649 | 0.504 | FullyAdjusted | PCAdeath |
| s_vldl_tg | 2035 | 1.140 | 0.793 | 1.638 | 0.479 | FullyAdjusted | PCAdeath |
| xs_vldl_p | 2038 | 1.039 | 0.713 | 1.513 | 0.843 | FullyAdjusted | PCAdeath |
| xs_vldl_1 | 2038 | 1.032 | 0.708 | 1.504 | 0.870 | FullyAdjusted | PCAdeath |
| xs_vldl_pl | 2039 | 0.969 | 0.665 | 1.413 | 0.872 | FullyAdjusted | PCAdeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xs_vldl_ce | 2039 | 1.043 | 0.712 | 1.526 | 0.830 | FullyAdjusted | PCAdeath |
| xs_vldl_fc | 2038 | 0.995 | 0.685 | 1.446 | 0.980 | FullyAdjusted | PCAdeath |
| xs_vldl_tg | 2038 | 1.106 | 0.766 | 1.598 | 0.591 | FullyAdjusted | PCAdeath |
| idl_p | 2039 | 0.944 | 0.643 | 1.385 | 0.768 | FullyAdjusted | PCAdeath |
| idl_1 | 2039 | 0.933 | 0.636 | 1.370 | 0.724 | FullyAdjusted | PCAdeath |
| idl_pl | 2039 | 0.904 | 0.615 | 1.327 | 0.605 | FullyAdjusted | PCAdeath |
| idl_c | 2039 | 0.935 | 0.637 | 1.373 | 0.733 | FullyAdjusted | PCAdeath |
| idl_ce | 2039 | 0.954 | 0.650 | 1.401 | 0.811 | FullyAdjusted | PCAdeath |
| idl_fc | 2039 | 0.897 | 0.614 | 1.312 | 0.576 | FullyAdjusted | PCAdeath |
| idl_tg | 2038 | 1.067 | 0.738 | 1.541 | 0.731 | FullyAdjusted | PCAdeath |
| l_ldl_p | 2039 | 0.947 | 0.646 | 1.386 | 0.778 | FullyAdjusted | PCAdeath |
| l_ldl_l | 2039 | 0.938 | 0.640 | 1.373 | 0.740 | FullyAdjusted | PCAdeath |
| l_ldl_pl | 2039 | 0.937 | 0.639 | 1.376 | 0.741 | FullyAdjusted | PCAdeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| l_ldl_ce | 2039 | 0.938 | 0.642 | 1.373 | 0.743 | FullyAdjusted | PCAdeath |
| l_ldl_fc | 2039 | 0.905 | 0.618 | 1.324 | 0.606 | FullyAdjusted | PCAdeath |
| l_ldl_tg | 2038 | 1.055 | 0.729 | 1.527 | 0.775 | FullyAdjusted | PCAdeath |
| m_ldl_p | 2039 | 0.961 | 0.658 | 1.402 | 0.835 | FullyAdjusted | PCAdeath |
| m_ldl_1 | 2039 | 0.958 | 0.657 | 1.398 | 0.824 | FullyAdjusted | PCAdeath |
| m_ldl_pl | 2039 | 1.027 | 0.702 | 1.503 | 0.891 | FullyAdjusted | PCAdeath |
| m_ldl_c | 2039 | 0.937 | 0.643 | 1.365 | 0.735 | FullyAdjusted | PCAdeath |
| m_ldl_ce | 2039 | 0.925 | 0.636 | 1.345 | 0.683 | FullyAdjusted | PCAdeath |
| m_ldl_fc | 2039 | 0.990 | 0.676 | 1.449 | 0.958 | FullyAdjusted | PCAdeath |
| m_ldl_tg | 2038 | 1.066 | 0.737 | 1.540 | 0.735 | FullyAdjusted | PCAdeath |
| s_ldl_p | 2039 | 0.996 | 0.683 | 1.453 | 0.983 | FullyAdjusted | PCAdeath |
| s_ldl_l | 2039 | 0.994 | 0.681 | 1.451 | 0.976 | FullyAdjusted | PCAdeath |
| s_ldl_pl | 2039 | 1.120 | 0.767 | 1.635 | 0.558 | FullyAdjusted | PCAdeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :---: |
| s_ldl_ce | 2039 | 0.928 | 0.639 | 1.349 | 0.697 | FullyAdjusted | PCAdeath |
| s_ldl_fc | 2039 | 1.036 | 0.708 | 1.518 | 0.854 | FullyAdjusted | PCAdeath |
| s_ldl_tg | 2038 | 1.154 | 0.803 | 1.658 | 0.440 | FullyAdjusted | PCAdeath |
| xl_hdl_p | 2037 | 1.104 | 0.763 | 1.595 | 0.600 | FullyAdjusted | PCAdeath |
| xl_hdl_1 | 2037 | 1.113 | 0.770 | 1.607 | 0.569 | FullyAdjusted | PCAdeath |
| xl_hdl_pl | 2037 | 1.013 | 0.696 | 1.476 | 0.946 | FullyAdjusted | PCAdeath |
| xl_hdl_c | 2037 | 1.223 | 0.850 | 1.760 | 0.279 | FullyAdjusted | PCAdeath |
| xl_hdl_ce | 2037 | 1.238 | 0.860 | 1.783 | 0.251 | FullyAdjusted | PCAdeath |
| xl_hdl_fc | 2037 | 1.185 | 0.824 | 1.704 | 0.361 | FullyAdjusted | PCAdeath |
| xl_hdl_tg | 2034 | 1.124 | 0.787 | 1.604 | 0.520 | FullyAdjusted | PCAdeath |
| l_hdl_p | 2038 | 0.992 | 0.684 | 1.438 | 0.965 | FullyAdjusted | PCAdeath |
| l_hdl_l | 2038 | 0.991 | 0.683 | 1.439 | 0.964 | FullyAdjusted | PCAdeath |
| l_hdl_pl | 2039 | 0.989 | 0.683 | 1.431 | 0.952 | FullyAdjusted | PCAdeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| l_hdl_ce | 2038 | 1.003 | 0.691 | 1.456 | 0.986 | FullyAdjusted | PCAdeath |
| l_hdl_fc | 2038 | 0.989 | 0.678 | 1.443 | 0.953 | FullyAdjusted | PCAdeath |
| l_hdl_tg | 2038 | 0.912 | 0.609 | 1.367 | 0.656 | FullyAdjusted | PCAdeath |
| m_hdl_p | 2039 | 1.079 | 0.772 | 1.506 | 0.657 | FullyAdjusted | PCAdeath |
| m_hdl_1 | 2039 | 1.073 | 0.767 | 1.502 | 0.679 | FullyAdjusted | PCAdeath |
| m_hdl_pl | 2038 | 1.099 | 0.786 | 1.536 | 0.580 | FullyAdjusted | PCAdeath |
| m_hdl_c | 2039 | 1.048 | 0.744 | 1.477 | 0.788 | FullyAdjusted | PCAdeath |
| m_hdl_ce | 2039 | 1.048 | 0.744 | 1.475 | 0.790 | FullyAdjusted | PCAdeath |
| m_hdl_fc | 2039 | 1.051 | 0.742 | 1.488 | 0.780 | FullyAdjusted | PCAdeath |
| m_hdl_tg | 2038 | 1.085 | 0.751 | 1.568 | 0.665 | FullyAdjusted | PCAdeath |
| s_hdl_p | 2039 | 1.205 | 0.851 | 1.706 | 0.294 | FullyAdjusted | PCAdeath |
| s_hdl_l | 2039 | 1.189 | 0.837 | 1.689 | 0.333 | FullyAdjusted | PCAdeath |
| s_hdl_pl | 2039 | 1.222 | 0.879 | 1.698 | 0.232 | FullyAdjusted | PCAdeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :---: |
| s_hdl_ce | 2033 | 0.909 | 0.633 | 1.307 | 0.608 | FullyAdjusted | PCAdeath |
| s_hdl_fc | 2039 | 1.277 | 0.909 | 1.794 | 0.159 | FullyAdjusted | PCAdeath |
| s_hdl_tg | 2037 | 1.166 | 0.809 | 1.681 | 0.410 | FullyAdjusted | PCAdeath |
| vldl_d | 2039 | 1.160 | 0.802 | 1.679 | 0.430 | FullyAdjusted | PCAdeath |
| ldl_d | 2039 | 0.877 | 0.601 | 1.280 | 0.495 | FullyAdjusted | PCAdeath |
| hdl_d | 2039 | 0.982 | 0.673 | 1.433 | 0.924 | FullyAdjusted | PCAdeath |
| serum_c | 2039 | 1.004 | 0.685 | 1.471 | 0.985 | FullyAdjusted | PCAdeath |
| vldl_c | 2039 | 1.120 | 0.770 | 1.630 | 0.553 | FullyAdjusted | PCAdeath |
| remnant_c | 2039 | 1.044 | 0.714 | 1.525 | 0.825 | FullyAdjusted | PCAdeath |
| ldl_c | 2039 | 0.934 | 0.640 | 1.363 | 0.723 | FullyAdjusted | PCAdeath |
| hdl_c | 2039 | 1.063 | 0.741 | 1.525 | 0.739 | FullyAdjusted | PCAdeath |
| hdl2_c | 2039 | 1.054 | 0.735 | 1.513 | 0.775 | FullyAdjusted | PCAdeath |
| hdl3_c | 2039 | 1.097 | 0.758 | 1.589 | 0.623 | FullyAdjusted | PCAdeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| freec | 2036 | 0.990 | 0.672 | 1.457 | 0.958 | FullyAdjusted | PCAdeath |
| serum_tg | 2039 | 1.176 | 0.828 | 1.671 | 0.366 | FullyAdjusted | PCAdeath |
| vldl_tg | 2039 | 1.186 | 0.839 | 1.676 | 0.335 | FullyAdjusted | PCAdeath |
| ldl_tg | 2039 | 1.081 | 0.747 | 1.564 | 0.680 | FullyAdjusted | PCAdeath |
| hdl_tg | 2039 | 1.088 | 0.748 | 1.582 | 0.658 | FullyAdjusted | PCAdeath |
| totpg | 2036 | 1.109 | 0.768 | 1.603 | 0.581 | FullyAdjusted | PCAdeath |
| tg_pg | 2036 | 1.105 | 0.771 | 1.584 | 0.587 | FullyAdjusted | PCAdeath |
| pc | 2036 | 1.065 | 0.733 | 1.548 | 0.741 | FullyAdjusted | PCAdeath |
| sm | 2036 | 1.207 | 0.818 | 1.782 | 0.342 | FullyAdjusted | PCAdeath |
| totcho | 2036 | 1.096 | 0.750 | 1.602 | 0.636 | FullyAdjusted | PCAdeath |
| apoa1 | 2039 | 1.103 | 0.770 | 1.578 | 0.593 | FullyAdjusted | PCAdeath |
| apob | 2038 | 1.055 | 0.725 | 1.535 | 0.781 | FullyAdjusted | PCAdeath |
| apob_apoa1 | 2038 | 1.007 | 0.696 | 1.457 | 0.969 | FullyAdjusted | PCAdeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| unsat | 2035 | 0.960 | 0.669 | 1.377 | 0.825 | FullyAdjusted | PCAdeath |
| dha | 2035 | 1.103 | 0.777 | 1.564 | 0.584 | FullyAdjusted | PCAdeath |
| la | 2035 | 0.989 | 0.669 | 1.461 | 0.956 | FullyAdjusted | PCAdeath |
| faw3 | 2035 | 1.136 | 0.797 | 1.620 | 0.480 | FullyAdjusted | PCAdeath |
| faw6 | 2035 | 1.040 | 0.706 | 1.531 | 0.843 | FullyAdjusted | PCAdeath |
| pufa | 2035 | 1.065 | 0.726 | 1.563 | 0.748 | FullyAdjusted | PCAdeath |
| mufa | 2035 | 1.119 | 0.788 | 1.590 | 0.530 | FullyAdjusted | PCAdeath |
| sfa | 2035 | 1.178 | 0.828 | 1.676 | 0.364 | FullyAdjusted | PCAdeath |
| dha_fa | 2035 | 1.040 | 0.728 | 1.486 | 0.830 | FullyAdjusted | PCAdeath |
| la_fa | 2035 | 0.796 | 0.565 | 1.121 | 0.192 | FullyAdjusted | PCAdeath |
| faw3_fa | 2035 | 1.089 | 0.770 | 1.541 | 0.630 | FullyAdjusted | PCAdeath |
| faw6_fa | 2035 | 0.848 | 0.597 | 1.204 | 0.356 | FullyAdjusted | PCAdeath |
| pufa_fa | 2035 | 0.886 | 0.624 | 1.256 | 0.495 | FullyAdjusted | PCAdeath |


| Metabolite | $\mathbf{N}$ | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :---: | ---: | :---: | :---: | :---: | :---: | :---: |
| sfa_fa | 2035 | 1.273 | 0.883 | 1.837 | 0.196 | FullyAdjusted | PCAdeath |
| glc | 2033 | 1.166 | 0.881 | 1.544 | 0.281 | FullyAdjusted | PCAdeath |
| lac | 2039 | 0.815 | 0.502 | 1.323 | 0.407 | FullyAdjusted | PCAdeath |
| pyr | 2037 | 0.995 | 0.724 | 1.369 | 0.977 | FullyAdjusted | PCAdeath |
| cit | 2038 | 0.980 | 0.671 | 1.433 | 0.919 | FullyAdjusted | PCAdeath |
| ala | 2039 | 0.834 | 0.569 | 1.222 | 0.352 | FullyAdjusted | PCAdeath |
| gln | 2038 | 0.967 | 0.668 | 1.400 | 0.857 | FullyAdjusted | PCAdeath |
| gly | 2030 | 0.552 | 0.352 | 0.866 | 0.010 | FullyAdjusted | PCAdeath |
| his | 2034 | 0.826 | 0.585 | 1.164 | 0.275 | FullyAdjusted | PCAdeath |
| ile | 2038 | 1.128 | 0.803 | 1.585 | 0.488 | FullyAdjusted | PCAdeath |
| leu | 2039 | 1.112 | 0.795 | 1.557 | 0.535 | FullyAdjusted | PCAdeath |
| val | 2038 | 1.074 | 0.762 | 1.515 | 0.683 | FullyAdjusted | PCAdeath |
| phe | 2039 | 0.707 | 0.475 | 1.050 | 0.086 | FullyAdjusted | PCAdeath |
| ph | 2034 | 0.954 | 0.663 | 1.372 | 0.798 | FullyAdjusted | PCAdeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper 95\% CI | p-value | Model | Outcome |
| :--- | :---: | ---: | ---: | ---: | ---: | :---: | :---: |
| ace | 2039 | 1.197 | 1.042 | 1.374 | 0.011 | FullyAdjusted | PCAdeath |
| acace | 2039 | 1.253 | 1.024 | 1.533 | 0.029 | FullyAdjusted | PCAdeath |
| bohbut | 1988 | 1.206 | 0.975 | 1.491 | 0.085 | FullyAdjusted | PCAdeath |
| crea | 2033 | 1.260 | 0.900 | 1.763 | 0.179 | FullyAdjusted | PCAdeath |
| alb | 2039 | 1.075 | 0.737 | 1.570 | 0.706 | FullyAdjusted | PCAdeath |
| gp | 2039 | 1.225 | 0.852 | 1.762 | 0.274 | FullyAdjusted | PCAdeath |

Appendix D Table D 2: Cox regression results for individual metabolites and all-cause death in the minimally and fully adjusted model, case only.

| Metabolite | N | Hazard <br> Ratio | Lower <br> $\mathbf{9 5 \%} \mathbf{C I}$ | Upper <br> $\mathbf{9 5 \%} \mathbf{C I}$ | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xxl_vldl_p_z | 2093 | 1.159 | 1.034 | 1.301 | 0.012 | Main | allcausedeath |
| xxl_vldl_l_z | 2085 | 1.171 | 1.040 | 1.319 | 0.009 | Main | allcausedeath |
| xxl_vldl_pl_z | 2085 | 1.176 | 1.045 | 1.323 | 0.007 | Main | allcausedeath |
| xxl_vldl_c_z | 2087 | 1.133 | 1.003 | 1.280 | 0.045 | Main | allcausedeath |
| xxl_vldl_ce_z | 2088 | 1.083 | 0.957 | 1.225 | 0.208 | Main | allcausedeath |
| xxl_vldl_fc_z | 2086 | 1.177 | 1.046 | 1.324 | 0.007 | Main | allcausedeath |
| xxl_vldl_tg_z | 2086 | 1.171 | 1.041 | 1.317 | 0.009 | Main | allcausedeath |
| xl_vldl_p_z | 2089 | 1.131 | 1.004 | 1.274 | 0.043 | Main | allcausedeath |
| xl_vldl_l_z | 2088 | 1.139 | 1.010 | 1.284 | 0.033 | Main | allcausedeath |
| xl_vldl_pl_z | 2087 | 1.145 | 1.016 | 1.291 | 0.026 | Main | allcausedeath |


| Metabolite | N | Hazard <br> Ratio | Lower <br> $\mathbf{9 5 \%} \mathbf{C I}$ | Upper <br> $\mathbf{9 5 \%} \mathbf{C I}$ | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :--- |
| l_vldl_p_z | 2089 | 1.111 | 0.983 | 1.256 | 0.091 | Main | allcausedeath |
| l_vldl_1_z | 2089 | 1.111 | 0.982 | 1.255 | 0.094 | Main | allcausedeath |
| l_vldl_pl_z | 2089 | 1.111 | 0.983 | 1.255 | 0.092 | Main | allcausedeath |
| l_vldl_c_z | 2088 | 1.091 | 0.964 | 1.235 | 0.167 | Main | allcausedeath |
| l_vldl_ce_z | 2088 | 1.056 | 0.930 | 1.199 | 0.399 | Main | allcausedeath |
| l_vldl_fc_z | 2089 | 1.120 | 0.992 | 1.264 | 0.067 | Main | allcausedeath |
| l_vldl_tg_z | 2089 | 1.118 | 0.989 | 1.263 | 0.074 | Main | allcausedeath |
| m_vldl_p_z | 2088 | 1.059 | 0.933 | 1.202 | 0.375 | Main | allcausedeath |
| m_vldl_1_z | 2088 | 1.057 | 0.931 | 1.200 | 0.393 | Main | allcausedeath |
| m_vldl_pl_z | 2089 | 1.049 | 0.924 | 1.192 | 0.457 | Main | allcausedeath |
| m_vldl_c_z | 2089 | 1.024 | 0.901 | 1.165 | 0.714 | Main | allcausedeath |
| m_vldl_ce_z | 2091 | 0.987 | 0.865 | 1.126 | 0.845 | Main | allcausedeath |
| m_vldl_fc_z | 2088 | 1.065 | 0.938 | 1.208 | 0.331 | Main | allcausedeath |
| m_vldl_tg_z | 2089 | 1.070 | 0.943 | 1.214 | 0.296 | Main | allcausedeath |


| Metabolite | N | Hazard <br> Ratio | Lower <br> $\mathbf{9 5 \%} \mathbf{C I}$ | Upper <br> $\mathbf{9 5 \%} \mathbf{C I}$ | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_vldl_p_z | 2091 | 0.977 | 0.857 | 1.114 | 0.728 | Main | allcausedeath |
| s_vldl_l_z | 2092 | 0.990 | 0.869 | 1.129 | 0.882 | Main | allcausedeath |
| s_vldl_pl_z | 2092 | 1.002 | 0.880 | 1.141 | 0.975 | Main | allcausedeath |
| s_vldl_c_z | 2092 | 0.917 | 0.803 | 1.047 | 0.202 | Main | allcausedeath |
| s_vldl_ce_z | 2092 | 0.882 | 0.772 | 1.009 | 0.067 | Main | allcausedeath |
| s_vldl_fc_z | 2092 | 0.986 | 0.865 | 1.123 | 0.828 | Main | allcausedeath |
| s_vldl_tg_z | 2089 | 1.013 | 0.891 | 1.153 | 0.839 | Main | allcausedeath |
| xs_vldl_p_z | 2092 | 0.946 | 0.829 | 1.079 | 0.409 | Main | allcausedeath |
| xs_vldl_l_z | 2092 | 0.937 | 0.821 | 1.069 | 0.331 | Main | allcausedeath |
| xs_vldl_pl_z | 2093 | 0.892 | 0.780 | 1.019 | 0.093 | Main | allcausedeath |


| Metabolite | N | Hazard <br> Ratio | Lower <br> $\mathbf{9 5 \%} \mathbf{C I}$ | Upper <br> $\mathbf{9 5 \%} \mathbf{C I}$ | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| idl_p_z | 2093 | 0.925 | 0.810 | 1.056 | 0.249 | Main | allcausedeath |
| idl_l_z | 2093 | 0.917 | 0.803 | 1.048 | 0.202 | Main | allcausedeath |
| idl_pl_z | 2093 | 0.889 | 0.777 | 1.016 | 0.084 | Main | allcausedeath |
| idl_c_z | 2093 | 0.915 | 0.801 | 1.045 | 0.190 | Main | allcausedeath |
| idl_ce_z | 2093 | 0.936 | 0.819 | 1.069 | 0.326 | Main | allcausedeath |
| idl_fc_z | 2093 | 0.871 | 0.763 | 0.995 | 0.042 | Main | allcausedeath |
| idl_tg_z | 2092 | 1.049 | 0.925 | 1.190 | 0.460 | Main | allcausedeath |
| l_ldl_p_z | 2093 | 0.892 | 0.781 | 1.020 | 0.096 | Main | allcausedeath |
| l_ldl_l_z | 2093 | 0.886 | 0.775 | 1.013 | 0.076 | Main | allcausedeath |
| l_ldl_pl_z | 2093 | 0.892 | 0.780 | 1.020 | 0.096 | Main | allcausedeath |


| Metabolite | N | Hazard <br> Ratio | Lower <br> $\mathbf{9 5 \%} \mathbf{C I}$ | Upper <br> $\mathbf{9 5 \%} \mathbf{C I}$ | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_ldl_p_z | 2093 | 0.861 | 0.753 | 0.985 | 0.029 | Main | allcausedeath |
| m_ldl_l_z | 2093 | 0.856 | 0.748 | 0.979 | 0.023 | Main | allcausedeath |
| m_ldl_pl_z | 2093 | 0.922 | 0.807 | 1.054 | 0.235 | Main | allcausedeath |
| m_ldl_c_z | 2093 | 0.833 | 0.728 | 0.952 | 0.008 | Main | allcausedeath |
| m_ldl_ce_z | 2093 | 0.826 | 0.722 | 0.944 | 0.005 | Main | allcausedeath |
| m_ldl_fc_z | 2093 | 0.867 | 0.758 | 0.991 | 0.037 | Main | allcausedeath |
| m_ldl_tg_z | 2092 | 1.037 | 0.913 | 1.177 | 0.579 | Main | allcausedeath |
| s_ldl_p_z | 2093 | 0.859 | 0.751 | 0.983 | 0.027 | Main | allcausedeath |
| s_ldl_l_z | 2093 | 0.855 | 0.747 | 0.978 | 0.022 | Main | allcausedeath |
| s_ldl_pl_z | 2093 | 0.936 | 0.819 | 1.069 | 0.328 | Main | allcausedeath |


| Metabolite | N | Hazard <br> Ratio | Lower <br> $\mathbf{9 5 \%} \mathbf{C I}$ | Upper <br> $\mathbf{9 5 \%} \mathbf{C I}$ | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_hdl_p_z | 2091 | 1.102 | 0.972 | 1.249 | 0.129 | Main | allcausedeath |
| xl_hdl_l_z | 2091 | 1.102 | 0.972 | 1.249 | 0.131 | Main | allcausedeath |
| xl_hdl_pl_z | 2091 | 1.088 | 0.959 | 1.234 | 0.190 | Main | allcausedeath |
| xl_hdl_c_z | 2091 | 1.095 | 0.965 | 1.244 | 0.159 | Main | allcausedeath |
| xl_hdl_ce_z | 2091 | 1.080 | 0.951 | 1.227 | 0.235 | Main | allcausedeath |
| xl_hdl_fc_z | 2091 | 1.131 | 0.998 | 1.282 | 0.053 | Main | allcausedeath |
| xl_hdl_tg_z | 2088 | 1.173 | 1.040 | 1.323 | 0.009 | Main | allcausedeath |
| l_hdl_p_z | 2092 | 1.129 | 0.996 | 1.281 | 0.058 | Main | allcausedeath |
| l_hdl_l_z | 2092 | 1.127 | 0.993 | 1.278 | 0.064 | Main | allcausedeath |


| Metabolite | N | Hazard <br> Ratio | Lower $95 \% \text { CI }$ | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_hdl_p_z | 2093 | 1.170 | 1.030 | 1.329 | 0.016 | Main | allcausedeath |
| m_hdl_l_z | 2093 | 1.161 | 1.022 | 1.319 | 0.022 | Main | allcausedeath |
| m_hdl_pl_z | 2092 | 1.170 | 1.030 | 1.330 | 0.016 | Main | allcausedeath |
| m_hdl_c_z | 2093 | 1.115 | 0.979 | 1.270 | 0.100 | Main | allcausedeath |
| m_hdl_ce_z | 2093 | 1.117 | 0.981 | 1.272 | 0.094 | Main | allcausedeath |
| m_hdl_fc_z | 2093 | 1.107 | 0.972 | 1.261 | 0.127 | Main | allcausedeath |
| m_hdl_tg_z | 2092 | 1.149 | 1.012 | 1.305 | 0.032 | Main | allcausedeath |
| s_hdl_p_z | 2093 | 1.064 | 0.934 | 1.213 | 0.348 | Main | allcausedeath |
| s_hdl_l_z | 2093 | 1.046 | 0.918 | 1.192 | 0.500 | Main | allcausedeath |
| s_hdl_pl_z | 2093 | 1.231 | 1.088 | 1.394 | 0.001 | Main | allcausedeath |
| s_hdl_c_z | 2087 | 0.824 | 0.726 | 0.937 | 0.003 | Main | allcausedeath |
| s_hdl_ce_z | 2087 | 0.803 | 0.710 | 0.909 | 0.001 | Main | allcausedeath |
| s_hdl_fc_z | 2093 | 1.141 | 1.004 | 1.296 | 0.043 | Main | allcausedeath |
| s_hdl_tg_z | 2091 | 1.091 | 0.961 | 1.238 | 0.177 | Main | allcausedeath |


| Metabolite | N | Hazard <br> Ratio | Lower <br> $\mathbf{9 5 \%} \mathbf{C I}$ | Upper <br> $\mathbf{9 5 \%} \mathbf{C I}$ | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| vldl_d_z | 2093 | 1.148 | 1.007 | 1.308 | 0.040 | Main | allcausedeath |
| ldl_d_z | 2093 | 1.257 | 1.110 | 1.422 | 0.000 | Main | allcausedeath |
| hdl_d_z | 2093 | 1.091 | 0.960 | 1.240 | 0.181 | Main | allcausedeath |
| serum_c_z | 2093 | 0.923 | 0.808 | 1.055 | 0.239 | Main | allcausedeath |
| vldl_c_z | 2093 | 1.013 | 0.889 | 1.155 | 0.842 | Main | allcausedeath |
| remnant_c_z | 2093 | 0.968 | 0.848 | 1.104 | 0.628 | Main | allcausedeath |
| ldl_c_z | 2093 | 0.849 | 0.743 | 0.971 | 0.017 | Main | allcausedeath |
| hdl_c_z | 2093 | 1.078 | 0.948 | 1.227 | 0.252 | Main | allcausedeath |
| hdl2_c_z | 2093 | 1.084 | 0.953 | 1.233 | 0.217 | Main | allcausedeath |
| hdl3_c_z | 2093 | 0.987 | 0.867 | 1.125 | 0.849 | Main | allcausedeath |


| Metabolite | N | Hazard <br> Ratio | Lower $95 \% \text { CI }$ | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| ldl_tg_z | 2093 | 1.055 | 0.929 | 1.197 | 0.410 | Main | allcausedeath |
| hdl_tg_z | 2093 | 1.176 | 1.041 | 1.327 | 0.009 | Main | allcausedeath |
| totpg_z | 2090 | 1.079 | 0.948 | 1.228 | 0.252 | Main | allcausedeath |
| tg_pg_z | 2090 | 1.091 | 0.960 | 1.240 | 0.182 | Main | allcausedeath |
| pc_Z | 2090 | 1.062 | 0.933 | 1.210 | 0.361 | Main | allcausedeath |
| sm_Z | 2090 | 0.976 | 0.854 | 1.115 | 0.721 | Main | allcausedeath |
| totcho_z | 2090 | 1.065 | 0.935 | 1.213 | 0.344 | Main | allcausedeath |
| apoa1_z | 2093 | 1.098 | 0.965 | 1.249 | 0.155 | Main | allcausedeath |
| apob_z | 2092 | 0.959 | 0.840 | 1.094 | 0.532 | Main | allcausedeath |
| apob_apoa1_z | 2092 | 0.931 | 0.816 | 1.062 | 0.288 | Main | allcausedeath |
| totfa_z | 2089 | 1.055 | 0.928 | 1.200 | 0.411 | Main | allcausedeath |
| unsat_z | 2089 | 0.775 | 0.683 | 0.879 | 0.000 | Main | allcausedeath |
| dha_z | 2089 | 0.893 | 0.776 | 1.026 | 0.110 | Main | allcausedeath |
| la_z | 2089 | 0.917 | 0.804 | 1.048 | 0.203 | Main | allcausedeath |


| Metabolite | N | Hazard <br> Ratio | Lower <br> $\mathbf{9 5 \%} \mathbf{C I}$ | Upper <br> $\mathbf{9 5 \%} \mathbf{C I}$ | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| faw3_z | 2089 | 0.888 | 0.773 | 1.020 | 0.092 | Main | allcausedeath |
| faw6_z | 2089 | 0.933 | 0.817 | 1.066 | 0.307 | Main | allcausedeath |
| pufa_z | 2089 | 0.920 | 0.805 | 1.051 | 0.220 | Main | allcausedeath |
| mufa_z | 2089 | 1.118 | 0.989 | 1.265 | 0.076 | Main | allcausedeath |
| sfa_z | 2089 | 1.097 | 0.967 | 1.244 | 0.149 | Main | allcausedeath |
| dha_fa_z | 2089 | 0.830 | 0.719 | 0.957 | 0.011 | Main | allcausedeath |
| la_fa_z | 0.795 | 0.707 | 0.893 | 0.000 | Main | allcausedeath |  |
| faw3_fa_z | 2089 | 0.789 | 0.681 | 0.914 | 0.002 | Main | allcausedeath |
| faw6_fa_z | 2089 | 0.791 | 0.701 | 0.893 | 0.000 | Main | allcausedeath |
| pufa_fa_z | 2089 | 0.765 | 0.678 | 0.862 | 0.000 | Main | allcausedeath |


| Metabolite | N | Hazard <br> Ratio | Lower <br> $\mathbf{9 5 \%} \mathbf{C I}$ | Upper <br> $\mathbf{9 5 \%} \mathbf{C I}$ | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :--- |
| pyr_z | 2091 | 0.998 | 0.900 | 1.106 | 0.964 | Main | allcausedeath |
| cit_z | 2092 | 1.050 | 0.921 | 1.196 | 0.467 | Main | allcausedeath |
| ala_z | 2093 | 0.991 | 0.872 | 1.126 | 0.887 | Main | allcausedeath |
| gln_z | 2092 | 0.890 | 0.782 | 1.014 | 0.079 | Main | allcausedeath |
| gly_z | 2084 | 0.944 | 0.827 | 1.077 | 0.392 | Main | allcausedeath |
| his_z | 2088 | 0.997 | 0.879 | 1.131 | 0.960 | Main | allcausedeath |
| ile_z | 2092 | 1.059 | 0.932 | 1.204 | 0.377 | Main | allcausedeath |
| leu_z | 2093 | 1.057 | 0.932 | 1.200 | 0.387 | Main | allcausedeath |
| val_z | 2092 | 0.957 | 0.839 | 1.091 | 0.507 | Main | allcausedeath |
| phe_z | 2093 | 1.165 | 1.035 | 1.312 | 0.012 | Main | allcausedeath |
| tyr_z | 2088 | 1.058 | 0.934 | 1.199 | 0.373 | Main | allcausedeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper $95 \% \text { CI }$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| crea_Z | 2087 | 0.899 | 0.783 | 1.032 | 0.131 | Main | allcausedeath |
| alb_z | 2093 | 1.030 | 0.901 | 1.176 | 0.665 | Main | allcausedeath |
| gp_z | 2093 | 1.142 | 1.007 | 1.294 | 0.038 | Main | allcausedeath |
| xxl_vldl_p_z | 2039 | 1.132 | 1.002 | 1.279 | 0.047 | FullyAdjusted | allcausedeath |
| xxl_vldl_l_z | 2031 | 1.134 | 1.001 | 1.284 | 0.048 | FullyAdjusted | allcausedeath |
| xxl_vldl_pl_z | 2031 | 1.138 | 1.006 | 1.288 | 0.040 | FullyAdjusted | allcausedeath |
| xxl_vldl_c_z | 2033 | 1.106 | 0.972 | 1.257 | 0.125 | FullyAdjusted | allcausedeath |
| xxl_vldl_ce_z | 2034 | 1.066 | 0.936 | 1.214 | 0.335 | FullyAdjusted | allcausedeath |
| xxl_vldl_fc_z | 2032 | 1.137 | 1.004 | 1.287 | 0.043 | FullyAdjusted | allcausedeath |
| xxl_vldl_tg_z | 2032 | 1.133 | 1.001 | 1.283 | 0.048 | FullyAdjusted | allcausedeath |
| xl_vldl_p_z | 2035 | 1.095 | 0.965 | 1.241 | 0.158 | FullyAdjusted | allcausedeath |
| xl_vldl_l_z | 2034 | 1.102 | 0.971 | 1.251 | 0.131 | FullyAdjusted | allcausedeath |
| xl_vldl_pl_z | 2033 | 1.111 | 0.979 | 1.260 | 0.103 | FullyAdjusted | allcausedeath |
| xl_vldl_c_z | 2035 | 1.076 | 0.947 | 1.223 | 0.262 | FullyAdjusted | allcausedeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | $\begin{gathered} \text { Upper } \\ 95 \% \text { CI } \end{gathered}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_vldl_ce_z | 2034 | 1.058 | 0.928 | 1.205 | 0.399 | FullyAdjusted | allcausedeath |
| xl_vldl_fc_z | 2034 | 1.105 | 0.974 | 1.254 | 0.120 | FullyAdjusted | allcausedeath |
| xl_vldl_tg_z | 2034 | 1.106 | 0.975 | 1.255 | 0.119 | FullyAdjusted | allcausedeath |
| l_vldl_p_z | 2035 | 1.077 | 0.946 | 1.225 | 0.263 | FullyAdjusted | allcausedeath |
| l_vldl_l_z | 2035 | 1.076 | 0.945 | 1.225 | 0.268 | FullyAdjusted | allcausedeath |
| 1_vldl_pl_z | 2035 | 1.078 | 0.947 | 1.226 | 0.257 | FullyAdjusted | allcausedeath |
| l_vldl_c_z | 2034 | 1.060 | 0.930 | 1.208 | 0.382 | FullyAdjusted | allcausedeath |
| 1_vldl_ce_z | 2034 | 1.030 | 0.901 | 1.177 | 0.669 | FullyAdjusted | allcausedeath |
| 1_vldl_fc_z | 2035 | 1.085 | 0.955 | 1.233 | 0.212 | FullyAdjusted | allcausedeath |
| l_vldl_tg_z | 2035 | 1.081 | 0.951 | 1.230 | 0.234 | FullyAdjusted | allcausedeath |
| m_vldl_p_z | 2034 | 1.032 | 0.903 | 1.179 | 0.646 | FullyAdjusted | allcausedeath |
| m_vldl_l_z | 2034 | 1.030 | 0.902 | 1.178 | 0.660 | FullyAdjusted | allcausedeath |
| m_vldl_pl_z | 2035 | 1.026 | 0.897 | 1.173 | 0.708 | FullyAdjusted | allcausedeath |
| m_vldl_c_z | 2035 | 1.009 | 0.881 | 1.155 | 0.901 | FullyAdjusted | allcausedeath |


| Metabolite | N | Hazard Ratio | Lower $95 \% \text { CI }$ | $\begin{gathered} \text { Upper } \\ 95 \% \text { CI } \end{gathered}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_vldl_ce_z | 2037 | 0.981 | 0.854 | 1.127 | 0.787 | FullyAdjusted | allcausedeath |
| m_vldl_fc_z | 2034 | 1.038 | 0.909 | 1.186 | 0.583 | FullyAdjusted | allcausedeath |
| m_vldl_tg_z | 2035 | 1.038 | 0.909 | 1.186 | 0.578 | FullyAdjusted | allcausedeath |
| s_vldl_p_z | 2037 | 0.968 | 0.843 | 1.112 | 0.646 | FullyAdjusted | allcausedeath |
| S_vldl_l_z | 2038 | 0.987 | 0.860 | 1.132 | 0.850 | FullyAdjusted | allcausedeath |
| s_vldl_pl_z | 2038 | 0.997 | 0.869 | 1.143 | 0.966 | FullyAdjusted | allcausedeath |
| S_vldl_c_z | 2038 | 0.939 | 0.817 | 1.080 | 0.378 | FullyAdjusted | allcausedeath |
| s_vldl_ce_z | 2038 | 0.915 | 0.795 | 1.053 | 0.217 | FullyAdjusted | allcausedeath |
| s_vldl_fc_z | 2038 | 0.985 | 0.859 | 1.130 | 0.833 | FullyAdjusted | allcausedeath |
| s_vldl_tg_z | 2035 | 0.990 | 0.865 | 1.135 | 0.890 | FullyAdjusted | allcausedeath |
| xs_vldl_p_z | 2038 | 0.979 | 0.852 | 1.124 | 0.761 | FullyAdjusted | allcausedeath |
| xs_vldl_1_Z | 2038 | 0.973 | 0.847 | 1.118 | 0.698 | FullyAdjusted | allcausedeath |
| xs_vldl_pl_z | 2039 | 0.937 | 0.815 | 1.078 | 0.362 | FullyAdjusted | allcausedeath |
| xs_vldl_c_z | 2039 | 0.980 | 0.852 | 1.127 | 0.778 | FullyAdjusted | allcausedeath |


| Metabolite | N | Hazard Ratio | Lower $95 \% \text { CI }$ | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xs_vldl_ce_z | 2039 | 1.014 | 0.882 | 1.166 | 0.841 | FullyAdjusted | allcausedeath |
| xs_vldl_fc_z | 2038 | 0.912 | 0.793 | 1.048 | 0.194 | FullyAdjusted | allcausedeath |
| xs_vldl_tg_z | 2038 | 1.012 | 0.884 | 1.158 | 0.866 | FullyAdjusted | allcausedeath |
| idl_p_z | 2039 | 0.970 | 0.843 | 1.116 | 0.673 | FullyAdjusted | allcausedeath |
| idl_1_z | 2039 | 0.964 | 0.838 | 1.110 | 0.613 | FullyAdjusted | allcausedeath |
| idl_pl_z | 2039 | 0.939 | 0.816 | 1.081 | 0.382 | FullyAdjusted | allcausedeath |
| idl_c_z | 2039 | 0.964 | 0.838 | 1.109 | 0.607 | FullyAdjusted | allcausedeath |
| idl_ce_z | 2039 | 0.982 | 0.854 | 1.130 | 0.799 | FullyAdjusted | allcausedeath |
| idl_fc_z | 2039 | 0.924 | 0.804 | 1.063 | 0.270 | FullyAdjusted | allcausedeath |
| idl_tg_z | 2038 | 1.053 | 0.922 | 1.202 | 0.448 | FullyAdjusted | allcausedeath |
| l_ldl_p_z | 2039 | 0.941 | 0.817 | 1.083 | 0.393 | FullyAdjusted | allcausedeath |
| 1_ldl_l_z | 2039 | 0.935 | 0.812 | 1.076 | 0.348 | FullyAdjusted | allcausedeath |
| 1_ldl_pl_z | 2039 | 0.942 | 0.818 | 1.084 | 0.403 | FullyAdjusted | allcausedeath |
| 1_ldl_c_z | 2039 | 0.923 | 0.802 | 1.062 | 0.263 | FullyAdjusted | allcausedeath |


| Metabolite | N | Hazard <br> Ratio | Lower $95 \% \text { CI }$ | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| l_ldl_ce_z | 2039 | 0.925 | 0.803 | 1.064 | 0.275 | FullyAdjusted | allcausedeath |
| 1_ldl_fc_z | 2039 | 0.919 | 0.798 | 1.057 | 0.237 | FullyAdjusted | allcausedeath |
| l_ldl_tg_z | 2038 | 1.070 | 0.937 | 1.221 | 0.319 | FullyAdjusted | allcausedeath |
| m_ldl_p_z | 2039 | 0.909 | 0.790 | 1.047 | 0.186 | FullyAdjusted | allcausedeath |
| m_ldl_l_z | 2039 | 0.905 | 0.786 | 1.042 | 0.163 | FullyAdjusted | allcausedeath |
| m_ldl_pl_z | 2039 | 0.962 | 0.836 | 1.107 | 0.591 | FullyAdjusted | allcausedeath |
| m_ldl_c_z | 2039 | 0.884 | 0.768 | 1.018 | 0.087 | FullyAdjusted | allcausedeath |
| m_ldl_ce_z | 2039 | 0.878 | 0.763 | 1.010 | 0.070 | FullyAdjusted | allcausedeath |
| m_ldl_fc_z | 2039 | 0.914 | 0.793 | 1.052 | 0.210 | FullyAdjusted | allcausedeath |
| m_ldl_tg_z | 2038 | 1.048 | 0.916 | 1.198 | 0.496 | FullyAdjusted | allcausedeath |
| s_ldl_p_z | 2039 | 0.906 | 0.787 | 1.043 | 0.170 | FullyAdjusted | allcausedeath |
| s_ldl_l_Z | 2039 | 0.902 | 0.784 | 1.039 | 0.153 | FullyAdjusted | allcausedeath |
| s_ldl_pl_z | 2039 | 0.972 | 0.845 | 1.119 | 0.695 | FullyAdjusted | allcausedeath |
| s_ldl_c_Z | 2039 | 0.876 | 0.762 | 1.008 | 0.065 | FullyAdjusted | allcausedeath |


| Metabolite | N | Hazard Ratio | Lower $95 \% \text { CI }$ | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_ldl_ce_z | 2039 | 0.871 | 0.757 | 1.002 | 0.053 | FullyAdjusted | allcausedeath |
| s_ldl_fc_z | 2039 | 0.903 | 0.784 | 1.040 | 0.156 | FullyAdjusted | allcausedeath |
| s_ldl_tg_z | 2038 | 1.053 | 0.921 | 1.204 | 0.449 | FullyAdjusted | allcausedeath |
| xl_hdl_p_z | 2037 | 1.110 | 0.973 | 1.267 | 0.121 | FullyAdjusted | allcausedeath |
| xl_hdl_l_z | 2037 | 1.110 | 0.973 | 1.267 | 0.122 | FullyAdjusted | allcausedeath |
| xl_hdl_pl_z | 2037 | 1.096 | 0.960 | 1.252 | 0.175 | FullyAdjusted | allcausedeath |
| xl_hdl_c_z | 2037 | 1.104 | 0.966 | 1.263 | 0.146 | FullyAdjusted | allcausedeath |
| xl_hdl_ce_z | 2037 | 1.090 | 0.953 | 1.247 | 0.208 | FullyAdjusted | allcausedeath |
| xl_hdl_fc_z | 2037 | 1.136 | 0.995 | 1.296 | 0.058 | FullyAdjusted | allcausedeath |
| xl_hdl_tg_z | 2034 | 1.156 | 1.017 | 1.313 | 0.026 | FullyAdjusted | allcausedeath |
| 1_hdl_p_z | 2038 | 1.140 | 1.000 | 1.300 | 0.050 | FullyAdjusted | allcausedeath |
| 1_hdl_1_z | 2038 | 1.137 | 0.997 | 1.297 | 0.055 | FullyAdjusted | allcausedeath |
| 1_hdl_pl_z | 2039 | 1.129 | 0.989 | 1.289 | 0.072 | FullyAdjusted | allcausedeath |
| 1_hdl_c_z | 2038 | 1.131 | 0.992 | 1.290 | 0.066 | FullyAdjusted | allcausedeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper $95 \% \text { CI }$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| l_hdl_ce_z | 2038 | 1.139 | 0.999 | 1.298 | 0.052 | FullyAdjusted | allcausedeath |
| l_hdl_fc_z | 2038 | 1.105 | 0.967 | 1.262 | 0.141 | FullyAdjusted | allcausedeath |
| 1_hdl_tg_z | 2038 | 1.166 | 1.025 | 1.326 | 0.020 | FullyAdjusted | allcausedeath |
| m_hdl_p_z | 2039 | 1.174 | 1.030 | 1.339 | 0.017 | FullyAdjusted | allcausedeath |
| m_hdl_l_z | 2039 | 1.167 | 1.023 | 1.332 | 0.021 | FullyAdjusted | allcausedeath |
| m_hdl_pl_z | 2038 | 1.172 | 1.027 | 1.338 | 0.018 | FullyAdjusted | allcausedeath |
| m_hdl_c_z | 2039 | 1.125 | 0.984 | 1.287 | 0.084 | FullyAdjusted | allcausedeath |
| m_hdl_ce_z | 2039 | 1.126 | 0.985 | 1.287 | 0.083 | FullyAdjusted | allcausedeath |
| m_hdl_fc_z | 2039 | 1.124 | 0.983 | 1.286 | 0.089 | FullyAdjusted | allcausedeath |
| m_hdl_tg_z | 2038 | 1.134 | 0.993 | 1.296 | 0.064 | FullyAdjusted | allcausedeath |
| s_hdl_p_z | 2039 | 1.074 | 0.937 | 1.230 | 0.306 | FullyAdjusted | allcausedeath |
| s_hdl_l_z | 2039 | 1.059 | 0.924 | 1.214 | 0.412 | FullyAdjusted | allcausedeath |
| s_hdl_pl_z | 2039 | 1.202 | 1.056 | 1.368 | 0.005 | FullyAdjusted | allcausedeath |
| s_hdl_c_z | 2033 | 0.868 | 0.759 | 0.993 | 0.039 | FullyAdjusted | allcausedeath |


| Metabolite | N | Hazard <br> Ratio | Lower $95 \% \text { CI }$ | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_hdl_ce_z | 2033 | 0.848 | 0.743 | 0.966 | 0.013 | FullyAdjusted | allcausedeath |
| s_hdl_fc_z | 2039 | 1.125 | 0.984 | 1.286 | 0.084 | FullyAdjusted | allcausedeath |
| s_hdl_tg_z | 2037 | 1.056 | 0.924 | 1.208 | 0.424 | FullyAdjusted | allcausedeath |
| vldl_d_z | 2039 | 1.094 | 0.954 | 1.254 | 0.200 | FullyAdjusted | allcausedeath |
| ldl_d_z | 2039 | 1.253 | 1.099 | 1.428 | 0.001 | FullyAdjusted | allcausedeath |
| hdl_d_z | 2039 | 1.101 | 0.963 | 1.260 | 0.160 | FullyAdjusted | allcausedeath |
| serum_c_z | 2039 | 0.973 | 0.845 | 1.119 | 0.698 | FullyAdjusted | allcausedeath |
| vldl_c_z | 2039 | 1.019 | 0.888 | 1.169 | 0.790 | FullyAdjusted | allcausedeath |
| remnant_c_z | 2039 | 0.996 | 0.866 | 1.144 | 0.950 | FullyAdjusted | allcausedeath |
| ldl_c_z | 2039 | 0.900 | 0.782 | 1.037 | 0.144 | FullyAdjusted | allcausedeath |
| hdl_c_z | 2039 | 1.100 | 0.962 | 1.257 | 0.165 | FullyAdjusted | allcausedeath |
| hdl2_c_z | 2039 | 1.103 | 0.965 | 1.261 | 0.150 | FullyAdjusted | allcausedeath |
| hdl3_c_z | 2039 | 1.025 | 0.894 | 1.175 | 0.724 | FullyAdjusted | allcausedeath |
| estc_z | 2036 | 0.966 | 0.839 | 1.111 | 0.625 | FullyAdjusted | allcausedeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper $95 \% \text { CI }$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| freec_z | 2036 | 0.980 | 0.852 | 1.127 | 0.777 | FullyAdjusted | allcausedeath |
| serum_tg_Z | 2039 | 1.081 | 0.947 | 1.234 | 0.246 | FullyAdjusted | allcausedeath |
| vldl_tg_Z | 2039 | 1.074 | 0.941 | 1.226 | 0.292 | FullyAdjusted | allcausedeath |
| ldl_tg_z | 2039 | 1.065 | 0.932 | 1.217 | 0.356 | FullyAdjusted | allcausedeath |
| hdl_tg_z | 2039 | 1.159 | 1.019 | 1.319 | 0.025 | FullyAdjusted | allcausedeath |
| totpg_z | 2036 | 1.098 | 0.958 | 1.260 | 0.180 | FullyAdjusted | allcausedeath |
| tg_pg_z | 2036 | 1.050 | 0.917 | 1.201 | 0.480 | FullyAdjusted | allcausedeath |
| pc_z | 2036 | 1.082 | 0.943 | 1.241 | 0.260 | FullyAdjusted | allcausedeath |
| sm_z | 2036 | 1.029 | 0.893 | 1.186 | 0.690 | FullyAdjusted | allcausedeath |
| totcho_z | 2036 | 1.094 | 0.953 | 1.256 | 0.200 | FullyAdjusted | allcausedeath |
| apoa1_z | 2039 | 1.126 | 0.984 | 1.288 | 0.084 | FullyAdjusted | allcausedeath |
| apob_z | 2038 | 0.984 | 0.857 | 1.131 | 0.822 | FullyAdjusted | allcausedeath |
| apob_apoa1_z | 2038 | 0.946 | 0.824 | 1.086 | 0.429 | FullyAdjusted | allcausedeath |
| totfa_z | 2035 | 1.055 | 0.922 | 1.208 | 0.437 | FullyAdjusted | allcausedeath |


| Metabolite | N | Hazard <br> Ratio | Lower <br> $\mathbf{9 5 \%} \mathbf{C I}$ | Upper <br> $\mathbf{9 5 \%} \mathbf{C I}$ | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :--- | :--- |
| unsat_z | 2035 | 0.814 | 0.712 | 0.929 | 0.002 | FullyAdjusted allcausedeath |  |
| dha_z | 2035 | 0.890 | 0.768 | 1.030 | 0.119 | FullyAdjusted allcausedeath |  |
| la_z | 2035 | 0.947 | 0.822 | 1.090 | 0.447 | FullyAdjusted allcausedeath |  |
| faw3_z | 2035 | 0.884 | 0.764 | 1.023 | 0.098 | FullyAdjusted allcausedeath |  |
| faw6_z | 2035 | 0.965 | 0.838 | 1.111 | 0.621 | FullyAdjusted allcausedeath |  |
| pufa_z | 2035 | 0.946 | 0.822 | 1.090 | 0.442 | FullyAdjusted allcausedeath |  |
| mufa_z | 2035 | 1.101 | 0.967 | 1.254 | 0.146 | FullyAdjusted allcausedeath |  |
| sfa_z | 2035 | 1.087 | 0.952 | 1.241 | 0.216 | FullyAdjusted allcausedeath |  |
| dha_fa_z | 2035 | 0.828 | 0.713 | 0.962 | 0.014 | FullyAdjusted allcausedeath |  |
| la_fa_z | 2035 | 0.833 | 0.734 | 0.945 | 0.005 | FullyAdjusted allcausedeath |  |
| faw3_fa_z | 2035 | 0.787 | 0.675 | 0.919 | 0.002 | FullyAdjusted allcausedeath |  |
| faw6_fa_z | 2035 | 0.840 | 0.738 | 0.956 | 0.008 | FullyAdjusted allcausedeath |  |


| Metabolite | N | Hazard Ratio | Lower $95 \% \text { CI }$ | Upper $95 \% \mathrm{CI}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| sfa_fa_z | 2035 | 1.199 | 1.049 | 1.370 | 0.008 | FullyAdjusted | allcausedeath |
| glc_z | 2033 | 1.126 | 1.009 | 1.256 | 0.034 | FullyAdjusted | allcausedeath |
| lac_z | 2039 | 1.024 | 0.908 | 1.154 | 0.704 | FullyAdjusted | allcausedeath |
| pyr_z | 2037 | 0.976 | 0.858 | 1.110 | 0.710 | FullyAdjusted | allcausedeath |
| cit_Z | 2038 | 1.039 | 0.904 | 1.193 | 0.593 | FullyAdjusted | allcausedeath |
| ala_z | 2039 | 1.001 | 0.876 | 1.144 | 0.990 | FullyAdjusted | allcausedeath |
| $g l n \_z$ | 2038 | 0.923 | 0.805 | 1.058 | 0.251 | FullyAdjusted | allcausedeath |
| gly_z | 2030 | 0.965 | 0.841 | 1.106 | 0.605 | FullyAdjusted | allcausedeath |
| his_z | 2034 | 0.962 | 0.843 | 1.098 | 0.568 | FullyAdjusted | allcausedeath |
| ile_z | 2038 | 0.985 | 0.860 | 1.129 | 0.828 | FullyAdjusted | allcausedeath |
| leu_z | 2039 | 0.999 | 0.874 | 1.142 | 0.987 | FullyAdjusted | allcausedeath |
| val_z | 2038 | 0.924 | 0.805 | 1.059 | 0.255 | FullyAdjusted | allcausedeath |
| phe_z | 2039 | 1.130 | 0.995 | 1.284 | 0.060 | FullyAdjusted | allcausedeath |
| tyr_z | 2034 | 1.003 | 0.877 | 1.147 | 0.963 | FullyAdjusted | allcausedeath |


| Metabolite | N | Hazard <br> Ratio | Lower <br> $\mathbf{9 5 \%} \mathbf{C I}$ | Upper <br> $\mathbf{9 5 \%} \mathbf{C I}$ | p-value | Model | Outcome |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| ace_z | 2039 | 1.127 | 1.057 | 1.202 | 0.000 | FullyAdjusted allcausedeath |  |
| acace_z | 2039 | 1.157 | 1.043 | 1.284 | 0.006 | FullyAdjusted allcausedeath |  |
| bohbut_z | 1988 | 1.104 | 0.990 | 1.231 | 0.075 | FullyAdjusted allcausedeath |  |
| crea_z | 2033 | 0.927 | 0.803 | 1.070 | 0.301 | FullyAdjusted allcausedeath |  |
| alb_z | 2039 | 1.043 | 0.907 | 1.199 | 0.553 | FullyAdjusted allcausedeath |  |
| gp_z | 2039 | 1.121 | 0.982 | 1.281 | 0.092 | FullyAdjusted allcausedeath |  |

Appendix D Table D 3: Cox regression results for PCa-mRS and All-cause mRS, for all-cause and PCa specific mortality, in the minimally and fully adjusted regression models, cases only.

| Metabolomic model | Outcome | Regression model | Hazard ratio | Lower 95\%CI | Upper 95\% CI | p-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| PCa-mRS | PCa death | Main* | 1.16 | 0.83 | 1.61 | 0.38 |
| PCa-mRS | PCa death | Fully adjusted** | 1.32 | 0.91 | 1.9 | 0.12 |
| PCa-mRS | All-cause <br> death | Main* | 1.17 | 1.02 | 1.33 | 0.015 |
| PCa-mRS | All-cause <br> death | Fully adjusted** | 1.14 | 0.99 | 1.31 | 0.063 |
| All-cause mRS | PCa death | Main* | 0.95 | 0.67 | 1.4 | 0.14 |
| All-cause mRS | PCa death | Fully adjusted** | 1.03 | 0.7 | 1.51 | 0.88 |
| All-cause mRS | All-cause | Main* | 1.34 | 1.19 | 1.51 | $<0.0001$ |

[^4]**Fully Adjusted regression: adjusted for age, centre, PSA at baseline, stage, Gleason score

Appendix D Table D 4: Correlation between individual metabolites and PCa mRS and all-cause mRS in the ProtecT trial, cases only.


Appendix D Table D 5: Cox regression results for individual metabolites and all-cause death in the minimally adjusted model, in controls and pooled

| Metabolite | N | Hazard Ratio |  | Lower $95 \% \mathrm{CI}$ | Upper $95 \% \text { CI }$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xxl_vldl_p | 2167 |  | 0.920 | 0.796 | 1.062 | 0.254 | Controls | allcausedeath |
| xxl_vldl_1 | 2153 |  | 0.921 | 0.801 | 1.059 | 0.247 | Controls | allcausedeath |
| xxl_vldl_pl | 2153 |  | 0.920 | 0.800 | 1.057 | 0.239 | Controls | allcausedeath |
| xxl_vldl_c | 2157 |  | 0.956 | 0.833 | 1.096 | 0.515 | Controls | allcausedeath |
| xxl_vldl_ce | 2160 |  | 0.957 | 0.835 | 1.097 | 0.530 | Controls | allcausedeath |
| xxl_vldl_fc | 2154 |  | 0.926 | 0.806 | 1.064 | 0.280 | Controls | allcausedeath |
| xxl_vldl_tg | 2152 |  | 0.917 | 0.798 | 1.055 | 0.226 | Controls | allcausedeath |
| xl_vldl_p | 2155 |  | 0.912 | 0.793 | 1.050 | 0.200 | Controls | allcausedeath |
| xl_vldl_1 | 2155 |  | 0.913 | 0.794 | 1.051 | 0.205 | Controls | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower <br> 95\% CI | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_vldl_pl | 2154 |  | 0.922 | 0.802 | 1.060 | 0.254 | Controls | allcausedeath |
| xl_vldl_c | 2158 |  | 0.940 | 0.819 | 1.079 | 0.379 | Controls | allcausedeath |
| xl_vldl_ce | 2160 |  | 0.935 | 0.814 | 1.073 | 0.339 | Controls | allcausedeath |
| xl_vldl_fc | 2156 |  | 0.922 | 0.802 | 1.061 | 0.258 | Controls | allcausedeath |
| xl_vldl_tg | 2155 |  | 0.909 | 0.790 | 1.046 | 0.183 | Controls | allcausedeath |
| 1_vldl_p | 2162 |  | 0.921 | 0.801 | 1.059 | 0.247 | Controls | allcausedeath |
| 1_vldl_1 | 2161 |  | 0.922 | 0.802 | 1.060 | 0.253 | Controls | allcausedeath |
| l_vldl_pl | 2161 |  | 0.926 | 0.805 | 1.064 | 0.276 | Controls | allcausedeath |
| 1_vldl_c | 2161 |  | 0.929 | 0.809 | 1.067 | 0.298 | Controls | allcausedeath |
| l_vldl_ce | 2162 |  | 0.931 | 0.810 | 1.070 | 0.314 | Controls | allcausedeath |
| l_vldl_fc | 2159 |  | 0.905 | 0.786 | 1.042 | 0.166 | Controls | allcausedeath |
| l_vldl_tg | 2160 |  | 0.895 | 0.777 | 1.032 | 0.127 | Controls | allcausedeath |
| m_vldl_p | 2163 |  | 0.908 | 0.789 | 1.045 | 0.179 | Controls | allcausedeath |
| m_vldl_l | 2163 |  | 0.910 | 0.791 | 1.047 | 0.189 | Controls | allcausedeath |


| Metabolite | N | Hazard Ratio | Lower $95 \% \mathrm{CI}$ | $\begin{gathered} \text { Upper } \\ 95 \% \text { CI } \end{gathered}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_vldl_pl | 2163 | 0.910 | 0.791 | 1.048 | 0.191 | Controls | allcausedeath |
| m_vldl_c | 2165 | 0.929 | 0.808 | 1.069 | 0.304 | Controls | allcausedeath |
| m_vldl_ce | 2166 | 0.948 | 0.824 | 1.090 | 0.453 | Controls | allcausedeath |
| m_vldl_fc | 2165 | 0.913 | 0.794 | 1.050 | 0.204 | Controls | allcausedeath |
| m_vldl_tg | 2163 | 0.902 | 0.783 | 1.038 | 0.150 | Controls | allcausedeath |
| s_vldl_p | 2167 | 0.908 | 0.789 | 1.046 | 0.182 | Controls | allcausedeath |
| S_vldl_l | 2167 | 0.910 | 0.790 | 1.048 | 0.190 | Controls | allcausedeath |
| s_vldl_pl | 2167 | 0.914 | 0.794 | 1.052 | 0.208 | Controls | allcausedeath |
| s_vldl_c | 2167 | 0.939 | 0.817 | 1.080 | 0.379 | Controls | allcausedeath |
| s_vldl_ce | 2167 | 0.958 | 0.834 | 1.100 | 0.541 | Controls | allcausedeath |
| s_vldl_fc | 2167 | 0.917 | 0.796 | 1.056 | 0.228 | Controls | allcausedeath |
| s_vldl_tg | 2165 | 0.903 | 0.785 | 1.039 | 0.154 | Controls | allcausedeath |
| xs_vldl_p | 2167 | 0.980 | 0.856 | 1.123 | 0.776 | Controls | allcausedeath |
| xs_vldl_1 | 2167 | $0.984$ | 0.859 | 1.127 | 0.814 | Controls | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower <br> 95\% CI | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xs_vldl_pl | 2167 |  | 0.982 | 0.857 | 1.124 | 0.788 | Controls | allcausedeath |
| xs_vldl_c | 2166 |  | 0.985 | 0.859 | 1.129 | 0.829 | Controls | allcausedeath |
| xs_vldl_ce | 2166 |  | 1.001 | 0.874 | 1.147 | 0.987 | Controls | allcausedeath |
| xs_vldl_fc | 2167 |  | 0.973 | 0.850 | 1.113 | 0.686 | Controls | allcausedeath |
| xs_vldl_tg | 2167 |  | 0.950 | 0.827 | 1.092 | 0.472 | Controls | allcausedeath |
| idl_p | 2166 |  | 0.986 | 0.860 | 1.130 | 0.837 | Controls | allcausedeath |
| idl_1 | 2166 |  | 0.985 | 0.860 | 1.129 | 0.829 | Controls | allcausedeath |
| idl_pl | 2167 |  | 0.987 | 0.862 | 1.130 | 0.849 | Controls | allcausedeath |
| idl_c | 2167 |  | 0.980 | 0.855 | 1.123 | 0.774 | Controls | allcausedeath |
| idl_ce | 2167 |  | 0.975 | 0.850 | 1.118 | 0.714 | Controls | allcausedeath |
| idl_fc | 2167 |  | 0.994 | 0.869 | 1.138 | 0.931 | Controls | allcausedeath |
| idl_tg | 2164 |  | 1.003 | 0.875 | 1.149 | 0.970 | Controls | allcausedeath |
| 1_1dl_p | 2167 |  | 0.956 | 0.834 | 1.097 | 0.521 | Controls | allcausedeath |
| 1_Idl_1 | 2167 |  | 0.955 | 0.833 | 1.095 | 0.511 | Controls | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower $95 \% \text { CI }$ | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1_ldl_pl | 2167 |  | 0.955 | 0.832 | 1.095 | 0.507 | Controls | allcausedeath |
| 1_ldl_c | 2167 |  | 0.948 | 0.826 | 1.087 | 0.445 | Controls | allcausedeath |
| 1_ldl_ce | 2167 |  | 0.940 | 0.819 | 1.079 | 0.378 | Controls | allcausedeath |
| 1_ldl_fc | 2167 |  | 0.971 | 0.848 | 1.112 | 0.670 | Controls | allcausedeath |
| 1_ldl_tg | 2165 |  | 1.012 | 0.885 | 1.158 | 0.859 | Controls | allcausedeath |
| m_ldl_p | 2167 |  | 0.918 | 0.799 | 1.054 | 0.223 | Controls | allcausedeath |
| m_ldl_1 | 2167 |  | 0.914 | 0.796 | 1.049 | 0.202 | Controls | allcausedeath |
| m_ldl_pl | 2167 |  | 0.913 | 0.794 | 1.050 | 0.202 | Controls | allcausedeath |
| m_ldl_c | 2167 |  | 0.909 | 0.792 | 1.043 | 0.173 | Controls | allcausedeath |
| m_ldl_ce | 2167 |  | 0.911 | 0.794 | 1.046 | 0.185 | Controls | allcausedeath |
| m_ldl_fc | 2167 |  | 0.897 | 0.781 | 1.031 | 0.125 | Controls | allcausedeath |
| m_ldl_tg | 2166 |  | 1.017 | 0.890 | 1.162 | 0.804 | Controls | allcausedeath |
| s_ldl_p | 2167 |  | 0.892 | 0.776 | 1.025 | 0.106 | Controls | allcausedeath |
| s_ldl_1 | 2167 |  | 0.890 | 0.774 | 1.022 | 0.100 | Controls | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower <br> 95\% CI | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_ldl_pl | 2167 |  | 0.874 | 0.760 | 1.006 | 0.060 | Controls | allcausedeath |
| s_ldl_c | 2167 |  | 0.897 | 0.781 | 1.029 | 0.120 | Controls | allcausedeath |
| s_ldl_ce | 2167 |  | 0.904 | 0.788 | 1.037 | 0.151 | Controls | allcausedeath |
| s_ldl_fc | 2167 |  | 0.864 | 0.752 | 0.992 | 0.039 | Controls | allcausedeath |
| s_ldl_tg | 2167 |  | 0.957 | 0.833 | 1.099 | 0.537 | Controls | allcausedeath |
| xl_hdl_p | 2165 |  | 1.061 | 0.931 | 1.209 | 0.375 | Controls | allcausedeath |
| xl_hdl_1 | 2165 |  | 1.054 | 0.924 | 1.202 | 0.432 | Controls | allcausedeath |
| xl_hdl_pl | 2165 |  | 1.127 | 0.993 | 1.281 | 0.065 | Controls | allcausedeath |
| xl_hdl_c | 2165 |  | 0.961 | 0.839 | 1.101 | 0.569 | Controls | allcausedeath |
| xl_hdl_ce | 2165 |  | 0.935 | 0.815 | 1.072 | 0.333 | Controls | allcausedeath |
| xl_hdl_fc | 2165 |  | 1.029 | 0.901 | 1.174 | 0.678 | Controls | allcausedeath |
| xl_hdl_tg | 2161 |  | 1.054 | 0.925 | 1.201 | 0.427 | Controls | allcausedeath |
| 1_hdl_p | 2167 |  | 1.163 | 1.024 | 1.321 | 0.020 | Controls | allcausedeath |
| 1_hdl_1 | 2167 |  | 1.159 | 1.020 | 1.316 | 0.023 | Controls | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower $95 \% \text { CI }$ | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1_hdl_pl | 2167 |  | 1.157 | 1.015 | 1.317 | 0.028 | Controls | allcausedeath |
| 1_hdl_c | 2167 |  | 1.147 | 1.010 | 1.302 | 0.035 | Controls | allcausedeath |
| 1_hdl_ce | 2167 |  | 1.151 | 1.014 | 1.306 | 0.030 | Controls | allcausedeath |
| 1_hdl_fc | 2167 |  | 1.134 | 0.998 | 1.289 | 0.053 | Controls | allcausedeath |
| 1_hdl_tg | 2163 |  | 1.240 | 1.093 | 1.405 | 0.001 | Controls | allcausedeath |
| m_hdl_p | 2167 |  | 1.089 | 0.950 | 1.248 | 0.222 | Controls | allcausedeath |
| m_hdl_1 | 2167 |  | 1.086 | 0.947 | 1.246 | 0.235 | Controls | allcausedeath |
| m_hdl_pl | 2167 |  | 1.103 | 0.963 | 1.263 | 0.156 | Controls | allcausedeath |
| m_hdl_c | 2167 |  | 1.060 | 0.923 | 1.217 | 0.411 | Controls | allcausedeath |
| m_hdl_ce | 2167 |  | 1.053 | 0.917 | 1.209 | 0.467 | Controls | allcausedeath |
| m_hdl_fc | 2167 |  | 1.083 | 0.944 | 1.243 | 0.255 | Controls | allcausedeath |
| m_hdl_tg | 2166 |  | 1.080 | 0.946 | 1.234 | 0.256 | Controls | allcausedeath |
| s_hdl_p | 2167 |  | 0.933 | 0.812 | 1.073 | 0.332 | Controls | allcausedeath |
| s_hdl_l | 2167 |  | 0.927 | 0.806 | 1.066 | 0.287 | Controls | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower 95\% CI | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_hdl_pl | 2167 |  | 0.996 | 0.870 | 1.140 | 0.948 | Controls | allcausedeath |
| s_hdl_c | 2161 |  | 0.907 | 0.795 | 1.035 | 0.149 | Controls | allcausedeath |
| s_hdl_ce | 2159 |  | 0.914 | 0.801 | 1.042 | 0.176 | Controls | allcausedeath |
| s_hdl_fc | 2167 |  | 0.930 | 0.810 | 1.068 | 0.303 | Controls | allcausedeath |
| s_hdl_tg | 2167 |  | 0.972 | 0.848 | 1.114 | 0.683 | Controls | allcausedeath |
| vldl_d | 2166 |  | 0.905 | 0.789 | 1.038 | 0.155 | Controls | allcausedeath |
| ldl_d | 2166 |  | 1.603 | 1.413 | 1.819 | 0.000 | Controls | allcausedeath |
| hdl_d | 2166 |  | 1.147 | 1.010 | 1.303 | 0.035 | Controls | allcausedeath |
| serum_c | 2166 |  | 0.954 | 0.832 | 1.093 | 0.497 | Controls | allcausedeath |
| vldl_c | 2166 |  | 0.933 | 0.811 | 1.075 | 0.338 | Controls | allcausedeath |
| remnant_c | 2166 |  | 0.945 | 0.822 | 1.086 | 0.422 | Controls | allcausedeath |
| ldl_c | 2166 |  | 0.924 | 0.805 | 1.061 | 0.261 | Controls | allcausedeath |
| hdl_c | 2166 |  | 1.064 | 0.930 | 1.217 | 0.368 | Controls | allcausedeath |
| hdl2_c | 2166 |  | 1.074 | 0.939 | 1.229 | 0.296 | Controls | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower $95 \% \text { CI }$ | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| hdl3_c | 2166 |  | 0.949 | 0.828 | 1.087 | 0.448 | Controls | allcausedeath |
| estc | 2165 |  | 0.951 | 0.829 | 1.090 | 0.467 | Controls | allcausedeath |
| freec | 2165 |  | 0.961 | 0.839 | 1.100 | 0.565 | Controls | allcausedeath |
| serum_tg | 2166 |  | 0.921 | 0.798 | 1.062 | 0.258 | Controls | allcausedeath |
| vldl_tg | 2166 |  | 0.901 | 0.780 | 1.042 | 0.159 | Controls | allcausedeath |
| ldl_tg | 2166 |  | 1.025 | 0.897 | 1.171 | 0.719 | Controls | allcausedeath |
| hdl_tg | 2166 |  | 1.093 | 0.958 | 1.246 | 0.185 | Controls | allcausedeath |
| totpg | 2165 |  | 0.999 | 0.872 | 1.143 | 0.983 | Controls | allcausedeath |
| tg_pg | 2164 |  | 0.895 | 0.776 | 1.031 | 0.124 | Controls | allcausedeath |
| pc | 2165 |  | 1.016 | 0.887 | 1.163 | 0.822 | Controls | allcausedeath |
| sm | 2165 |  | 0.977 | 0.853 | 1.117 | 0.730 | Controls | allcausedeath |
| totcho | 2165 |  | 1.044 | 0.913 | 1.193 | 0.532 | Controls | allcausedeath |
| apoa1 | 2166 |  | 1.027 | 0.897 | 1.177 | 0.697 | Controls | allcausedeath |
| apob | 2166 |  | 0.909 | 0.789 | 1.046 | 0.184 | Controls | allcausedeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | $\begin{gathered} \text { Upper } \\ 95 \% \text { CI } \end{gathered}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| apob_apoa1 | 2166 | 0.905 | 0.787 | 1.041 | 0.162 | Controls | allcausedeath |
| totfa | 2162 | 0.921 | 0.800 | 1.061 | 0.254 | Controls | allcausedeath |
| unsat | 2162 | 0.931 | 0.814 | 1.064 | 0.295 | Controls | allcausedeath |
| dha | 2162 | 0.930 | 0.808 | 1.070 | 0.311 | Controls | allcausedeath |
| la | 2162 | 0.883 | 0.766 | 1.017 | 0.083 | Controls | allcausedeath |
| faw3 | 2162 | 0.865 | 0.748 | 1.001 | 0.052 | Controls | allcausedeath |
| faw6 | 2162 | 0.895 | 0.778 | 1.030 | 0.121 | Controls | allcausedeath |
| pufa | 2162 | 0.883 | 0.766 | 1.017 | 0.084 | Controls | allcausedeath |
| mufa | 2162 | 0.955 | 0.830 | 1.099 | 0.519 | Controls | allcausedeath |
| sfa | 2162 | 0.934 | 0.811 | 1.075 | 0.339 | Controls | allcausedeath |
| dha_fa | 2162 | 0.970 | 0.848 | 1.108 | 0.650 | Controls | allcausedeath |
| la_fa | 2162 | 0.927 | 0.811 | 1.060 | 0.268 | Controls | allcausedeath |
| faw3_fa | 2162 | 0.868 | 0.751 | 1.004 | 0.056 | Controls | allcausedeath |
| faw6_fa | 2162 | 0.957 | 0.837 | 1.094 | 0.519 | Controls | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower <br> 95\% CI | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| pufa_fa | 2162 |  | 0.921 | 0.806 | 1.052 | 0.223 | Controls | allcausedeath |
| mufa_fa | 2162 |  | 1.074 | 0.937 | 1.230 | 0.305 | Controls | allcausedeath |
| sfa_fa | 2162 |  | 1.065 | 0.927 | 1.224 | 0.375 | Controls | allcausedeath |
| glc | 2158 |  | 1.110 | 1.001 | 1.231 | 0.048 | Controls | allcausedeath |
| lac | 2165 |  | 1.127 | 1.022 | 1.243 | 0.016 | Controls | allcausedeath |
| pyr | 2163 |  | 1.073 | 0.969 | 1.188 | 0.175 | Controls | allcausedeath |
| cit | 2164 |  | 0.978 | 0.850 | 1.126 | 0.755 | Controls | allcausedeath |
| ala | 2165 |  | 0.982 | 0.856 | 1.126 | 0.793 | Controls | allcausedeath |
| gln | 2164 |  | 0.909 | 0.792 | 1.043 | 0.173 | Controls | allcausedeath |
| gly | 2158 |  | 1.103 | 0.971 | 1.253 | 0.133 | Controls | allcausedeath |
| his | 2158 |  | 1.026 | 0.903 | 1.166 | 0.690 | Controls | allcausedeath |
| ile | 2165 |  | 0.969 | 0.844 | 1.112 | 0.651 | Controls | allcausedeath |
| leu | 2165 |  | 0.937 | 0.815 | 1.076 | 0.357 | Controls | allcausedeath |
| val | 2164 |  | 0.956 | 0.832 | 1.099 | 0.529 | Controls | allcausedeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | $\begin{gathered} \text { Upper } \\ 95 \% \text { CI } \end{gathered}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| phe | 2165 | 1.219 | 1.076 | 1.381 | 0.002 | Controls | allcausedeath |
| tyr | 2158 | 1.017 | 0.891 | 1.161 | 0.799 | Controls | allcausedeath |
| ace | 2165 | 1.085 | 1.019 | 1.155 | 0.010 | Controls | allcausedeath |
| acace | 2165 | 1.026 | 0.898 | 1.172 | 0.709 | Controls | allcausedeath |
| bohbut | 2095 | 0.965 | 0.827 | 1.125 | 0.647 | Controls | allcausedeath |
| crea | 2158 | 1.003 | 0.873 | 1.153 | 0.967 | Controls | allcausedeath |
| alb | 2166 | 0.963 | 0.838 | 1.107 | 0.596 | Controls | allcausedeath |
| gp | 2165 | 1.056 | 0.923 | 1.207 | 0.427 | Controls | allcausedeath |
| xxl_vldl_p | 2093 | 1.178 | 1.037 | 1.337 | 0.012 | Cases | allcausedeath |
| xxl_vldl_1 | 2085 | 1.187 | 1.043 | 1.350 | 0.009 | Cases | allcausedeath |
| xxl_vldl_pl | 2085 | 1.193 | 1.049 | 1.356 | 0.007 | Cases | allcausedeath |
| xxl_vldl_c | 2087 | 1.143 | 1.003 | 1.303 | 0.045 | Cases | allcausedeath |
| xxl_vldl_ce | 2088 | 1.088 | 0.954 | 1.240 | 0.208 | Cases | allcausedeath |
| xxl_vldl_fc | 2086 | 1.193 | 1.050 | 1.356 | 0.007 | Cases | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower $95 \% \text { CI }$ | $\begin{gathered} \text { Upper } \\ 95 \% \text { CI } \end{gathered}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xxl_vldl_tg | 2086 |  | 1.186 | 1.044 | 1.347 | 0.009 | Cases | allcausedeath |
| xl_vldl_p | 2089 |  | 1.141 | 1.004 | 1.296 | 0.043 | Cases | allcausedeath |
| xl_vldl_1 | 2088 |  | 1.150 | 1.011 | 1.309 | 0.033 | Cases | allcausedeath |
| xl_vldl_pl | 2087 |  | 1.157 | 1.017 | 1.316 | 0.026 | Cases | allcausedeath |
| xl_vldl_c | 2089 |  | 1.114 | 0.978 | 1.268 | 0.104 | Cases | allcausedeath |
| xl_vldl_ce | 2088 |  | 1.094 | 0.959 | 1.248 | 0.183 | Cases | allcausedeath |
| xl_vldl_fc | 2088 |  | 1.147 | 1.009 | 1.305 | 0.036 | Cases | allcausedeath |
| xl_vldl_tg | 2088 |  | 1.157 | 1.017 | 1.317 | 0.026 | Cases | allcausedeath |
| 1_vldl_p | 2089 |  | 1.121 | 0.982 | 1.280 | 0.091 | Cases | allcausedeath |
| 1_vldl_1 | 2089 |  | 1.120 | 0.981 | 1.278 | 0.094 | Cases | allcausedeath |
| l_vldl_pl | 2089 |  | 1.120 | 0.982 | 1.277 | 0.092 | Cases | allcausedeath |
| 1_vldl_c | 2088 |  | 1.098 | 0.961 | 1.254 | 0.167 | Cases | allcausedeath |
| 1_vldl_ce | 2088 |  | 1.060 | 0.926 | 1.212 | 0.399 | Cases | allcausedeath |
| 1_vldl_fc | 2089 |  | 1.129 | 0.991 | 1.286 | 0.067 | Cases | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower 95\% CI | $\begin{gathered} \text { Upper } \\ 95 \% \text { CI } \end{gathered}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| l_vldl_tg | 2089 |  | 1.127 | 0.989 | 1.286 | 0.074 | Cases | allcausedeath |
| m_vldl_p | 2088 |  | 1.063 | 0.929 | 1.218 | 0.375 | Cases | allcausedeath |
| m_vldl_1 | 2088 |  | 1.061 | 0.926 | 1.215 | 0.393 | Cases | allcausedeath |
| m_vldl_pl | 2089 |  | 1.053 | 0.920 | 1.204 | 0.457 | Cases | allcausedeath |
| m_vldl_c | 2089 |  | 1.026 | 0.896 | 1.175 | 0.714 | Cases | allcausedeath |
| m_vldl_ce | 2091 |  | 0.986 | 0.861 | 1.131 | 0.845 | Cases | allcausedeath |
| m_vldl_fc | 2088 |  | 1.069 | 0.934 | 1.225 | 0.331 | Cases | allcausedeath |
| m_vldl_tg | 2089 |  | 1.075 | 0.939 | 1.230 | 0.296 | Cases | allcausedeath |
| s_vldl_p | 2091 |  | 0.976 | 0.852 | 1.118 | 0.728 | Cases | allcausedeath |
| s_vldl_1 | 2092 |  | 0.990 | 0.865 | 1.132 | 0.882 | Cases | allcausedeath |
| s_vldl_pl | 2092 |  | 1.002 | 0.877 | 1.145 | 0.975 | Cases | allcausedeath |
| s_vldl_c | 2092 |  | 0.917 | 0.802 | 1.048 | 0.202 | Cases | allcausedeath |
| s_vldl_ce | 2092 |  | 0.882 | 0.771 | 1.009 | 0.067 | Cases | allcausedeath |
| s_vldl_fc | 2092 |  | 0.985 | 0.863 | 1.126 | 0.828 | Cases | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower <br> 95\% CI | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_vldl_tg | 2089 |  | 1.014 | 0.886 | 1.161 | 0.839 | Cases | allcausedeath |
| xs_vldl_p | 2092 |  | 0.945 | 0.827 | 1.080 | 0.409 | Cases | allcausedeath |
| xs_vldl_1 | 2092 |  | 0.936 | 0.818 | 1.070 | 0.331 | Cases | allcausedeath |
| xs_vldl_pl | 2093 |  | 0.890 | 0.777 | 1.020 | 0.093 | Cases | allcausedeath |
| xs_vldl_c | 2093 |  | 0.934 | 0.816 | 1.067 | 0.315 | Cases | allcausedeath |
| xs_vldl_ce | 2093 |  | 0.969 | 0.848 | 1.107 | 0.643 | Cases | allcausedeath |
| xs_vldl_fc | 2092 |  | 0.864 | 0.753 | 0.990 | 0.036 | Cases | allcausedeath |
| xs_vldl_tg | 2092 |  | 1.020 | 0.895 | 1.163 | 0.767 | Cases | allcausedeath |
| idl_p | 2093 |  | 0.924 | 0.808 | 1.057 | 0.249 | Cases | allcausedeath |
| idl_1 | 2093 |  | 0.916 | 0.801 | 1.048 | 0.202 | Cases | allcausedeath |
| idl_pl | 2093 |  | 0.887 | 0.775 | 1.016 | 0.084 | Cases | allcausedeath |
| idl_c | 2093 |  | 0.914 | 0.799 | 1.045 | 0.190 | Cases | allcausedeath |
| idl_ce | 2093 |  | 0.935 | 0.818 | 1.069 | 0.326 | Cases | allcausedeath |
| idl_fc | 2093 |  | 0.869 | 0.759 | 0.995 | 0.042 | Cases | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower 95\% CI | $\begin{gathered} \text { Upper } \\ 95 \% \text { CI } \end{gathered}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| idl_tg | 2092 |  | 1.049 | 0.924 | 1.192 | 0.460 | Cases | allcausedeath |
| 1_1dl_p | 2093 |  | 0.891 | 0.779 | 1.020 | 0.096 | Cases | allcausedeath |
| 1_ldl_1 | 2093 |  | 0.885 | 0.773 | 1.013 | 0.076 | Cases | allcausedeath |
| 1_ldl_pl | 2093 |  | 0.892 | 0.779 | 1.021 | 0.096 | Cases | allcausedeath |
| 1_ldl_c | 2093 |  | 0.871 | 0.761 | 0.997 | 0.045 | Cases | allcausedeath |
| 1_ldl_ce | 2093 |  | 0.874 | 0.763 | 1.000 | 0.051 | Cases | allcausedeath |
| 1_ldl_fc | 2093 |  | 0.864 | 0.755 | 0.990 | 0.035 | Cases | allcausedeath |
| 1_ldl_tg | 2092 |  | 1.060 | 0.933 | 1.204 | 0.373 | Cases | allcausedeath |
| m_ldl_p | 2093 |  | 0.860 | 0.751 | 0.985 | 0.029 | Cases | allcausedeath |
| m_ldl_1 | 2093 |  | 0.855 | 0.746 | 0.979 | 0.023 | Cases | allcausedeath |
| m_ldl_pl | 2093 |  | 0.922 | 0.807 | 1.054 | 0.235 | Cases | allcausedeath |
| m_ldl_c | 2093 |  | 0.831 | 0.725 | 0.952 | 0.008 | Cases | allcausedeath |
| m_ldl_ce | 2093 |  | 0.824 | 0.719 | 0.943 | 0.005 | Cases | allcausedeath |
| m_ldl_fc | 2093 |  | 0.866 | 0.757 | 0.991 | 0.037 | Cases | allcausedeath |


| Metabolite | N | Hazard Ratio | Lower $95 \% \text { CI }$ | $\begin{gathered} \text { Upper } \\ 95 \% \text { CI } \end{gathered}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_ldl_tg | 2092 | 1.037 | 0.911 | 1.181 | 0.579 | Cases | allcausedeath |
| s_ldl_p | 2093 | 0.858 | 0.749 | 0.983 | 0.027 | Cases | allcausedeath |
| s_ldl_1 | 2093 | 0.853 | 0.745 | 0.977 | 0.022 | Cases | allcausedeath |
| s_ldl_pl | 2093 | 0.936 | 0.819 | 1.069 | 0.328 | Cases | allcausedeath |
| s_ldl_c | 2093 | 0.823 | 0.718 | 0.942 | 0.005 | Cases | allcausedeath |
| s_ldl_ce | 2093 | 0.816 | 0.712 | 0.935 | 0.003 | Cases | allcausedeath |
| s_ldl_fc | 2093 | 0.857 | 0.749 | 0.981 | 0.025 | Cases | allcausedeath |
| s_ldl_tg | 2092 | 1.056 | 0.927 | 1.203 | 0.410 | Cases | allcausedeath |
| xl_hdl_p | 2091 | 1.104 | 0.972 | 1.254 | 0.129 | Cases | allcausedeath |
| xl_hdl_l | 2091 | 1.103 | 0.971 | 1.254 | 0.131 | Cases | allcausedeath |
| xl_hdl_pl | 2091 | 1.090 | 0.958 | 1.239 | 0.190 | Cases | allcausedeath |
| xl_hdl_c | 2091 | 1.098 | 0.964 | 1.249 | 0.159 | Cases | allcausedeath |
| xl_hdl_ce | 2091 | 1.082 | 0.950 | 1.232 | 0.235 | Cases | allcausedeath |
| xl_hdl_fc | 2091 | 1.134 | 0.998 | 1.288 | 0.053 | Cases | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower <br> 95\% CI | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_hdl_tg | 2088 |  | 1.183 | 1.042 | 1.343 | 0.009 | Cases | allcausedeath |
| 1_hdl_p | 2092 |  | 1.132 | 0.996 | 1.287 | 0.058 | Cases | allcausedeath |
| 1_hdl_1 | 2092 |  | 1.129 | 0.993 | 1.283 | 0.064 | Cases | allcausedeath |
| 1_hdl_pl | 2093 |  | 1.115 | 0.980 | 1.269 | 0.097 | Cases | allcausedeath |
| 1_hdl_c | 2092 |  | 1.127 | 0.992 | 1.280 | 0.067 | Cases | allcausedeath |
| 1_hdl_ce | 2092 |  | 1.136 | 1.000 | 1.290 | 0.049 | Cases | allcausedeath |
| 1_hdl_fc | 2092 |  | 1.096 | 0.962 | 1.248 | 0.167 | Cases | allcausedeath |
| 1_hdl_tg | 2091 |  | 1.136 | 1.004 | 1.287 | 0.044 | Cases | allcausedeath |
| m_hdl_p | 2093 |  | 1.171 | 1.030 | 1.331 | 0.016 | Cases | allcausedeath |
| m_hdl_1 | 2093 |  | 1.162 | 1.022 | 1.322 | 0.022 | Cases | allcausedeath |
| m_hdl_pl | 2092 |  | 1.172 | 1.030 | 1.333 | 0.016 | Cases | allcausedeath |
| m_hdl_c | 2093 |  | 1.116 | 0.979 | 1.272 | 0.100 | Cases | allcausedeath |
| m_hdl_ce | 2093 |  | 1.118 | 0.981 | 1.274 | 0.094 | Cases | allcausedeath |
| m_hdl_fc | 2093 |  | 1.108 | 0.971 | 1.265 | 0.127 | Cases | allcausedeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | $\begin{gathered} \text { Upper } \\ 95 \% \text { CI } \end{gathered}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_hdl_tg | 2092 | 1.156 | 1.013 | 1.319 | 0.032 | Cases | allcausedeath |
| s_hdl_p | 2093 | 1.064 | 0.934 | 1.213 | 0.348 | Cases | allcausedeath |
| s_hdl_l | 2093 | 1.046 | 0.918 | 1.193 | 0.500 | Cases | allcausedeath |
| s_hdl_pl | 2093 | 1.237 | 1.090 | 1.404 | 0.001 | Cases | allcausedeath |
| s_hdl_c | 2087 | 0.817 | 0.715 | 0.934 | 0.003 | Cases | allcausedeath |
| s_hdl_ce | 2087 | 0.795 | 0.698 | 0.905 | 0.001 | Cases | allcausedeath |
| s_hdl_fc | 2093 | 1.142 | 1.004 | 1.300 | 0.043 | Cases | allcausedeath |
| s_hdl_tg | 2091 | 1.095 | 0.960 | 1.249 | 0.177 | Cases | allcausedeath |
| vldl_d | 2093 | 1.153 | 1.007 | 1.321 | 0.040 | Cases | allcausedeath |
| ldl_d | 2093 | 1.260 | 1.111 | 1.428 | 0.000 | Cases | allcausedeath |
| hdl_d | 2093 | 1.092 | 0.960 | 1.244 | 0.181 | Cases | allcausedeath |
| serum_c | 2093 | 0.922 | 0.806 | 1.055 | 0.239 | Cases | allcausedeath |
| vldl_c | 2093 | 1.014 | 0.886 | 1.160 | 0.842 | Cases | allcausedeath |
| remnant_c | 2093 | 0.968 | 0.847 | 1.106 | 0.628 | Cases | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower $95 \% \text { CI }$ | $\begin{gathered} \text { Upper } \\ 95 \% \text { CI } \end{gathered}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| ldl_c | 2093 |  | 0.848 | 0.740 | 0.971 | 0.017 | Cases | allcausedeath |
| hdl_c | 2093 |  | 1.080 | 0.947 | 1.231 | 0.252 | Cases | allcausedeath |
| hdl2_c | 2093 |  | 1.086 | 0.953 | 1.237 | 0.217 | Cases | allcausedeath |
| hdl3_c | 2093 |  | 0.987 | 0.864 | 1.127 | 0.849 | Cases | allcausedeath |
| estc | 2090 |  | 0.918 | 0.802 | 1.050 | 0.214 | Cases | allcausedeath |
| freec | 2090 |  | 0.933 | 0.815 | 1.068 | 0.313 | Cases | allcausedeath |
| serum_tg | 2093 |  | 1.112 | 0.975 | 1.270 | 0.114 | Cases | allcausedeath |
| vldl_tg | 2093 |  | 1.109 | 0.970 | 1.267 | 0.130 | Cases | allcausedeath |
| ldl_tg | 2093 |  | 1.055 | 0.928 | 1.200 | 0.410 | Cases | allcausedeath |
| hdl_tg | 2093 |  | 1.182 | 1.042 | 1.341 | 0.009 | Cases | allcausedeath |
| totpg | 2090 |  | 1.080 | 0.947 | 1.233 | 0.252 | Cases | allcausedeath |
| tg_pg | 2090 |  | 1.097 | 0.958 | 1.257 | 0.182 | Cases | allcausedeath |
| pc | 2090 |  | 1.063 | 0.932 | 1.213 | 0.361 | Cases | allcausedeath |
| sm | 2090 |  | 0.976 | 0.852 | 1.117 | 0.721 | Cases | allcausedeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | Upper $95 \% \text { CI }$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| totcho | 2090 | 1.066 | 0.934 | 1.217 | 0.344 | Cases | allcausedeath |
| apoa1 | 2093 | 1.099 | 0.965 | 1.252 | 0.155 | Cases | allcausedeath |
| apob | 2092 | 0.958 | 0.838 | 1.096 | 0.532 | Cases | allcausedeath |
| apob_apoa1 | 2092 | 0.930 | 0.814 | 1.063 | 0.288 | Cases | allcausedeath |
| totfa | 2089 | 1.057 | 0.926 | 1.208 | 0.411 | Cases | allcausedeath |
| unsat | 2089 | 0.771 | 0.678 | 0.877 | 0.000 | Cases | allcausedeath |
| dha | 2089 | 0.890 | 0.771 | 1.027 | 0.110 | Cases | allcausedeath |
| la | 2089 | 0.917 | 0.803 | 1.048 | 0.203 | Cases | allcausedeath |
| faw3 | 2089 | 0.885 | 0.768 | 1.020 | 0.092 | Cases | allcausedeath |
| faw6 | 2089 | 0.933 | 0.816 | 1.066 | 0.307 | Cases | allcausedeath |
| pufa | 2089 | 0.919 | 0.803 | 1.052 | 0.220 | Cases | allcausedeath |
| mufa | 2089 | 1.124 | 0.988 | 1.278 | 0.076 | Cases | allcausedeath |
| sfa | 2089 | 1.102 | 0.966 | 1.256 | 0.149 | Cases | allcausedeath |
| dha_fa | 2089 | 0.826 | 0.713 | 0.956 | 0.011 | Cases | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower <br> 95\% CI | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| la_fa | 2089 |  | 0.792 | 0.704 | 0.892 | 0.000 | Cases | allcausedeath |
| faw3_fa | 2089 |  | 0.787 | 0.677 | 0.913 | 0.002 | Cases | allcausedeath |
| faw6_fa | 2089 |  | 0.785 | 0.693 | 0.890 | 0.000 | Cases | allcausedeath |
| pufa_fa | 2089 |  | 0.759 | 0.671 | 0.858 | 0.000 | Cases | allcausedeath |
| mufa_fa | 2089 |  | 1.239 | 1.096 | 1.400 | 0.001 | Cases | allcausedeath |
| sfa_fa | 2089 |  | 1.253 | 1.107 | 1.419 | 0.000 | Cases | allcausedeath |
| glc | 2087 |  | 1.198 | 1.082 | 1.328 | 0.001 | Cases | allcausedeath |
| lac | 2093 |  | 1.025 | 0.916 | 1.146 | 0.669 | Cases | allcausedeath |
| pyr | 2091 |  | 0.998 | 0.909 | 1.095 | 0.964 | Cases | allcausedeath |
| cit | 2092 |  | 1.049 | 0.921 | 1.195 | 0.467 | Cases | allcausedeath |
| ala | 2093 |  | 0.991 | 0.874 | 1.124 | 0.887 | Cases | allcausedeath |
| gln | 2092 |  | 0.890 | 0.782 | 1.014 | 0.079 | Cases | allcausedeath |
| gly | 2084 |  | 0.944 | 0.827 | 1.077 | 0.392 | Cases | allcausedeath |
| his | 2088 |  | 0.997 | 0.876 | 1.134 | 0.960 | Cases | allcausedeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | $\begin{gathered} \text { Upper } \\ 95 \% \text { CI } \end{gathered}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| ile | 2092 | 1.062 | 0.929 | 1.214 | 0.377 | Cases | allcausedeath |
| leu | 2093 | 1.060 | 0.929 | 1.209 | 0.387 | Cases | allcausedeath |
| val | 2092 | 0.956 | 0.836 | 1.093 | 0.507 | Cases | allcausedeath |
| phe | 2093 | 1.166 | 1.035 | 1.314 | 0.012 | Cases | allcausedeath |
| tyr | 2088 | 1.060 | 0.932 | 1.205 | 0.373 | Cases | allcausedeath |
| ace | 2093 | 1.140 | 1.069 | 1.217 | 0.000 | Cases | allcausedeath |
| acace | 2093 | 1.160 | 1.052 | 1.279 | 0.003 | Cases | allcausedeath |
| bohbut | 2041 | 1.081 | 0.975 | 1.199 | 0.140 | Cases | allcausedeath |
| crea | 2087 | 0.901 | 0.787 | 1.032 | 0.131 | Cases | allcausedeath |
| alb | 2093 | 1.029 | 0.903 | 1.174 | 0.665 | Cases | allcausedeath |
| gp | 2093 | 1.149 | 1.008 | 1.310 | 0.038 | Cases | allcausedeath |
| xxl_vldl_p | 4260 | 1.036 | 0.943 | 1.139 | 0.461 | Pooled | allcausedeath |
| xxl_vldl_1 | 4238 | 1.039 | 0.946 | 1.141 | 0.425 | Pooled | allcausedeath |
| xxl_vldl_pl | 4238 | 1.041 | 0.948 | 1.143 | 0.400 | Pooled | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower <br> 95\% CI | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xxl_vldl_c | 4244 |  | 1.040 | 0.947 | 1.142 | 0.415 | Pooled | allcausedeath |
| xxl_vldl_ce | 4248 |  | 1.017 | 0.926 | 1.118 | 0.720 | Pooled | allcausedeath |
| xxl_vldl_fc | 4240 |  | 1.045 | 0.952 | 1.147 | 0.352 | Pooled | allcausedeath |
| xxl_vldl_tg | 4238 |  | 1.038 | 0.945 | 1.139 | 0.436 | Pooled | allcausedeath |
| xl_vldl_p | 4244 |  | 1.018 | 0.927 | 1.118 | 0.712 | Pooled | allcausedeath |
| xl_vldl_1 | 4243 |  | 1.021 | 0.929 | 1.122 | 0.667 | Pooled | allcausedeath |
| xl_vldl_pl | 4241 |  | 1.029 | 0.936 | 1.130 | 0.555 | Pooled | allcausedeath |
| xl_vldl_c | 4247 |  | 1.020 | 0.928 | 1.120 | 0.686 | Pooled | allcausedeath |
| xl_vldl_ce | 4248 |  | 1.007 | 0.916 | 1.107 | 0.887 | Pooled | allcausedeath |
| xl_vldl_fc | 4244 |  | 1.026 | 0.933 | 1.127 | 0.599 | Pooled | allcausedeath |
| xl_vldl_tg | 4243 |  | 1.021 | 0.929 | 1.122 | 0.662 | Pooled | allcausedeath |
| 1_vldl_p | 4251 |  | 1.010 | 0.919 | 1.111 | 0.831 | Pooled | allcausedeath |
| 1_vldl_1 | 4250 |  | 1.011 | 0.919 | 1.112 | 0.825 | Pooled | allcausedeath |
| 1_vldl_pl | 4250 |  | 1.013 | 0.921 | 1.114 | 0.789 | Pooled | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower <br> 95\% CI | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| l_vldl_c | 4249 |  | 1.005 | 0.914 | 1.106 | 0.911 | Pooled | allcausedeath |
| 1_vldl_ce | 4250 |  | 0.990 | 0.899 | 1.090 | 0.834 | Pooled | allcausedeath |
| l_vldl_fc | 4248 |  | 1.007 | 0.916 | 1.108 | 0.880 | Pooled | allcausedeath |
| 1_vldl_tg | 4249 |  | 1.001 | 0.909 | 1.101 | 0.991 | Pooled | allcausedeath |
| m_vldl_p | 4251 |  | 0.979 | 0.888 | 1.078 | 0.662 | Pooled | allcausedeath |
| m_vldl_1 | 4251 |  | 0.979 | 0.889 | 1.078 | 0.666 | Pooled | allcausedeath |
| m_vldl_pl | 4252 |  | 0.976 | 0.886 | 1.075 | 0.625 | Pooled | allcausedeath |
| m_vldl_c | 4254 |  | 0.974 | 0.884 | 1.073 | 0.597 | Pooled | allcausedeath |
| m_vldl_ce | 4257 |  | 0.966 | 0.877 | 1.065 | 0.488 | Pooled | allcausedeath |
| m_vldl_fc | 4253 |  | 0.984 | 0.893 | 1.084 | 0.741 | Pooled | allcausedeath |
| m_vldl_tg | 4252 |  | 0.980 | 0.889 | 1.080 | 0.681 | Pooled | allcausedeath |
| s_vldl_p | 4258 |  | 0.941 | 0.854 | 1.037 | 0.221 | Pooled | allcausedeath |
| s_vldl_1 | 4259 |  | 0.949 | 0.861 | 1.046 | 0.291 | Pooled | allcausedeath |
| s_vldl_pl | 4259 |  | 0.957 | 0.869 | 1.054 | 0.375 | Pooled | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower <br> 95\% CI | $\begin{aligned} & \hline \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_vldl_c | 4259 |  | 0.929 | 0.843 | 1.023 | 0.133 | Pooled | allcausedeath |
| s_vldl_ce | 4259 |  | 0.920 | 0.836 | 1.013 | 0.091 | Pooled | allcausedeath |
| s_vldl_fc | 4259 |  | 0.951 | 0.864 | 1.048 | 0.309 | Pooled | allcausedeath |
| s_vldl_tg | 4254 |  | 0.955 | 0.867 | 1.052 | 0.350 | Pooled | allcausedeath |
| xs_vldl_p | 4259 |  | 0.965 | 0.877 | 1.061 | 0.464 | Pooled | allcausedeath |
| xs_vldl_1 | 4259 |  | 0.962 | 0.875 | 1.058 | 0.427 | Pooled | allcausedeath |
| xs_vldl_pl | 4260 |  | 0.939 | 0.854 | 1.034 | 0.200 | Pooled | allcausedeath |
| xs_vldl_c | 4259 |  | 0.963 | 0.875 | 1.059 | 0.436 | Pooled | allcausedeath |
| xs_vldl_ce | 4259 |  | 0.989 | 0.899 | 1.087 | 0.813 | Pooled | allcausedeath |
| xs_vldl_fc | 4259 |  | 0.921 | 0.837 | 1.014 | 0.093 | Pooled | allcausedeath |
| xs_vldl_tg | 4259 |  | 0.985 | 0.895 | 1.083 | 0.753 | Pooled | allcausedeath |
| idl_p | 4259 |  | 0.959 | 0.872 | 1.054 | 0.385 | Pooled | allcausedeath |
| idl_1 | 4259 |  | 0.954 | 0.868 | 1.050 | 0.338 | Pooled | allcausedeath |
| idl_pl | 4260 |  | 0.941 | 0.856 | 1.035 | 0.213 | Pooled | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower $95 \% \text { CI }$ | $\begin{gathered} \text { Upper } \\ 95 \% \text { CI } \end{gathered}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| idl_c | 4260 |  | 0.951 | 0.865 | 1.046 | 0.301 | Pooled | allcausedeath |
| idl_ce | 4260 |  | 0.959 | 0.872 | 1.055 | 0.387 | Pooled | allcausedeath |
| idl_fc | 4260 |  | 0.935 | 0.851 | 1.029 | 0.169 | Pooled | allcausedeath |
| idl_tg | 4256 |  | 1.029 | 0.937 | 1.129 | 0.551 | Pooled | allcausedeath |
| 1_ldl_p | 4260 |  | 0.928 | 0.843 | 1.021 | 0.124 | Pooled | allcausedeath |
| 1_ldl_1 | 4260 |  | 0.924 | 0.839 | 1.017 | 0.105 | Pooled | allcausedeath |
| 1_ldl_pl | 4260 |  | 0.927 | 0.843 | 1.021 | 0.122 | Pooled | allcausedeath |
| 1_ldı_c | 4260 |  | 0.913 | 0.830 | 1.005 | 0.063 | Pooled | allcausedeath |
| 1_ldl_ce | 4260 |  | 0.911 | 0.827 | 1.002 | 0.056 | Pooled | allcausedeath |
| 1_ldl_fc | 4260 |  | 0.922 | 0.838 | 1.014 | 0.094 | Pooled | allcausedeath |
| 1_ldl_tg | 4257 |  | 1.039 | 0.948 | 1.140 | 0.410 | Pooled | allcausedeath |
| m_ldl_p | 4260 |  | 0.893 | 0.810 | 0.983 | 0.021 | Pooled | allcausedeath |
| m_ldl_1 | 4260 |  | 0.888 | 0.806 | 0.978 | 0.016 | Pooled | allcausedeath |
| m_ldl_pl | 4260 |  | 0.922 | 0.837 | 1.015 | 0.096 | Pooled | allcausedeath |


| Metabolite | N | Hazard Ratio | Lower $95 \% \mathrm{CI}$ | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_ldl_c | 4260 | 0.873 | 0.793 | 0.962 | 0.006 | Pooled | allcausedeath |
| m_ldl_ce | 4260 | 0.871 | 0.791 | 0.959 | 0.005 | Pooled | allcausedeath |
| m_ldl_fc | 4260 | 0.886 | 0.805 | 0.976 | 0.014 | Pooled | allcausedeath |
| m_ldl_tg | 4258 | 1.030 | 0.939 | 1.130 | 0.532 | Pooled | allcausedeath |
| s_ldl_p | 4260 | 0.879 | 0.798 | 0.968 | 0.009 | Pooled | allcausedeath |
| s_ldl_l | 4260 | 0.876 | 0.795 | 0.965 | 0.007 | Pooled | allcausedeath |
| s_ldl_pl | 4260 | 0.909 | 0.826 | 1.001 | 0.052 | Pooled | allcausedeath |
| s_ldl_c | 4260 | 0.863 | 0.784 | 0.951 | 0.003 | Pooled | allcausedeath |
| s_ldl_ce | 4260 | 0.864 | 0.785 | 0.951 | 0.003 | Pooled | allcausedeath |
| s_ldl_fc | 4260 | 0.865 | 0.786 | 0.953 | 0.003 | Pooled | allcausedeath |
| s_ldl_tg | 4259 | 1.007 | 0.916 | 1.107 | 0.885 | Pooled | allcausedeath |
| xl_hdl_p | 4256 | 1.083 | 0.989 | 1.186 | 0.084 | Pooled | allcausedeath |
| xl_hdl_1 | 4256 | 1.080 | 0.986 | 1.183 | 0.099 | Pooled | allcausedeath |
| xl_hdl_pl | 4256 | $1.110$ | 1.014 | 1.215 | 0.023 | Pooled | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower <br> 95\% CI | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_hdl_c | 4256 |  | 1.029 | 0.937 | 1.130 | 0.547 | Pooled | allcausedeath |
| xl_hdl_ce | 4256 |  | 1.008 | 0.917 | 1.107 | 0.874 | Pooled | allcausedeath |
| xl_hdl_fc | 4256 |  | 1.081 | 0.987 | 1.185 | 0.093 | Pooled | allcausedeath |
| xl_hdl_tg | 4249 |  | 1.113 | 1.017 | 1.218 | 0.021 | Pooled | allcausedeath |
| 1_hdl_p | 4259 |  | 1.150 | 1.051 | 1.258 | 0.002 | Pooled | allcausedeath |
| 1_hdl_1 | 4259 |  | 1.146 | 1.048 | 1.254 | 0.003 | Pooled | allcausedeath |
| 1_hdl_pl | 4260 |  | 1.139 | 1.040 | 1.248 | 0.005 | Pooled | allcausedeath |
| 1_hdl_c | 4259 |  | 1.139 | 1.041 | 1.246 | 0.005 | Pooled | allcausedeath |
| 1_hdl_ce | 4259 |  | 1.145 | 1.047 | 1.252 | 0.003 | Pooled | allcausedeath |
| 1_hdl_fc | 4259 |  | 1.117 | 1.020 | 1.223 | 0.017 | Pooled | allcausedeath |
| 1_hdl_tg | 4254 |  | 1.188 | 1.088 | 1.297 | 0.000 | Pooled | allcausedeath |
| m_hdl_p | 4260 |  | 1.135 | 1.034 | 1.247 | 0.008 | Pooled | allcausedeath |
| m_hdl_l | 4260 |  | 1.130 | 1.029 | 1.241 | 0.011 | Pooled | allcausedeath |
| m_hdl_pl | 4259 |  | 1.142 | 1.040 | 1.254 | 0.005 | Pooled | allcausedeath |


| Metabolite | N | Hazard Ratio | Lower $95 \% \mathrm{CI}$ | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_hdl_c | 4260 | 1.093 | 0.994 | 1.202 | 0.066 | Pooled | allcausedeath |
| m_hdl_ce | 4260 | 1.090 | 0.992 | 1.199 | 0.074 | Pooled | allcausedeath |
| m_hdl_fc | 4260 | 1.101 | 1.001 | 1.211 | 0.047 | Pooled | allcausedeath |
| m_hdl_tg | 4258 | 1.117 | 1.017 | 1.227 | 0.020 | Pooled | allcausedeath |
| s_hdl_p | 4260 | 1.004 | 0.913 | 1.105 | 0.927 | Pooled | allcausedeath |
| s_hdl_l | 4260 | 0.992 | 0.902 | 1.092 | 0.874 | Pooled | allcausedeath |
| s_hdl_pl | 4260 | 1.114 | 1.016 | 1.222 | 0.022 | Pooled | allcausedeath |
| s_hdl_c | 4248 | 0.870 | 0.793 | 0.954 | 0.003 | Pooled | allcausedeath |
| s_hdl_ce | 4246 | 0.860 | 0.785 | 0.942 | 0.001 | Pooled | allcausedeath |
| s_hdl_fc | 4260 | 1.038 | 0.944 | 1.141 | 0.442 | Pooled | allcausedeath |
| s_hdl_tg | 4258 | 1.030 | 0.937 | 1.132 | 0.544 | Pooled | allcausedeath |
| vldl_d | 4259 | 1.015 | 0.922 | 1.117 | 0.761 | Pooled | allcausedeath |
| ldl_d | 4259 | 1.415 | 1.295 | 1.546 | 0.000 | Pooled | allcausedeath |
| hdl_d | 4259 | 1.121 | 1.024 | 1.228 | 0.013 | Pooled | allcausedeath |


| Metabolite | N | Hazard Ratio | Lower 95\% CI | $\begin{gathered} \text { Upper } \\ 95 \% \text { CI } \end{gathered}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| serum_c | 4259 | 0.943 | 0.857 | 1.037 | 0.224 | Pooled | allcausedeath |
| vldl_c | 4259 | 0.973 | 0.883 | 1.072 | 0.581 | Pooled | allcausedeath |
| remnant_c | 4259 | 0.958 | 0.870 | 1.055 | 0.383 | Pooled | allcausedeath |
| ldl_c | 4259 | 0.890 | 0.808 | 0.979 | 0.017 | Pooled | allcausedeath |
| hdl_c | 4259 | 1.076 | 0.980 | 1.181 | 0.127 | Pooled | allcausedeath |
| hdl2_c | 4259 | 1.084 | 0.987 | 1.190 | 0.092 | Pooled | allcausedeath |
| hdl3_c | 4259 | 0.973 | 0.885 | 1.069 | 0.564 | Pooled | allcausedeath |
| estc | 4255 | 0.939 | 0.854 | 1.033 | 0.198 | Pooled | allcausedeath |
| freec | 4255 | 0.951 | 0.865 | 1.046 | 0.304 | Pooled | allcausedeath |
| serum_tg | 4259 | 1.010 | 0.917 | 1.113 | 0.833 | Pooled | allcausedeath |
| vldl_tg | 4259 | 0.997 | 0.904 | 1.099 | 0.952 | Pooled | allcausedeath |
| ldl_tg | 4259 | 1.042 | 0.950 | 1.142 | 0.381 | Pooled | allcausedeath |
| hdl_tg | 4259 | 1.136 | 1.038 | 1.244 | 0.006 | Pooled | allcausedeath |
| totpg | 4255 | 1.043 | 0.949 | 1.146 | 0.385 | Pooled | allcausedeath |


| Metabolite | N | Hazard Ratio |  | Lower <br> 95\% CI | $\begin{gathered} \text { Upper } \\ 95 \% \text { CI } \end{gathered}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| tg_pg | 4254 |  | 0.985 | 0.894 | 1.086 | 0.762 | Pooled | allcausedeath |
| pc | 4255 |  | 1.043 | 0.950 | 1.146 | 0.379 | Pooled | allcausedeath |
| sm | 4255 |  | 0.980 | 0.891 | 1.077 | 0.671 | Pooled | allcausedeath |
| totcho | 4255 |  | 1.059 | 0.964 | 1.162 | 0.234 | Pooled | allcausedeath |
| apoa1 | 4259 |  | 1.067 | 0.972 | 1.172 | 0.173 | Pooled | allcausedeath |
| apob | 4258 |  | 0.935 | 0.849 | 1.031 | 0.176 | Pooled | allcausedeath |
| apob_apoa1 | 4258 |  | 0.918 | 0.833 | 1.010 | 0.080 | Pooled | allcausedeath |
| totfa | 4251 |  | 0.989 | 0.898 | 1.089 | 0.823 | Pooled | allcausedeath |
| unsat | 4251 |  | 0.850 | 0.775 | 0.933 | 0.001 | Pooled | allcausedeath |
| dha | 4251 |  | 0.911 | 0.824 | 1.007 | 0.069 | Pooled | allcausedeath |
| la | 4251 |  | 0.903 | 0.820 | 0.995 | 0.039 | Pooled | allcausedeath |
| faw3 | 4251 |  | 0.876 | 0.792 | 0.970 | 0.011 | Pooled | allcausedeath |
| faw6 | 4251 |  | 0.918 | 0.833 | 1.011 | 0.081 | Pooled | allcausedeath |
| pufa | 4251 |  | 0.905 | 0.821 | 0.997 | 0.043 | Pooled | allcausedeath |


| Metabolite | N | Hazard Ratio | Lower $95 \% \mathrm{CI}$ | Upper $95 \% \text { CI }$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| mufa | 4251 | 1.038 | 0.944 | 1.140 | 0.441 | Pooled | allcausedeath |
| sfa | 4251 | 1.016 | 0.923 | 1.117 | 0.751 | Pooled | allcausedeath |
| dha_fa | 4251 | 0.900 | 0.815 | 0.994 | 0.038 | Pooled | allcausedeath |
| la_fa | 4251 | 0.855 | 0.782 | 0.935 | 0.001 | Pooled | allcausedeath |
| faw3_fa | 4251 | 0.828 | 0.746 | 0.919 | 0.000 | Pooled | allcausedeath |
| faw6_fa | 4251 | 0.870 | 0.794 | 0.952 | 0.003 | Pooled | allcausedeath |
| pufa_fa | 4251 | 0.837 | 0.765 | 0.916 | 0.000 | Pooled | allcausedeath |
| mufa_fa | 4251 | 1.155 | 1.054 | 1.265 | 0.002 | Pooled | allcausedeath |
| sfa_fa | 4251 | 1.158 | 1.054 | 1.272 | 0.002 | Pooled | allcausedeath |
| glc | 4245 | 1.145 | 1.066 | 1.230 | 0.000 | Pooled | allcausedeath |
| lac | 4258 | 1.075 | 0.999 | 1.157 | 0.052 | Pooled | allcausedeath |
| pyr | 4254 | 1.022 | 0.958 | 1.090 | 0.506 | Pooled | allcausedeath |
| cit | 4256 | 1.016 | 0.923 | 1.117 | 0.748 | Pooled | allcausedeath |
| ala | 4258 | 0.985 | 0.898 | 1.081 | 0.757 | Pooled | allcausedeath |


| Metabolite | N | Hazard Ratio | Lower $95 \% \mathrm{CI}$ | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $g \ln$ | 4256 | 0.901 | 0.820 | 0.990 | 0.030 | Pooled | allcausedeath |
| gly | 4242 | 1.022 | 0.933 | 1.119 | 0.641 | Pooled | allcausedeath |
| his | 4246 | 1.013 | 0.925 | 1.110 | 0.773 | Pooled | allcausedeath |
| ile | 4257 | 1.011 | 0.919 | 1.112 | 0.823 | Pooled | allcausedeath |
| leu | 4258 | 0.995 | 0.905 | 1.094 | 0.917 | Pooled | allcausedeath |
| val | 4256 | 0.954 | 0.866 | 1.051 | 0.339 | Pooled | allcausedeath |
| phe | 4258 | 1.193 | 1.094 | 1.300 | 0.000 | Pooled | allcausedeath |
| tyr | 4246 | 1.036 | 0.945 | 1.136 | 0.446 | Pooled | allcausedeath |
| ace | 4258 | 1.105 | 1.058 | 1.155 | 0.000 | Pooled | allcausedeath |
| acace | 4258 | 1.100 | 1.014 | 1.194 | 0.022 | Pooled | allcausedeath |
| bohbut | 4136 | 1.037 | 0.949 | 1.133 | 0.424 | Pooled | allcausedeath |
| crea | 4245 | 0.952 | 0.864 | 1.048 | 0.315 | Pooled | allcausedeath |
| alb | 4259 | 1.001 | 0.909 | 1.101 | 0.991 | Pooled | allcausedeath |
| gp | 4258 | 1.100 | 1.002 | 1.207 | 0.046 | Pooled | allcausedeath |

Appendix D Table D 6:Cox regression results for PCa mRS and All-cause mRS for all-cause mortality, in controls and cases and controls (pooled)

| Metabolomic model | Outcome | Regression model | participant category | Hazard ratio | Lower 95\%CI | $\begin{gathered} \text { Upper } 95 \% \\ \text { CI } \end{gathered}$ | p-value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| PCa mRS | All-cause death | Minimally adjusted | Controls | 1.15 | 1.01 | 1.32 | 0.039 |


| PCa mRS | All-cause death | Minimally adjusted | Cases and controls(pooled) | 1.16 | 1.06 | 1.27 | 0.002 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :---: |
| All-cause mRS | All-cause death | Minimally adjusted | Controls | 1.37 | 1.22 | 1.55 | $<0.0001$ |
| All-cause mRS | All-cause death | Minimally adjusted | Cases and controls(pooled) | 1.36 | 1.25 | 1.48 | $<0.0001$ |

Appendix D Table D 7: Linear regression results for the associations between individual biomarkers and age in cases, controls and pooled cases and controls with follow-up data in the CAP trial (4,260 participants: 2,167 controls, 2,093 cases)

| Metabolite | N | Beta | Lower 95\% CI | $\begin{array}{\|l\|} \hline \text { Upper } \\ 95 \% \text { CI } \end{array}$ | pvalue | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| acace | 2093 | -0.009 | -0.018 | 0.000 | 0.04591 | Cases | age |
| acace | 4258 | -0.006 | -0.012 | 0.000 | 0.04864 | Pooled | age |
| acace | 2165 | -0.004 | -0.012 | 0.005 | 0.40141 | Controls | age |
| ace | 4258 | 0.004 | -0.002 | 0.010 | 0.15627 | Pooled | age |
| ace | 2093 | 0.004 | -0.003 | 0.012 | 0.26777 | Cases | age |
| ace | 2165 | 0.004 | -0.005 | 0.013 | 0.34481 | Controls | age |
| ala | 4258 | 0.010 | 0.004 | 0.016 | 0.00130 | Pooled | age |
| ala | 2165 | 0.011 | 0.003 | 0.019 | 0.00998 | Controls | age |
| ala | 2093 | 0.009 | 0.000 | 0.017 | 0.05312 | Cases | age |
| alb | 4259 | -0.022 | -0.028 | -0.016 | 0.00000 | Pooled | age |
| alb | 2093 | -0.024 | -0.033 | -0.016 | 0.00000 | Cases | age |
| alb | 2166 | -0.020 | -0.029 | -0.012 | 0.00000 | Controls | age |
| apoa1 | 2166 | -0.020 | -0.028 | -0.011 | 0.00001 | Controls | age |
| apoa1 | 4259 | -0.013 | -0.019 | -0.007 | 0.00003 | Pooled | age |
| apoa1 | 2093 | -0.005 | -0.014 | 0.003 | 0.20894 | Cases | age |
| apob_apoa1 | 4258 | -0.012 | -0.018 | -0.006 | 0.00011 | Pooled | age |
| apob_apoa1 | 2166 | -0.014 | -0.022 | -0.005 | 0.00152 | Controls | age |
| apob_apoa1 | 2092 | -0.010 | -0.018 | -0.001 | 0.02231 | Cases | age |
| apob | 4258 | -0.018 | -0.024 | -0.012 | 0.00000 | Pooled | age |


| Metabolite | N | Beta | Lower 95\% CI | $\begin{aligned} & \hline \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | $\mathbf{p}-$ <br> value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| apob | 2166 | -0.023 | -0.031 | -0.014 | 0.00000 | Controls | age |
| apob | 2092 | -0.013 | -0.021 | -0.004 | 0.00370 | Cases | age |
| bohbut | 2095 | 0.006 | -0.002 | 0.015 | 0.13725 | Controls | age |
| bohbut | 4136 | 0.003 | -0.004 | 0.009 | 0.42075 | Pooled | age |
| bohbut | 2041 | -0.002 | -0.011 | 0.007 | 0.63748 | Cases | age |
| cit | 4256 | 0.027 | 0.021 | 0.033 | 0.00000 | Pooled | age |
| cit | 2092 | 0.028 | 0.020 | 0.037 | 0.00000 | Cases | age |
| cit | 2164 | 0.026 | 0.018 | 0.035 | 0.00000 | Controls | age |
| crea | 4245 | 0.021 | 0.015 | 0.027 | 0.00000 | Pooled | age |
| crea | 2087 | 0.022 | 0.013 | 0.031 | 0.00000 | Cases | age |
| crea | 2158 | 0.020 | 0.012 | 0.028 | 0.00000 | Controls | age |
| dha_fa | 4251 | 0.019 | 0.013 | 0.025 | 0.00000 | Pooled | age |
| dha_fa | 2162 | 0.024 | 0.015 | 0.033 | 0.00000 | Controls | age |
| dha_fa | 2089 | 0.013 | 0.004 | 0.021 | 0.00339 | Cases | age |
| dha | 2089 | 0.004 | -0.004 | 0.012 | 0.34732 | Cases | age |
| dha | 4251 | 0.003 | -0.003 | 0.009 | 0.40705 | Pooled | age |
| dha | 2162 | 0.001 | -0.007 | 0.010 | 0.77957 | Controls | age |
| estc | 4255 | -0.020 | -0.026 | -0.014 | 0.00000 | Pooled | age |
| estc | 2165 | -0.026 | -0.034 | -0.017 | 0.00000 | Controls | age |
| estc | 2090 | -0.013 | -0.022 | -0.005 | 0.00235 | Cases | age |
| faw3_fa | 4251 | 0.017 | 0.011 | 0.023 | 0.00000 | Pooled | age |
| faw3_fa | 2162 | 0.020 | 0.012 | 0.029 | 0.00000 | Controls | age |



| Metabolite | N | Beta | Lower $95 \% \text { CI }$ | Upper $95 \% \text { CI }$ | $\left\lvert\, \begin{aligned} & \mathrm{p}- \\ & \text { value } \end{aligned}\right.$ | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| gp | 2165 | -0.006 | -0.015 | 0.003 | 0.18165 | Controls | age |
| gp | 4258 | -0.003 | -0.009 | 0.003 | 0.27483 | Pooled | age |
| gp | 2093 | -0.001 | -0.009 | 0.007 | 0.84316 | Cases | age |
| hdl_c | 2166 | -0.007 | -0.016 | 0.001 | 0.09103 | Controls | age |
| hdl_c | 4259 | -0.004 | -0.010 | 0.002 | 0.16884 | Pooled | age |
| hdl_c | 2093 | -0.001 | -0.009 | 0.008 | 0.88002 | Cases | age |
| hdl_d | 2093 | 0.010 | 0.002 | 0.019 | 0.01928 | Cases | age |
| hdl_d | 4259 | 0.005 | -0.001 | 0.012 | 0.07518 | Pooled | age |
| hdl_d | 2166 | 0.001 | -0.007 | 0.010 | 0.79371 | Controls | age |
| hdl_tg | 2166 | -0.011 | -0.019 | -0.002 | 0.01570 | Controls | age |
| hdl_tg | 4259 | -0.004 | -0.010 | 0.002 | 0.18245 | Pooled | age |
| hdl_tg | 2093 | 0.003 | -0.006 | 0.011 | 0.54597 | Cases | age |
| hdl2_c | 2166 | -0.007 | -0.016 | 0.002 | 0.10682 | Controls | age |
| hdl2_c | 4259 | -0.004 | -0.010 | 0.002 | 0.20021 | Pooled | age |
| hdl2_c | 2093 | 0.000 | -0.009 | 0.008 | 0.90870 | Cases | age |
| hdl3_c | 2166 | -0.009 | -0.018 | 0.000 | 0.03885 | Controls | age |
| hdl3_c | 4259 | -0.006 | -0.012 | 0.000 | 0.05234 | Pooled | age |
| hdl3_c | 2093 | -0.002 | -0.011 | 0.006 | 0.56791 | Cases | age |
| his | 2088 | 0.003 | -0.006 | 0.011 | 0.51289 | Cases | age |
| his | 2158 | -0.002 | -0.010 | 0.007 | 0.70642 | Controls | age |
| his | 4246 | 0.000 | -0.006 | 0.007 | 0.88277 | Pooled | age |
| idl_c | 4260 | -0.017 | -0.023 | -0.011 | 0.00000 | Pooled | age |


| Metabolite | N |  | Beta |  | Lower $95 \% \text { CI }$ | Upper $95 \% \text { CI }$ | $\mathrm{p}-$ <br> value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| idl_c |  | 2167 |  | -0.021 | -0.030 | -0.013 | 0.00000 | Controls | age |
| idl_c |  | 2093 |  | -0.012 | -0.021 | -0.004 | 0.00576 | Cases | age |
| idl_ce |  | 4260 |  | -0.018 | -0.024 | -0.012 | 0.00000 | Pooled | age |
| idl_ce |  | 2167 |  | -0.023 | -0.031 | -0.014 | 0.00000 | Controls | age |
| idl_ce |  | 2093 |  | -0.013 | -0.021 | -0.004 | 0.00333 | Cases | age |
| idl_fc |  | 4260 |  | -0.013 | -0.019 | -0.007 | 0.00003 | Pooled | age |
| idl_fc |  | 2167 |  | -0.016 | -0.025 | -0.008 | 0.00024 | Controls | age |
| idl_fc |  | 2093 |  | -0.009 | -0.018 | -0.001 | 0.03310 | Cases | age |
| idl_l |  | 4259 |  | -0.016 | -0.022 | -0.010 | 0.00000 | Pooled | age |
| idl_l |  | 2166 |  | -0.021 | -0.029 | -0.012 | 0.00000 | Controls | age |
| idl_l |  | 2093 |  | -0.011 | -0.020 | -0.003 | 0.01074 | Cases | age |
| idl_p |  | 4259 |  | -0.016 | -0.022 | -0.009 | 0.00000 | Pooled | age |
| idl_p |  | 2166 |  | -0.020 | -0.028 | -0.011 | 0.00001 | Controls | age |
| idl_p |  | 2093 |  | -0.011 | -0.019 | -0.002 | 0.01291 | Cases | age |
| idl_pl |  | 4260 |  | -0.016 | -0.022 | -0.010 | 0.00000 | Pooled | age |
| idl_pl |  | 2167 |  | -0.020 | -0.028 | -0.011 | 0.00001 | Controls | age |
| idl_pl |  | 2093 |  | -0.011 | -0.020 | -0.003 | 0.00993 | Cases | age |
| idl_tg |  | 2164 |  | -0.006 | -0.014 | 0.003 | 0.19023 | Controls | age |
| idl_tg |  | 2092 |  | 0.003 | -0.005 | 0.012 | 0.44699 | Cases | age |
| idl_tg |  | 4256 |  | -0.001 | -0.007 | 0.005 | 0.66290 | Pooled | age |
| ile |  | 2165 |  | -0.011 | -0.020 | -0.002 | 0.01418 | Controls | age |
| ile |  | 4257 |  | -0.005 | -0.011 | 0.001 | 0.09324 | Pooled | age |


| Metabolite | N | Beta | Lower $95 \% \text { CI }$ | Upper $95 \% \text { CI }$ | $\left\lvert\, \begin{aligned} & \mathrm{p}- \\ & \text { value } \end{aligned}\right.$ | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| ile | 2092 | 0.001 | -0.008 | 0.009 | 0.90158 | Cases | age |
| 1_hdl_c | 2092 | 0.009 | 0.000 | 0.017 | 0.04905 | Cases | age |
| 1_hdl_c | 4259 | 0.005 | -0.001 | 0.011 | 0.08775 | Pooled | age |
| 1_hdl_c | 2167 | 0.002 | -0.006 | 0.011 | 0.61720 | Controls | age |
| l_hdl_ce | 2092 | 0.009 | 0.000 | 0.017 | 0.04568 | Cases | age |
| 1_hdl_ce | 4259 | 0.005 | -0.001 | 0.011 | 0.08671 | Pooled | age |
| 1_hdl_ce | 2167 | 0.002 | -0.007 | 0.011 | 0.63306 | Controls | age |
| 1_hdl_fc | 2092 | 0.008 | 0.000 | 0.017 | 0.06019 | Cases | age |
| 1_hdl_fc | 4259 | 0.005 | -0.001 | 0.011 | 0.10449 | Pooled | age |
| 1_hdl_fc | 2167 | 0.002 | -0.006 | 0.011 | 0.61902 | Controls | age |
| 1_hdl_1 | 2092 | 0.008 | -0.001 | 0.016 | 0.08054 | Cases | age |
| 1_hdl_1 | 4259 | 0.004 | -0.002 | 0.010 | 0.19967 | Pooled | age |
| 1_hdl_1 | 2167 | 0.001 | -0.008 | 0.009 | 0.89558 | Controls | age |
| 1_hdl_p | 2092 | 0.008 | -0.001 | 0.016 | 0.08088 | Cases | age |
| 1_hdl_p | 4259 | 0.004 | -0.002 | 0.010 | 0.25025 | Pooled | age |
| 1_hdl_p | 2167 | 0.000 | -0.009 | 0.008 | 0.97866 | Controls | age |
| 1_hdl_pl | 2093 | 0.006 | -0.003 | 0.014 | 0.19901 | Cases | age |
| 1_hdl_pl | 4260 | 0.002 | -0.004 | 0.008 | 0.53914 | Pooled | age |
| 1_hdl_pl | 2167 | -0.001 | -0.010 | 0.007 | 0.73974 | Controls | age |
| 1_hdl_tg | 2091 | 0.017 | 0.009 | 0.026 | 0.00008 | Cases | age |
| l_hdl_tg | 4254 | 0.009 | 0.002 | 0.015 | 0.00562 | Pooled | age |
| 1_hdl_tg | 2163 | 0.000 | -0.008 | 0.009 | 0.93531 | Controls | age |


| Metabolite | N | Beta | Lower 95\% CI | $\begin{aligned} & \hline \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | $\begin{aligned} & \mathrm{p}- \\ & \text { value } \end{aligned}$ | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1_ldl_c | 4260 | -0.018 | -0.024 | -0.011 | 0.00000 | Pooled | age |
| 1_ldl_c | 2167 | -0.021 | -0.030 | -0.013 | 0.00000 | Controls | age |
| 1_ldl_c | 2093 | -0.013 | -0.022 | -0.005 | 0.00273 | Cases | age |
| 1_ldl_ce | 4260 | -0.018 | -0.024 | -0.012 | 0.00000 | Pooled | age |
| 1_ldl_ce | 2167 | -0.022 | -0.031 | -0.014 | 0.00000 | Controls | age |
| 1_ldl_ce | 2093 | -0.014 | -0.022 | -0.005 | 0.00179 | Cases | age |
| 1_ldl_fc | 4260 | -0.015 | -0.021 | -0.009 | 0.00000 | Pooled | age |
| 1_ldl_fc | 2167 | -0.018 | -0.027 | -0.009 | 0.00004 | Controls | age |
| 1_ldl_fc | 2093 | -0.011 | -0.020 | -0.003 | 0.01027 | Cases | age |
| 1_ldl_1 | 4260 | -0.017 | -0.023 | -0.011 | 0.00000 | Pooled | age |
| 1_ldl_1 | 2167 | -0.022 | -0.030 | -0.013 | 0.00000 | Controls | age |
| 1_ldl_1 | 2093 | -0.013 | -0.021 | -0.004 | 0.00412 | Cases | age |
| 1_ldl_p | 4260 | -0.017 | -0.023 | -0.011 | 0.00000 | Pooled | age |
| 1_ldl_p | 2167 | -0.021 | -0.029 | -0.012 | 0.00000 | Controls | age |
| 1_ldl_p | 2093 | -0.012 | -0.021 | -0.004 | 0.00477 | Cases | age |
| 1_ldl_pl | 4260 | -0.019 | -0.025 | -0.013 | 0.00000 | Pooled | age |
| 1_ldl_pl | 2167 | -0.023 | -0.031 | -0.014 | 0.00000 | Controls | age |
| 1_ldl_pl | 2093 | -0.014 | -0.022 | -0.005 | 0.00153 | Cases | age |
| 1_ldl_tg | 2165 | -0.008 | -0.017 | 0.000 | 0.06308 | Controls | age |
| 1_ldl_tg | 4257 | -0.004 | -0.010 | 0.003 | 0.25410 | Pooled | age |
| 1_ldl_tg | 2092 | 0.001 | -0.007 | 0.010 | 0.73093 | Cases | age |
| 1_vldl_c | 2161 | -0.012 | -0.021 | -0.003 | 0.00980 | Controls | age |


| Metabolite | N |  | Beta |  | Lower $95 \% \text { CI }$ | $\begin{aligned} & \hline \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | $\begin{array}{\|l\|} \mathrm{p}- \\ \text { value } \end{array}$ | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1_vldl_c |  | 4249 |  | -0.008 | -0.014 | -0.002 | 0.01063 | Pooled | age |
| 1_vldl_c |  | 2088 |  | -0.004 | -0.012 | 0.004 | 0.31413 | Cases | age |
| 1_vldl_ce |  | 2162 |  | -0.013 | -0.022 | -0.004 | 0.00334 | Controls | age |
| 1_vldl_ce |  | 4250 |  | -0.009 | -0.015 | -0.003 | 0.00438 | Pooled | age |
| l_vldl_ce |  | 2088 |  | -0.004 | -0.013 | 0.004 | 0.29096 | Cases | age |
| 1_vldl_fc |  | 2159 |  | -0.011 | -0.020 | -0.002 | 0.01501 | Controls | age |
| 1_vldl_fc |  | 4248 |  | -0.007 | -0.014 | -0.001 | 0.01595 | Pooled | age |
| 1_vldl_fc |  | 2089 |  | -0.004 | -0.012 | 0.004 | 0.34117 | Cases | age |
| 1_vldl_1 |  | 4250 |  | -0.007 | -0.013 | -0.001 | 0.01846 | Pooled | age |
| 1_vldl_1 |  | 2161 |  | -0.010 | -0.019 | -0.001 | 0.02303 | Controls | age |
| 1_vldl_1 |  | 2089 |  | -0.004 | -0.012 | 0.004 | 0.28224 | Cases | age |
| 1_vldl_p |  | 4251 |  | -0.008 | -0.014 | -0.002 | 0.01045 | Pooled | age |
| 1_vldl_p |  | 2162 |  | -0.012 | -0.021 | -0.003 | 0.01164 | Controls | age |
| 1_vldl_p |  | 2089 |  | -0.004 | -0.012 | 0.004 | 0.28069 | Cases | age |
| 1_vldl_pl |  | 4250 |  | -0.008 | -0.014 | -0.002 | 0.01103 | Pooled | age |
| 1_vldl_pl |  | 2161 |  | -0.011 | -0.020 | -0.002 | 0.01341 | Controls | age |
| 1_vldl_pl |  | 2089 |  | -0.005 | -0.013 | 0.004 | 0.26970 | Cases | age |
| l_vldl_tg |  | 4249 |  | -0.007 | -0.013 | -0.001 | 0.01654 | Pooled | age |
| 1_vldl_tg |  | 2160 |  | -0.010 | -0.019 | -0.002 | 0.02132 | Controls | age |
| 1_vldl_tg |  | 2089 |  | -0.004 | -0.012 | 0.004 | 0.27783 | Cases | age |
| la_fa |  | 4251 |  | -0.008 | -0.014 | -0.002 | 0.00865 | Pooled | age |
| la_fa |  | 2162 |  | -0.009 | -0.018 | -0.001 | 0.03069 | Controls | age |


| Metabolite | N | Beta | Lower 95\% CI | $\begin{array}{\|l\|} \hline \text { Upper } \\ 95 \% \text { CI } \end{array}$ | $\mathbf{p}-$ value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| la_fa | 2089 | -0.007 | -0.016 | 0.001 | 0.10277 | Cases | age |
| la | 4251 | -0.020 | -0.026 | -0.014 | 0.00000 | Pooled | age |
| la | 2162 | -0.027 | -0.036 | -0.019 | 0.00000 | Controls | age |
| la | 2089 | -0.012 | -0.021 | -0.003 | 0.00631 | Cases | age |
| lac | 2165 | 0.003 | -0.005 | 0.012 | 0.45743 | Controls | age |
| lac | 4258 | 0.001 | -0.005 | 0.007 | 0.65526 | Pooled | age |
| lac | 2093 | -0.001 | -0.010 | 0.008 | 0.76516 | Cases | age |
| ldl_c | 4259 | -0.017 | -0.023 | -0.011 | 0.00000 | Pooled | age |
| ldl_c | 2166 | -0.021 | -0.029 | -0.012 | 0.00000 | Controls | age |
| ldl_c | 2093 | -0.013 | -0.022 | -0.005 | 0.00235 | Cases | age |
| ldl_d | 4259 | 0.009 | 0.003 | 0.015 | 0.00467 | Pooled | age |
| ldl_d | 2093 | 0.010 | 0.001 | 0.019 | 0.02215 | Cases | age |
| ldl_d | 2166 | 0.008 | -0.001 | 0.016 | 0.07728 | Controls | age |
| ldl_tg | 2166 | -0.010 | -0.019 | -0.002 | 0.01672 | Controls | age |
| ldl_tg | 4259 | -0.006 | -0.012 | 0.000 | 0.07080 | Pooled | age |
| ldl_tg | 2093 | 0.000 | -0.009 | 0.008 | 0.94514 | Cases | age |
| leu | 2165 | -0.011 | -0.020 | -0.003 | 0.00990 | Controls | age |
| leu | 4258 | -0.006 | -0.012 | 0.000 | 0.04959 | Pooled | age |
| leu | 2093 | -0.001 | -0.009 | 0.008 | 0.88497 | Cases | age |
| m_hdl_c | 4260 | -0.013 | -0.019 | -0.007 | 0.00003 | Pooled | age |
| m_hdl_c | 2167 | -0.015 | -0.023 | -0.006 | 0.00073 | Controls | age |
| m_hdl_c | 2093 | -0.011 | -0.019 | -0.002 | 0.01440 | Cases | age |


| Metabolite | N |  | Beta | Lower $95 \% \text { CI }$ | $\begin{array}{\|l\|} \hline \text { Upper } \\ 95 \% \text { CI } \end{array}$ | $\mathrm{p}-$ <br> value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_hdl_ce |  | 4260 | -0.013 | -0.019 | -0.007 | 0.00001 | Pooled | age |
| m_hdl_ce |  | 2167 | -0.015 | -0.024 | -0.007 | 0.00046 | Controls | age |
| m_hdl_ce |  | 2093 | -0.011 | -0.020 | -0.003 | 0.00899 | Cases | age |
| m_hdl_fc |  | 4260 | -0.010 | -0.016 | -0.004 | 0.00090 | Pooled | age |
| m_hdl_fc |  | 2167 | -0.013 | -0.021 | -0.004 | 0.00346 | Controls | age |
| m_hdl_fc |  | 2093 | -0.007 | -0.016 | 0.001 | 0.08918 | Cases | age |
| m_hdl_1 |  | 4260 | -0.013 | -0.019 | -0.007 | 0.00002 | Pooled | age |
| m_hdl_1 |  | 2167 | -0.016 | -0.025 | -0.008 | 0.00017 | Controls | age |
| m_hdl_1 |  | 2093 | -0.010 | -0.018 | -0.001 | 0.02733 | Cases | age |
| m_hdl_p |  | 4260 | -0.014 | -0.020 | -0.008 | 0.00001 | Pooled | age |
| m_hdl_p |  | 2167 | -0.017 | -0.026 | -0.009 | 0.00008 | Controls | age |
| m_hdl_p |  | 2093 | -0.010 | -0.018 | -0.001 | 0.02623 | Cases | age |
| m_hdl_pl |  | 4259 | -0.011 | -0.017 | -0.005 | 0.00026 | Pooled | age |
| m_hdl_pl |  | 2167 | -0.015 | -0.024 | -0.006 | 0.00056 | Controls | age |
| m_hdl_pl |  | 2092 | -0.007 | -0.016 | 0.001 | 0.10143 | Cases | age |
| m_hdl_tg |  | 2166 | -0.016 | -0.025 | -0.007 | 0.00039 | Controls | age |
| m_hdl_tg |  | 4258 | -0.011 | -0.017 | -0.005 | 0.00048 | Pooled | age |
| m_hdl_tg |  | 2092 | -0.006 | -0.014 | 0.003 | 0.19102 | Cases | age |
| m_ldl_c |  | 4260 | -0.017 | -0.023 | -0.011 | 0.00000 | Pooled | age |
| m_ldl_c |  | 2167 | -0.020 | -0.028 | -0.011 | 0.00001 | Controls | age |
| m_ldl_c |  | 2093 | -0.013 | -0.022 | -0.005 | 0.00235 | Cases | age |
| m_ldl_ce |  | 4260 | -0.016 | -0.022 | -0.010 | 0.00000 | Pooled | age |


| Metabolite | N |  | Beta |  | Lower $95 \% \text { CI }$ | $\begin{array}{\|l\|} \hline \text { Upper } \\ 95 \% \text { CI } \end{array}$ | $\mathrm{p}-$ <br> value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_ldl_ce |  | 2167 |  | -0.019 | -0.028 | -0.010 | 0.00001 | Controls | age |
| m_ldl_ce |  | 2093 |  | -0.013 | -0.022 | -0.004 | 0.00286 | Cases | age |
| m_ldl_fc |  | 4260 |  | -0.018 | -0.024 | -0.012 | 0.00000 | Pooled | age |
| m_ldl_fc |  | 2167 |  | -0.022 | -0.030 | -0.013 | 0.00000 | Controls | age |
| m_ldl_fc |  | 2093 |  | -0.014 | -0.023 | -0.006 | 0.00119 | Cases | age |
| m_ldl_1 |  | 4260 |  | -0.017 | -0.023 | -0.011 | 0.00000 | Pooled | age |
| m_ldl_1 |  | 2167 |  | -0.021 | -0.030 | -0.013 | 0.00000 | Controls | age |
| m_ldl_1 |  | 2093 |  | -0.013 | -0.022 | -0.005 | 0.00218 | Cases | age |
| m_ldl_p |  | 4260 |  | -0.017 | -0.023 | -0.011 | 0.00000 | Pooled | age |
| m_ldl_p |  | 2167 |  | -0.020 | -0.029 | -0.012 | 0.00000 | Controls | age |
| m_ldl_p |  | 2093 |  | -0.013 | -0.022 | -0.005 | 0.00251 | Cases | age |
| m_ldl_pl |  | 4260 |  | -0.021 | -0.027 | -0.015 | 0.00000 | Pooled | age |
| m_ldl_pl |  | 2167 |  | -0.025 | -0.034 | -0.017 | 0.00000 | Controls | age |
| m_ldl_pl |  | 2093 |  | -0.016 | -0.024 | -0.007 | 0.00033 | Cases | age |
| m_ldl_tg |  | 2166 |  | -0.010 | -0.018 | -0.001 | 0.02326 | Controls | age |
| m_ldl_tg |  | 4258 |  | -0.005 | -0.011 | 0.001 | 0.09374 | Pooled | age |
| m_ldl_tg |  | 2092 |  | 0.000 | -0.009 | 0.008 | 0.99276 | Cases | age |
| m_vldl_c |  | 2165 |  | -0.014 | -0.023 | -0.005 | 0.00156 | Controls | age |
| m_vldl_c |  | 4254 |  | -0.009 | -0.015 | -0.003 | 0.00269 | Pooled | age |
| m_vldl_c |  | 2089 |  | -0.004 | -0.013 | 0.004 | 0.30705 | Cases | age |
| m_vldl_ce |  | 2166 |  | -0.017 | -0.026 | -0.008 | 0.00012 | Controls | age |
| m_vldl_ce |  | 4257 |  | -0.012 | -0.018 | -0.006 | 0.00013 | Pooled | age |


| Metabolite | N | Beta | Lower $95 \% \text { CI }$ | Upper $95 \% \text { CI }$ | $\left\lvert\, \begin{aligned} & \mathrm{p}- \\ & \text { value } \end{aligned}\right.$ | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_vldl_ce | 2091 | -0.006 | -0.015 | 0.002 | 0.12972 | Cases | age |
| m_vldl_fc | 2165 | -0.011 | -0.020 | -0.003 | 0.01156 | Controls | age |
| m_vldl_fc | 4253 | -0.008 | -0.014 | -0.002 | 0.01433 | Pooled | age |
| m_vldl_fc | 2088 | -0.004 | -0.012 | 0.004 | 0.36605 | Cases | age |
| m_vldl_1 | 4251 | -0.008 | -0.014 | -0.002 | 0.00877 | Pooled | age |
| m_vldl_1 | 2163 | -0.012 | -0.020 | -0.003 | 0.00971 | Controls | age |
| m_vldl_1 | 2088 | -0.005 | -0.013 | 0.003 | 0.25500 | Cases | age |
| m_vldl_p | 4251 | -0.009 | -0.015 | -0.003 | 0.00509 | Pooled | age |
| m_vldl_p | 2163 | -0.013 | -0.022 | -0.004 | 0.00533 | Controls | age |
| m_vldl_p | 2088 | -0.005 | -0.013 | 0.003 | 0.25369 | Cases | age |
| m_vldl_pl | 2163 | -0.013 | -0.022 | -0.004 | 0.00388 | Controls | age |
| m_vldl_pl | 4252 | -0.009 | -0.015 | -0.003 | 0.00435 | Pooled | age |
| m_vldl_pl | 2089 | -0.005 | -0.013 | 0.004 | 0.27461 | Cases | age |
| m_vldl_tg | 4252 | -0.008 | -0.014 | -0.001 | 0.01480 | Pooled | age |
| m_vldl_tg | 2163 | -0.011 | -0.019 | -0.002 | 0.01865 | Controls | age |
| m_vldl_tg | 2089 | -0.005 | -0.013 | 0.004 | 0.27115 | Cases | age |
| mufa_fa | 2162 | -0.003 | -0.012 | 0.005 | 0.43426 | Controls | age |
| mufa_fa | 4251 | -0.001 | -0.008 | 0.005 | 0.63224 | Pooled | age |
| mufa_fa | 2089 | 0.000 | -0.009 | 0.008 | 0.99359 | Cases | age |
| mufa | 2162 | -0.019 | -0.028 | -0.011 | 0.00001 | Controls | age |
| mufa | 4251 | -0.013 | -0.019 | -0.007 | 0.00001 | Pooled | age |
| mufa | 2089 | -0.007 | -0.015 | 0.001 | 0.09244 | Cases | age |


| Metabolite | N | Beta | Lower $95 \% \text { CI }$ | Upper $95 \% \text { CI }$ | $\mathbf{p}-$ <br> value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| pc | 2165 | -0.025 | -0.033 | -0.016 | 0.00000 | Controls | age |
| pc | 4255 | -0.016 | -0.022 | -0.010 | 0.00000 | Pooled | age |
| pc | 2090 | -0.006 | -0.014 | 0.003 | 0.20165 | Cases | age |
| phe | 4258 | 0.017 | 0.011 | 0.023 | 0.00000 | Pooled | age |
| phe | 2093 | 0.020 | 0.011 | 0.028 | 0.00001 | Cases | age |
| phe | 2165 | 0.015 | 0.007 | 0.024 | 0.00036 | Controls | age |
| pufa_fa | 2089 | -0.001 | -0.010 | 0.007 | 0.77512 | Cases | age |
| pufa_fa | 4251 | 0.000 | -0.006 | 0.006 | 0.90770 | Pooled | age |
| pufa_fa | 2162 | 0.000 | -0.008 | 0.009 | 0.95104 | Controls | age |
| pufa | 4251 | -0.017 | -0.023 | -0.011 | 0.00000 | Pooled | age |
| pufa | 2162 | -0.024 | -0.032 | -0.015 | 0.00000 | Controls | age |
| pufa | 2089 | -0.010 | -0.018 | -0.001 | 0.02605 | Cases | age |
| pyr | 4254 | 0.004 | -0.002 | 0.010 | 0.22801 | Pooled | age |
| pyr | 2091 | 0.005 | -0.005 | 0.014 | 0.34645 | Cases | age |
| pyr | 2163 | 0.003 | -0.005 | 0.011 | 0.46742 | Controls | age |
| remnant_c | 4259 | -0.017 | -0.023 | -0.011 | 0.00000 | Pooled | age |
| remnant_c | 2166 | -0.023 | -0.031 | -0.014 | 0.00000 | Controls | age |
| remnant_c | 2093 | -0.011 | -0.019 | -0.002 | 0.01205 | Cases | age |
| s_hdl_c | 4248 | -0.011 | -0.017 | -0.005 | 0.00022 | Pooled | age |
| s_hdl_c | 2087 | -0.013 | -0.021 | -0.005 | 0.00216 | Cases | age |
| s_hdl_c | 2161 | -0.010 | -0.019 | -0.001 | 0.02633 | Controls | age |
| s_hdl_ce | 4246 | -0.010 | -0.016 | -0.004 | 0.00157 | Pooled | age |


| Metabolite | N |  | Beta | Lower $95 \% \mathrm{CI}$ | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | $\mathrm{p}-$ <br> value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_hdl_ce |  | 2087 | -0.012 | -0.020 | -0.003 | 0.00613 | Cases | age |
| s_hdl_ce |  | 2159 | -0.008 | -0.017 | 0.001 | 0.07052 | Controls | age |
| s_hdl_fc |  | 4260 | -0.010 | -0.016 | -0.004 | 0.00170 | Pooled | age |
| s_hdl_fc |  | 2167 | -0.011 | -0.020 | -0.002 | 0.01165 | Controls | age |
| s_hdl_fc |  | 2093 | -0.008 | -0.017 | 0.000 | 0.06077 | Cases | age |
| s_hdl_1 |  | 4260 | -0.017 | -0.023 | -0.011 | 0.00000 | Pooled | age |
| s_hdl_1 |  | 2167 | -0.017 | -0.026 | -0.009 | 0.00005 | Controls | age |
| s_hdl_1 |  | 2093 | -0.015 | -0.024 | -0.007 | 0.00045 | Cases | age |
| s_hdl_p |  | 4260 | -0.017 | -0.023 | -0.011 | 0.00000 | Pooled | age |
| s_hdl_p |  | 2167 | -0.018 | -0.026 | -0.009 | 0.00003 | Controls | age |
| s_hdl_p |  | 2093 | -0.015 | -0.024 | -0.007 | 0.00056 | Cases | age |
| s_hdl_pl |  | 4260 | -0.011 | -0.017 | -0.005 | 0.00057 | Pooled | age |
| s_hdl_pl |  | 2167 | -0.012 | -0.021 | -0.004 | 0.00500 | Controls | age |
| s_hdl_pl |  | 2093 | -0.009 | -0.017 | 0.000 | 0.04471 | Cases | age |
| s_hdl_tg |  | 2167 | -0.006 | -0.015 | 0.002 | 0.15713 | Controls | age |
| s_hdl_tg |  | 4258 | -0.003 | -0.009 | 0.003 | 0.37593 | Pooled | age |
| s_hdl_tg |  | 2091 | 0.001 | -0.008 | 0.009 | 0.85211 | Cases | age |
| s_ldl_c |  | 4260 | -0.016 | -0.023 | -0.010 | 0.00000 | Pooled | age |
| s_ldl_c |  | 2167 | -0.019 | -0.028 | -0.011 | 0.00001 | Controls | age |
| s_ldl_c |  | 2093 | -0.013 | -0.022 | -0.005 | 0.00212 | Cases | age |
| s_ldl_ce |  | 4260 | -0.016 | -0.022 | -0.010 | 0.00000 | Pooled | age |
| s_ldl_ce |  | 2167 | -0.019 | -0.027 | -0.010 | 0.00002 | Controls | age |


| Metabolite | N | Beta | Lower 95\% CI | $\begin{aligned} & \hline \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | $\mathbf{p}-$ <br> value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_ldl_ce | 2093 | -0.013 | -0.021 | -0.004 | 0.00281 | Cases | age |
| s_ldl_fc | 4260 | -0.018 | -0.024 | -0.012 | 0.00000 | Pooled | age |
| s_ldl_fc | 2167 | -0.022 | -0.030 | -0.013 | 0.00000 | Controls | age |
| s_ldl_fc | 2093 | -0.015 | -0.023 | -0.006 | 0.00086 | Cases | age |
| s_ldl_l | 4260 | -0.018 | -0.024 | -0.012 | 0.00000 | Pooled | age |
| s_ldl_1 | 2167 | -0.022 | -0.031 | -0.014 | 0.00000 | Controls | age |
| s_ldl_1 | 2093 | -0.014 | -0.023 | -0.006 | 0.00114 | Cases | age |
| s_ldl_p | 4260 | -0.018 | -0.024 | -0.012 | 0.00000 | Pooled | age |
| s_ldl_p | 2167 | -0.022 | -0.030 | -0.013 | 0.00000 | Controls | age |
| s_ldl_p | 2093 | -0.014 | -0.023 | -0.006 | 0.00121 | Cases | age |
| s_ldl_pl | 4260 | -0.022 | -0.028 | -0.016 | 0.00000 | Pooled | age |
| s_ldl_pl | 2167 | -0.027 | -0.036 | -0.019 | 0.00000 | Controls | age |
| s_ldl_pl | 2093 | -0.017 | -0.026 | -0.009 | 0.00009 | Cases | age |
| s_ldl_tg | 2167 | -0.015 | -0.024 | -0.007 | 0.00048 | Controls | age |
| s_ldl_tg | 4259 | -0.010 | -0.016 | -0.004 | 0.00131 | Pooled | age |
| s_ldl_tg | 2092 | -0.004 | -0.013 | 0.004 | 0.32228 | Cases | age |
| s_vldl_c | 4259 | -0.012 | -0.018 | -0.006 | 0.00006 | Pooled | age |
| s_vldl_c | 2167 | -0.017 | -0.026 | -0.009 | 0.00007 | Controls | age |
| s_vldl_c | 2092 | -0.007 | -0.016 | 0.001 | 0.09867 | Cases | age |
| s_vldl_ce | 4259 | -0.012 | -0.018 | -0.006 | 0.00005 | Pooled | age |
| s_vldl_ce | 2167 | -0.017 | -0.026 | -0.009 | 0.00007 | Controls | age |
| s_vldl_ce | 2092 | -0.007 | -0.016 | 0.001 | 0.09011 | Cases | age |


| Metabolite | N |  | Beta | Lower $95 \% \mathrm{CI}$ | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | $\mathrm{p}-$ <br> value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_vldl_fc |  | 2167 | -0.015 | -0.024 | -0.007 | 0.00051 | Controls | age |
| s_vldl_fc |  | 4259 | -0.011 | -0.017 | -0.005 | 0.00059 | Pooled | age |
| s_vldl_fc |  | 2092 | -0.006 | -0.014 | 0.003 | 0.17350 | Cases | age |
| s_vldl_1 |  | 4259 | -0.010 | -0.016 | -0.004 | 0.00094 | Pooled | age |
| s_vldl_l |  | 2167 | -0.014 | -0.023 | -0.006 | 0.00119 | Controls | age |
| s_vldl_1 |  | 2092 | -0.006 | -0.015 | 0.002 | 0.14423 | Cases | age |
| s_vldl_p |  | 4258 | -0.010 | -0.016 | -0.004 | 0.00121 | Pooled | age |
| s_vldl_p |  | 2167 | -0.014 | -0.023 | -0.006 | 0.00134 | Controls | age |
| s_vldl_p |  | 2091 | -0.006 | -0.014 | 0.003 | 0.18092 | Cases | age |
| s_vldl_pl |  | 4259 | -0.012 | -0.018 | -0.006 | 0.00016 | Pooled | age |
| s_vldl_pl |  | 2167 | -0.016 | -0.025 | -0.007 | 0.00027 | Controls | age |
| s_vldl_pl |  | 2092 | -0.007 | -0.016 | 0.001 | 0.09630 | Cases | age |
| s_vldl_tg |  | 2165 | -0.009 | -0.018 | 0.000 | 0.03910 | Controls | age |
| s_vldl_tg |  | 4254 | -0.006 | -0.012 | 0.000 | 0.05734 | Pooled | age |
| s_vldl_tg |  | 2089 | -0.003 | -0.011 | 0.006 | 0.54199 | Cases | age |
| serum_c |  | 4259 | -0.019 | -0.025 | -0.013 | 0.00000 | Pooled | age |
| serum_c |  | 2166 | -0.025 | -0.033 | -0.016 | 0.00000 | Controls | age |
| serum_c |  | 2093 | -0.013 | -0.021 | -0.004 | 0.00404 | Cases | age |
| serum_tg |  | 4259 | -0.008 | -0.014 | -0.002 | 0.01157 | Pooled | age |
| serum_tg |  | 2166 | -0.011 | -0.020 | -0.002 | 0.01233 | Controls | age |
| serum_tg |  | 2093 | -0.005 | -0.013 | 0.003 | 0.21413 | Cases | age |
| sfa_fa |  | 2162 | 0.008 | -0.001 | 0.016 | 0.06735 | Controls | age |


| Metabolite | N | Beta | Lower 95\% CI | $\begin{aligned} & \hline \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | $\mathrm{p}-$ <br> value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| sfa_fa | 4251 | 0.006 | 0.000 | 0.012 | 0.06920 | Pooled | age |
| sfa_fa | 2089 | 0.003 | -0.006 | 0.011 | 0.51849 | Cases | age |
| sfa | 4251 | -0.015 | -0.021 | -0.009 | 0.00000 | Pooled | age |
| sfa | 2162 | -0.021 | -0.029 | -0.012 | 0.00000 | Controls | age |
| sfa | 2089 | -0.008 | -0.016 | 0.000 | 0.05412 | Cases | age |
| sm | 4255 | -0.018 | -0.024 | -0.012 | 0.00000 | Pooled | age |
| sm | 2165 | -0.022 | -0.031 | -0.013 | 0.00000 | Controls | age |
| sm | 2090 | -0.014 | -0.022 | -0.005 | 0.00168 | Cases | age |
| tg_pg | 4254 | -0.003 | -0.009 | 0.003 | 0.36190 | Pooled | age |
| tg_pg | 2164 | -0.004 | -0.012 | 0.005 | 0.42776 | Controls | age |
| tg_pg | 2090 | -0.003 | -0.011 | 0.005 | 0.47450 | Cases | age |
| totcho | 2165 | -0.022 | -0.030 | -0.013 | 0.00000 | Controls | age |
| totcho | 4255 | -0.013 | -0.019 | -0.007 | 0.00004 | Pooled | age |
| totcho | 2090 | -0.003 | -0.012 | 0.005 | 0.47900 | Cases | age |
| totfa | 4251 | -0.016 | -0.022 | -0.010 | 0.00000 | Pooled | age |
| totfa | 2162 | -0.023 | -0.031 | -0.014 | 0.00000 | Controls | age |
| totfa | 2089 | -0.009 | -0.017 | 0.000 | 0.03828 | Cases | age |
| totpg | 2165 | -0.026 | -0.034 | -0.017 | 0.00000 | Controls | age |
| totpg | 4255 | -0.016 | -0.022 | -0.010 | 0.00000 | Pooled | age |
| totpg | 2090 | -0.005 | -0.014 | 0.003 | 0.24566 | Cases | age |
| tyr | 4246 | 0.011 | 0.005 | 0.017 | 0.00048 | Pooled | age |
| tyr | 2088 | 0.015 | 0.007 | 0.023 | 0.00048 | Cases | age |


| Metabolite | N | Beta | Lower 95\% CI | $\begin{aligned} & \hline \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | $\mathbf{p}-$ <br> value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| tyr | 2158 | 0.007 | -0.002 | 0.015 | 0.13958 | Controls | age |
| unsat | 4251 | 0.006 | 0.000 | 0.012 | 0.03864 | Pooled | age |
| unsat | 2162 | 0.007 | -0.001 | 0.016 | 0.10349 | Controls | age |
| unsat | 2089 | 0.006 | -0.003 | 0.014 | 0.17766 | Cases | age |
| val | 2164 | -0.005 | -0.014 | 0.004 | 0.25222 | Controls | age |
| val | 2092 | 0.004 | -0.005 | 0.012 | 0.36267 | Cases | age |
| val | 4256 | -0.001 | -0.007 | 0.005 | 0.85882 | Pooled | age |
| vldl_c | 4259 | -0.013 | -0.019 | -0.007 | 0.00003 | Pooled | age |
| vldl_c | 2166 | -0.018 | -0.027 | -0.009 | 0.00004 | Controls | age |
| vldl_c | 2093 | -0.008 | -0.016 | 0.001 | 0.07172 | Cases | age |
| vldl_d | 4259 | -0.003 | -0.009 | 0.003 | 0.29005 | Pooled | age |
| vldl_d | 2166 | -0.004 | -0.013 | 0.005 | 0.37627 | Controls | age |
| vldl_d | 2093 | -0.003 | -0.011 | 0.005 | 0.46474 | Cases | age |
| vldl_tg | 4259 | -0.008 | -0.014 | -0.002 | 0.00928 | Pooled | age |
| vldl_tg | 2166 | -0.011 | -0.019 | -0.002 | 0.01559 | Controls | age |
| vldl_tg | 2093 | -0.006 | -0.014 | 0.002 | 0.14114 | Cases | age |
| xl_hdl_c | 2165 | -0.008 | -0.016 | 0.001 | 0.07275 | Controls | age |
| xl_hdl_c | 4256 | -0.003 | -0.009 | 0.003 | 0.37346 | Pooled | age |
| xl_hdl_c | 2091 | 0.003 | -0.005 | 0.011 | 0.48525 | Cases | age |
| xl_hdl_ce | 2165 | -0.009 | -0.017 | 0.000 | 0.05138 | Controls | age |
| xl_hdl_ce | 4256 | -0.003 | -0.010 | 0.003 | 0.26017 | Pooled | age |
| xl_hdl_ce | 2091 | 0.002 | -0.006 | 0.011 | 0.62091 | Cases | age |


| Metabolite | N | Beta | Lower 95\% CI | $\begin{aligned} & \hline \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | $\mathbf{p}-$ <br> value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_hdl_fc | 2165 | -0.006 | -0.014 | 0.003 | 0.20291 | Controls | age |
| xl_hdl_fc | 2091 | 0.005 | -0.003 | 0.014 | 0.23299 | Cases | age |
| xl_hdl_fc | 4256 | 0.000 | -0.006 | 0.006 | 0.89581 | Pooled | age |
| xl_hdl_1 | 2091 | 0.007 | -0.002 | 0.015 | 0.10776 | Cases | age |
| xl_hdl_1 | 2165 | -0.004 | -0.013 | 0.004 | 0.34074 | Controls | age |
| xl_hdl_l | 4256 | 0.001 | -0.005 | 0.007 | 0.74605 | Pooled | age |
| xl_hdl_p | 2091 | 0.007 | -0.001 | 0.016 | 0.09054 | Cases | age |
| xl_hdl_p | 2165 | -0.003 | -0.012 | 0.005 | 0.44389 | Controls | age |
| xl_hdl_p | 4256 | 0.002 | -0.004 | 0.008 | 0.55028 | Pooled | age |
| xl_hdl_pl | 2091 | 0.009 | 0.001 | 0.018 | 0.02913 | Cases | age |
| xl_hdl_pl | 4256 | 0.005 | -0.001 | 0.011 | 0.11588 | Pooled | age |
| xl_hdl_pl | 2165 | 0.001 | -0.008 | 0.009 | 0.87787 | Controls | age |
| xl_hdl_tg | 2161 | -0.008 | -0.016 | 0.001 | 0.08989 | Controls | age |
| xl_hdl_tg | 2088 | 0.007 | -0.002 | 0.015 | 0.11105 | Cases | age |
| xl_hdl_tg | 4249 | -0.001 | -0.007 | 0.006 | 0.86608 | Pooled | age |
| xl_vldl_c | 2158 | -0.012 | -0.021 | -0.003 | 0.00894 | Controls | age |
| xl_vldl_c | 4247 | -0.008 | -0.014 | -0.002 | 0.01048 | Pooled | age |
| xl_vldl_c | 2089 | -0.004 | -0.012 | 0.004 | 0.32945 | Cases | age |
| xl_vldl_ce | 2160 | -0.011 | -0.020 | -0.002 | 0.01242 | Controls | age |
| xl_vldl_ce | 4248 | -0.007 | -0.013 | -0.001 | 0.01604 | Pooled | age |
| xl_vldl_ce | 2088 | -0.004 | -0.012 | 0.004 | 0.37961 | Cases | age |
| xl_vldl_fc | 4244 | -0.009 | -0.015 | -0.003 | 0.00419 | Pooled | age |


| Metabolite | N | Beta | Lower 95\% CI | $\begin{aligned} & \hline \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | $\mathbf{p}-$ <br> value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xl_vldl_fc | 2156 | -0.013 | -0.022 | -0.004 | 0.00489 | Controls | age |
| xl_vldl_fc | 2088 | -0.005 | -0.013 | 0.003 | 0.22837 | Cases | age |
| xl_vldl_1 | 4243 | -0.008 | -0.014 | -0.002 | 0.01269 | Pooled | age |
| xl_vldl_1 | 2155 | -0.011 | -0.019 | -0.002 | 0.02127 | Controls | age |
| xl_vldl_1 | 2088 | -0.005 | -0.013 | 0.003 | 0.21857 | Cases | age |
| xl_vldl_p | 4244 | -0.007 | -0.014 | -0.001 | 0.01538 | Pooled | age |
| xl_vldl_p | 2155 | -0.011 | -0.020 | -0.002 | 0.02033 | Controls | age |
| xl_vldl_p | 2089 | -0.005 | -0.013 | 0.004 | 0.26777 | Cases | age |
| xl_vldl_pl | 4241 | -0.009 | -0.015 | -0.003 | 0.00343 | Pooled | age |
| xl_vldl_pl | 2154 | -0.012 | -0.021 | -0.003 | 0.00770 | Controls | age |
| xl_vldl_pl | 2087 | -0.006 | -0.014 | 0.002 | 0.13868 | Cases | age |
| xl_vldl_tg | 4243 | -0.007 | -0.013 | -0.001 | 0.01656 | Pooled | age |
| xl_vldl_tg | 2155 | -0.010 | -0.019 | -0.001 | 0.02887 | Controls | age |
| xl_vldl_tg | 2088 | -0.005 | -0.013 | 0.003 | 0.22235 | Cases | age |
| xs_vldl_c | 4259 | -0.012 | -0.018 | -0.006 | 0.00013 | Pooled | age |
| xs_vldl_c | 2166 | -0.016 | -0.025 | -0.008 | 0.00021 | Controls | age |
| xs_vldl_c | 2093 | -0.007 | -0.016 | 0.002 | 0.10704 | Cases | age |
| xs_vldl_ce | 4259 | -0.012 | -0.018 | -0.006 | 0.00011 | Pooled | age |
| xs_vldl_ce | 2166 | -0.016 | -0.025 | -0.008 | 0.00018 | Controls | age |
| xs_vldl_ce | 2093 | -0.007 | -0.016 | 0.001 | 0.10048 | Cases | age |
| xs_vldl_fc | 4259 | -0.011 | -0.017 | -0.005 | 0.00059 | Pooled | age |
| xs_vldl_fc | 2167 | -0.014 | -0.023 | -0.006 | 0.00108 | Controls | age |


| Metabolite | N | Beta | Lower $95 \% \text { CI }$ | Upper $95 \% \text { CI }$ | $\begin{array}{\|l} \mathrm{p}- \\ \text { value } \end{array}$ | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xs_vldl_fc | 2092 | -0.006 | -0.015 | 0.002 | 0.13514 | Cases | age |
| xs_vldl_1 | 4259 | -0.012 | -0.018 | -0.006 | 0.00008 | Pooled | age |
| xs_vldl_1 | 2167 | -0.017 | -0.025 | -0.008 | 0.00012 | Controls | age |
| xs_vldl_1 | 2092 | -0.007 | -0.016 | 0.001 | 0.10295 | Cases | age |
| xs_vldl_p | 4259 | -0.012 | -0.018 | -0.006 | 0.00016 | Pooled | age |
| xs_vldl_p | 2167 | -0.016 | -0.025 | -0.008 | 0.00023 | Controls | age |
| xs_vldl_p | 2092 | -0.007 | -0.015 | 0.002 | 0.11313 | Cases | age |
| xs_vldl_pl | 4260 | -0.013 | -0.019 | -0.007 | 0.00002 | Pooled | age |
| xs_vldl_pl | 2167 | -0.017 | -0.025 | -0.008 | 0.00012 | Controls | age |
| xs_vldl_pl | 2093 | -0.009 | -0.017 | -0.001 | 0.03732 | Cases | age |
| xs_vldl_tg | 2167 | -0.009 | -0.018 | 0.000 | 0.03963 | Controls | age |
| xs_vldl_tg | 4259 | -0.005 | -0.011 | 0.001 | 0.07721 | Pooled | age |
| xs_vldl_tg | 2092 | -0.002 | -0.010 | 0.007 | 0.67524 | Cases | age |
| xxl_vldl_c | 4244 | -0.009 | -0.015 | -0.003 | 0.00464 | Pooled | age |
| xxl_vldl_c | 2157 | -0.013 | -0.022 | -0.004 | 0.00542 | Controls | age |
| xxl_vldl_c | 2087 | -0.005 | -0.013 | 0.003 | 0.23177 | Cases | age |
| xxl_vldl_ce | 2160 | -0.013 | -0.022 | -0.004 | 0.00326 | Controls | age |
| xxl_vldl_ce | 4248 | -0.009 | -0.015 | -0.003 | 0.00506 | Pooled | age |
| xxl_vldl_ce | 2088 | -0.004 | -0.012 | 0.004 | 0.33255 | Cases | age |
| xxl_vldl_fc | 4240 | -0.008 | -0.014 | -0.002 | 0.01110 | Pooled | age |
| xxl_vldl_fc | 2154 | -0.011 | -0.020 | -0.002 | 0.02078 | Controls | age |
| xxl_vldl_fc | 2086 | -0.005 | -0.013 | 0.003 | 0.19336 | Cases | age |


| Metabolite | N |  | Beta |  | Lower $95 \% \text { CI }$ | Upper $95 \% \text { CI }$ | $\mathrm{p}-$ <br> value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xxl_vldl_1 |  | 4238 |  | -0.010 | -0.016 | -0.004 | 0.00145 | Pooled | age |
| xxl_vldl_1 |  | 2153 |  | -0.013 | -0.022 | -0.004 | 0.00462 | Controls | age |
| xxl_vldl_l |  | 2085 |  | -0.007 | -0.015 | 0.001 | 0.09303 | Cases | age |
| xxl_vldl_p |  | 4260 |  | -0.012 | -0.018 | -0.006 | 0.00015 | Pooled | age |
| xxl_vldl_p |  | 2167 |  | -0.016 | -0.025 | -0.007 | 0.00064 | Controls | age |
| xxl_vldl_p |  | 2093 |  | -0.008 | -0.015 | 0.000 | 0.05525 | Cases | age |
| xxl_vldl_pl |  | 4238 |  | -0.010 | -0.016 | -0.004 | 0.00190 | Pooled | age |
| xxl_vldl_pl |  | 2153 |  | -0.013 | -0.022 | -0.004 | 0.00525 | Controls | age |
| xxl_vldl_pl |  | 2085 |  | -0.007 | -0.015 | 0.001 | 0.10758 | Cases | age |
| xxl_vldl_tg |  | 4238 |  | -0.010 | -0.016 | -0.003 | 0.00206 | Pooled | age |
| xxl_vldl_tg |  | 2152 |  | -0.012 | -0.021 | -0.003 | 0.00688 | Controls | age |
| xxl_vldl_tg |  | 2086 |  | -0.007 | -0.015 | 0.001 | 0.09430 | Cases | age |

Appendix D Table D 8: Linear regression results for the associations between individual biomarkers and BMI in cases, controls and pooled cases and controls with follow-up data in the CAP trial (4,260 participants: 2,167 controls, 2,093 cases)

| Metabolite | N | Beta | Lower $95 \% \text { CI }$ | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xxl_vldl_p |  |  | 0 | 0 |  | Cases | bmi |


| Metabolite | N |  | Beta |  | Lower 95\% CI | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xxl_vldl_1 |  | 1421 |  | 0.059 | 0.046 | 0.072 | 0.000 | Cases | bmi |
| xxl_vldl_pl |  | 1421 |  | 0.058 | 0.045 | 0.071 | 0.000 | Cases | bmi |
| xxl_vldl_c |  | 1422 |  | 0.055 | 0.042 | 0.068 | 0.000 | Cases | bmi |
| xxl_vldl_ce |  | 1421 |  | 0.048 | 0.035 | 0.061 | 0.000 | Cases | bmi |
| xxl_vldl_fc |  | 1421 |  | 0.060 | 0.047 | 0.073 | 0.000 | Cases | bmi |
| xxl_vldl_tg |  | 1422 |  | 0.062 | 0.049 | 0.075 | 0.000 | Cases | bmi |
| xl_vldl_p |  | 1422 |  | 0.065 | 0.052 | 0.077 | 0.000 | Cases | bmi |
| xl_vldl_1 |  | 1422 |  | 0.065 | 0.052 | 0.079 | 0.000 | Cases | bmi |
| xl_vldl_pl |  | 1421 |  | 0.060 | 0.047 | 0.073 | 0.000 | Cases | bmi |
| xl_vldl_c |  | 1422 |  | 0.059 | 0.046 | 0.072 | 0.000 | Cases | bmi |
| xl_vldl_ce |  | 1421 |  | 0.059 | 0.046 | 0.072 | 0.000 | Cases | bmi |
| xl_vldl_fc |  | 1422 |  | 0.059 | 0.046 | 0.072 | 0.000 | Cases | bmi |
| xl_vldl_tg |  | 1422 |  | 0.067 | 0.054 | 0.080 | 0.000 | Cases | bmi |
| 1_vldl_p |  | 1422 |  | 0.067 | 0.054 | 0.080 | 0.000 | Cases | bmi |
| 1_vldl_1 |  | 1422 |  | 0.070 | 0.057 | 0.083 | 0.000 | Cases | bmi |
| 1_vldl_pl |  | 1422 |  | 0.067 | 0.054 | 0.080 | 0.000 | Cases | bmi |
| 1_vldl_c |  | 1421 |  | 0.063 | 0.050 | 0.076 | 0.000 | Cases | bmi |
| 1_vldl_ce |  | 1421 |  | 0.059 | 0.046 | 0.073 | 0.000 | Cases | bmi |
| 1_vldl_fc |  | 1422 |  | 0.065 | 0.052 | 0.078 | 0.000 | Cases | bmi |
| 1_vldl_tg |  | 1422 |  | 0.071 | 0.058 | 0.084 | 0.000 | Cases | bmi |
| m_vldl_p |  | 1421 |  | 0.063 | 0.050 | 0.076 | 0.000 | Cases | bmi |
| m_vldl_1 |  | 1421 |  | 0.066 | 0.052 | 0.079 | 0.000 | Cases | bmi |


| Metabolite | N |  | Beta |  | Lower $95 \% \text { CI }$ | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_vldl_pl |  | 1422 |  | 0.062 | 0.048 | 0.075 | 0.000 | Cases | bmi |
| m_vldl_c |  | 1422 |  | 0.052 | 0.038 | 0.066 | 0.000 | Cases | bmi |
| m_vldl_ce |  | 1423 |  | 0.039 | 0.026 | 0.053 | 0.000 | Cases | bmi |
| m_vldl_fc |  | 1421 |  | 0.062 | 0.048 | 0.075 | 0.000 | Cases | bmi |
| m_vldl_tg |  | 1422 |  | 0.069 | 0.056 | 0.082 | 0.000 | Cases | bmi |
| s_vldl_p |  | 1423 |  | 0.047 | 0.033 | 0.061 | 0.000 | Cases | bmi |
| s_vldl_1 |  | 1424 |  | 0.046 | 0.032 | 0.060 | 0.000 | Cases | bmi |
| s_vldl_pl |  | 1424 |  | 0.043 | 0.029 | 0.057 | 0.000 | Cases | bmi |
| s_vldl_c |  | 1424 |  | 0.015 | 0.000 | 0.029 | 0.046 | Cases | bmi |
| s_vldl_ce |  | 1424 |  | 0.000 | -0.014 | 0.014 | 0.986 | Cases | bmi |
| s_vldl_fc |  | 1424 |  | 0.037 | 0.023 | 0.051 | 0.000 | Cases | bmi |
| s_vldl_tg |  | 1422 |  | 0.061 | 0.047 | 0.074 | 0.000 | Cases | bmi |
| xs_vldl_p |  | 1424 |  | -0.005 | -0.019 | 0.010 | 0.520 | Cases | bmi |
| xs_vldl_1 |  | 1424 |  | -0.011 | -0.025 | 0.003 | 0.139 | Cases | bmi |
| xs_vldl_pl |  | 1425 |  | -0.029 | -0.042 | -0.015 | 0.000 | Cases | bmi |
| xs_vldl_c |  | 1425 |  | -0.024 | -0.038 | -0.010 | 0.001 | Cases | bmi |
| xs_vldl_ce |  | 1425 |  | -0.018 | -0.032 | -0.004 | 0.012 | Cases | bmi |
| xs_vldl_fc |  | 1424 |  | -0.035 | -0.049 | -0.021 | 0.000 | Cases | bmi |
| xs_vldl_tg |  | 1424 |  | 0.041 | 0.027 | 0.055 | 0.000 | Cases | bmi |
| idl_p |  | 1425 |  | -0.030 | -0.044 | -0.016 | 0.000 | Cases | bmi |
| idl_1 |  | 1425 |  | -0.034 | -0.048 | -0.020 | 0.000 | Cases | bmi |
| idl_pl |  | 1425 |  | -0.041 | -0.054 | -0.027 | 0.000 | Cases | bmi |


| Metabolite | N |  | Beta |  | Lower $95 \% \text { CI }$ | $\begin{array}{\|l\|} \hline \text { Upper } \\ 95 \% \text { CI } \end{array}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| idl_c |  | 1425 |  | -0.036 | -0.050 | -0.022 | 0.000 | Cases | bmi |
| idl_ce |  | 1425 |  | -0.027 | -0.041 | -0.014 | 0.000 | Cases | bmi |
| idl_fc |  | 1425 |  | -0.053 | -0.066 | -0.040 | 0.000 | Cases | bmi |
| idl_tg |  | 1424 |  | 0.013 | -0.001 | 0.027 | 0.076 | Cases | bmi |
| 1_ldl_p |  | 1425 |  | -0.030 | -0.044 | -0.016 | 0.000 | Cases | bmi |
| 1_ldl_1 |  | 1425 |  | -0.034 | -0.047 | -0.020 | 0.000 | Cases | bmi |
| 1_ldl_pl |  | 1425 |  | -0.032 | -0.046 | -0.018 | 0.000 | Cases | bmi |
| 1_ldl_c |  | 1425 |  | -0.036 | -0.049 | -0.022 | 0.000 | Cases | bmi |
| 1_ldl_ce |  | 1425 |  | -0.031 | -0.045 | -0.017 | 0.000 | Cases | bmi |
| 1_ldl_fc |  | 1425 |  | -0.046 | -0.060 | -0.033 | 0.000 | Cases | bmi |
| 1_ldl_tg |  | 1424 |  | -0.002 | -0.016 | 0.012 | 0.775 | Cases | bmi |
| m_ldl_p |  | 1425 |  | -0.029 | -0.043 | -0.015 | 0.000 | Cases | bmi |
| m_ldl_1 |  | 1425 |  | -0.031 | -0.044 | -0.017 | 0.000 | Cases | bmi |
| m_ldl_pl |  | 1425 |  | -0.014 | -0.028 | 0.000 | 0.049 | Cases | bmi |
| m_ldl_c |  | 1425 |  | -0.035 | -0.049 | -0.022 | 0.000 | Cases | bmi |
| m_ldl_ce |  | 1425 |  | -0.036 | -0.049 | -0.022 | 0.000 | Cases | bmi |
| m_ldl_fc |  | 1425 |  | -0.033 | -0.047 | -0.019 | 0.000 | Cases | bmi |
| m_ldl_tg |  | 1424 |  | -0.008 | -0.022 | 0.007 | 0.289 | Cases | bmi |
| s_ldl_p |  | 1425 |  | -0.029 | -0.043 | -0.015 | 0.000 | Cases | bmi |
| s_ldl_1 |  | 1425 |  | -0.032 | -0.045 | -0.018 | 0.000 | Cases | bmi |
| s_ldl_pl |  | 1425 |  | -0.016 | -0.030 | -0.002 | 0.026 | Cases | bmi |
| s_ldl_c |  | 1425 |  | -0.039 | -0.052 | -0.025 | 0.000 | Cases | bmi |


| Metabolite | N |  | Beta |  | Lower $95 \% \text { CI }$ | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_ldl_ce |  | 1425 |  | -0.039 | -0.052 | -0.025 | 0.000 | Cases | bmi |
| s_ldl_fc |  | 1425 |  | -0.036 | -0.050 | -0.023 | 0.000 | Cases | bmi |
| s_ldl_tg |  | 1424 |  | 0.021 | 0.007 | 0.035 | 0.003 | Cases | bmi |
| xl_hdl_p |  | 1423 |  | -0.075 | -0.089 | -0.062 | 0.000 | Cases | bmi |
| xl_hdl_1 |  | 1423 |  | -0.076 | -0.089 | -0.062 | 0.000 | Cases | bmi |
| xl_hdl_pl |  | 1423 |  | -0.082 | -0.096 | -0.068 | 0.000 | Cases | bmi |
| xl_hdl_c |  | 1423 |  | -0.065 | -0.079 | -0.051 | 0.000 | Cases | bmi |
| xl_hdl_ce |  | 1423 |  | -0.063 | -0.077 | -0.049 | 0.000 | Cases | bmi |
| xl_hdl_fc |  | 1423 |  | -0.068 | -0.082 | -0.054 | 0.000 | Cases | bmi |
| xl_hdl_tg |  | 1422 |  | 0.015 | 0.001 | 0.028 | 0.038 | Cases | bmi |
| 1_hdl_p |  | 1424 |  | -0.078 | -0.092 | -0.064 | 0.000 | Cases | bmi |
| 1_hdl_1 |  | 1424 |  | -0.079 | -0.092 | -0.065 | 0.000 | Cases | bmi |
| l_hdl_pl |  | 1425 |  | -0.078 | -0.091 | -0.064 | 0.000 | Cases | bmi |
| 1_hdl_c |  | 1424 |  | -0.079 | -0.092 | -0.065 | 0.000 | Cases | bmi |
| 1_hdl_ce |  | 1424 |  | -0.077 | -0.091 | -0.063 | 0.000 | Cases | bmi |
| 1_hdl_fc |  | 1424 |  | -0.083 | -0.096 | -0.069 | 0.000 | Cases | bmi |
| 1_hdl_tg |  | 1423 |  | -0.046 | -0.061 | -0.032 | 0.000 | Cases | bmi |
| m_hdl_p |  | 1425 |  | -0.028 | -0.041 | -0.014 | 0.000 | Cases | bmi |
| m_hdl_1 |  | 1425 |  | -0.030 | -0.044 | -0.017 | 0.000 | Cases | bmi |
| m_hdl_pl |  | 1425 |  | -0.028 | -0.042 | -0.015 | 0.000 | Cases | bmi |
| m_hdl_c |  | 1425 |  | -0.039 | -0.052 | -0.025 | 0.000 | Cases | bmi |
| m_hdl_ce |  | 1425 |  | -0.037 | -0.050 | -0.023 | 0.000 | Cases | bmi |


| Metabolite | N |  | Beta |  | Lower $95 \% \text { CI }$ | $\begin{array}{\|l\|} \hline \text { Upper } \\ 95 \% \text { CI } \end{array}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_hdl_fc |  | 1425 |  | -0.045 | -0.058 | -0.031 | 0.000 | Cases | bmi |
| m_hdl_tg |  | 1424 |  | 0.049 | 0.036 | 0.063 | 0.000 | Cases | bmi |
| s_hdl_p |  | 1425 |  | 0.023 | 0.009 | 0.037 | 0.002 | Cases | bmi |
| s_hdl_1 |  | 1425 |  | 0.016 | 0.002 | 0.031 | 0.021 | Cases | bmi |
| s_hdl_pl |  | 1425 |  | 0.039 | 0.025 | 0.052 | 0.000 | Cases | bmi |
| s_hdl_c |  | 1422 |  | -0.034 | -0.047 | -0.021 | 0.000 | Cases | bmi |
| s_hdl_ce |  | 1422 |  | -0.039 | -0.052 | -0.025 | 0.000 | Cases | bmi |
| s_hdl_fc |  | 1425 |  | 0.019 | 0.006 | 0.033 | 0.006 | Cases | bmi |
| s_hdl_tg |  | 1423 |  | 0.072 | 0.058 | 0.085 | 0.000 | Cases | bmi |
| vldl_d |  | 1425 |  | 0.082 | 0.068 | 0.095 | 0.000 | Cases | bmi |
| ldl_d |  | 1425 |  | -0.007 | -0.021 | 0.006 | 0.290 | Cases | bmi |
| hdl_d |  | 1425 |  | -0.086 | -0.100 | -0.072 | 0.000 | Cases | bmi |
| serum_c |  | 1425 |  | -0.043 | -0.057 | -0.029 | 0.000 | Cases | bmi |
| vldl_c |  | 1425 |  | 0.036 | 0.022 | 0.050 | 0.000 | Cases | bmi |
| remnant_c |  | 1425 |  | 0.007 | -0.007 | 0.021 | 0.326 | Cases | bmi |
| ldl_c |  | 1425 |  | -0.037 | -0.051 | -0.023 | 0.000 | Cases | bmi |
| hdl_c |  | 1425 |  | -0.075 | -0.088 | -0.061 | 0.000 | Cases | bmi |
| hdl2_c |  | 1425 |  | -0.074 | -0.088 | -0.061 | 0.000 | Cases | bmi |
| hdl3_c |  | 1425 |  | -0.062 | -0.075 | -0.048 | 0.000 | Cases | bmi |
| estc |  | 1424 |  | -0.045 | -0.058 | -0.031 | 0.000 | Cases | bmi |
| freec |  | 1424 |  | -0.039 | -0.053 | -0.026 | 0.000 | Cases | bmi |
| serum_tg |  | 1425 |  | 0.063 | 0.050 | 0.077 | 0.000 | Cases | bmi |


| Metabolite | N | Beta | Lower 95\% CI | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| vldl_tg | 1425 | 0.068 | 0.055 | 0.082 | 0.000 | Cases | bmi |
| 1dl_tg | 1425 | 0.002 | -0.013 | 0.016 | 0.802 | Cases | bmi |
| hdl_tg | 1425 | 0.035 | 0.021 | 0.049 | 0.000 | Cases | bmi |
| totpg | 1424 | -0.022 | -0.036 | -0.008 | 0.002 | Cases | bmi |
| tg_pg | 1424 | 0.076 | 0.063 | 0.089 | 0.000 | Cases | bmi |
| pc | 1424 | -0.031 | -0.045 | -0.017 | 0.000 | Cases | bmi |
| sm | 1424 | -0.040 | -0.054 | -0.026 | 0.000 | Cases | bmi |
| totcho | 1424 | -0.033 | -0.047 | -0.019 | 0.000 | Cases | bmi |
| apoa1 | 1425 | -0.065 | -0.078 | -0.051 | 0.000 | Cases | bmi |
| apob | 1424 | 0.014 | 0.000 | 0.028 | 0.055 | Cases | bmi |
| apob_apoa1 | 1424 | 0.041 | 0.027 | 0.055 | 0.000 | Cases | bmi |
| totfa | 1423 | 0.022 | 0.008 | 0.036 | 0.002 | Cases | bmi |
| unsat | 1423 | -0.070 | -0.083 | -0.056 | 0.000 | Cases | bmi |
| dha | 1423 | -0.007 | -0.021 | 0.007 | 0.323 | Cases | bmi |
| la | 1423 | -0.015 | -0.030 | -0.001 | 0.035 | Cases | bmi |
| faw3 | 1423 | 0.004 | -0.010 | 0.019 | 0.548 | Cases | bmi |
| faw6 | 1423 | -0.019 | -0.033 | -0.004 | 0.010 | Cases | bmi |
| pufa | 1423 | -0.015 | -0.029 | -0.001 | 0.037 | Cases | bmi |
| mufa | 1423 | 0.041 | 0.027 | 0.055 | 0.000 | Cases | bmi |
| sfa | 1423 | 0.030 | 0.016 | 0.044 | 0.000 | Cases | bmi |
| dha_fa | 1423 | -0.030 | -0.044 | -0.016 | 0.000 | Cases | bmi |
| la_fa | 1423 | -0.065 | -0.079 | -0.052 | 0.000 | Cases | bmi |


| Metabolite | N |  | Beta |  | Lower $95 \% \text { CI }$ | $\begin{array}{\|l\|} \hline \text { Upper } \\ 95 \% \text { CI } \end{array}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| faw3_fa |  | 1423 |  | -0.017 | -0.031 | -0.003 | 0.015 | Cases | bmi |
| faw6_fa |  | 1423 |  | -0.078 | -0.091 | -0.065 | 0.000 | Cases | bmi |
| pufa_fa |  | 1423 |  | -0.077 | -0.090 | -0.064 | 0.000 | Cases | bmi |
| mufa_fa |  | 1423 |  | 0.073 | 0.059 | 0.087 | 0.000 | Cases | bmi |
| sfa_fa |  | 1423 |  | 0.048 | 0.033 | 0.062 | 0.000 | Cases | bmi |
| glc |  | 1421 |  | 0.055 | 0.042 | 0.067 | 0.000 | Cases | bmi |
| lac |  | 1425 |  | 0.014 | 0.002 | 0.026 | 0.023 | Cases | bmi |
| pyr |  | 1423 |  | 0.033 | 0.022 | 0.044 | 0.000 | Cases | bmi |
| cit |  | 1424 |  | -0.011 | -0.025 | 0.004 | 0.160 | Cases | bmi |
| ala |  | 1425 |  | 0.027 | 0.012 | 0.041 | 0.000 | Cases | bmi |
| gln |  | 1424 |  | -0.057 | -0.070 | -0.043 | 0.000 | Cases | bmi |
| gly |  | 1416 |  | -0.041 | -0.054 | -0.027 | 0.000 | Cases | bmi |
| his |  | 1421 |  | 0.004 | -0.010 | 0.019 | 0.532 | Cases | bmi |
| ile |  | 1424 |  | 0.081 | 0.067 | 0.094 | 0.000 | Cases | bmi |
| leu |  | 1425 |  | 0.078 | 0.064 | 0.091 | 0.000 | Cases | bmi |
| val |  | 1425 |  | 0.081 | 0.067 | 0.095 | 0.000 | Cases | bmi |
| phe |  | 1425 |  | 0.061 | 0.047 | 0.075 | 0.000 | Cases | bmi |
| tyr |  | 1421 |  | 0.061 | 0.048 | 0.075 | 0.000 | Cases | bmi |
| ace |  | 1425 |  | 0.003 | -0.011 | 0.016 | 0.706 | Cases | bmi |
| acace |  | 1425 |  | 0.037 | 0.023 | 0.052 | 0.000 | Cases | bmi |
| bohbut |  | 1387 |  | 0.003 | -0.011 | 0.018 | 0.662 | Cases | bmi |
| crea |  | 1421 |  | 0.007 | -0.006 | 0.021 | 0.284 | Cases | bmi |


| Metabolite | N | Beta | Lower 95\% CI | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| alb | 1425 | -0.007 | -0.022 | 0.007 | 0.334 | Cases | bmi |
| gp | 1425 | 0.060 | 0.047 | 0.074 | 0.000 | Cases | bmi |
| xxl_vldl_p | 1395 | 0.059 | 0.046 | 0.073 | 0.000 | Controls | bmi |
| xxl_vldl_1 | 1390 | 0.064 | 0.050 | 0.079 | 0.000 | Controls | bmi |
| xxl_vldl_pl | 1390 | 0.064 | 0.050 | 0.079 | 0.000 | Controls | bmi |
| xxl_vldl_c | 1392 | 0.059 | 0.044 | 0.073 | 0.000 | Controls | bmi |
| xxl_vldl_ce | 1393 | 0.051 | 0.037 | 0.065 | 0.000 | Controls | bmi |
| xxl_vldl_fc | 1391 | 0.065 | 0.051 | 0.080 | 0.000 | Controls | bmi |
| xxl_vldl_tg | 1390 | 0.065 | 0.050 | 0.079 | 0.000 | Controls | bmi |
| xl_vldl_p | 1391 | 0.068 | 0.054 | 0.082 | 0.000 | Controls | bmi |
| xl_vldl_1 | 1391 | 0.069 | 0.055 | 0.084 | 0.000 | Controls | bmi |
| xl_vldl_pl | 1391 | 0.066 | 0.052 | 0.080 | 0.000 | Controls | bmi |
| xl_vldl_c | 1392 | 0.062 | 0.048 | 0.076 | 0.000 | Controls | bmi |
| xl_vldl_ce | 1393 | 0.061 | 0.047 | 0.075 | 0.000 | Controls | bmi |
| xl_vldl_fc | 1392 | 0.063 | 0.048 | 0.077 | 0.000 | Controls | bmi |
| xl_vldl_tg | 1391 | 0.071 | 0.057 | 0.085 | 0.000 | Controls | bmi |
| 1_vldl_p | 1393 | 0.070 | 0.056 | 0.084 | 0.000 | Controls | bmi |
| 1_vldl_1 | 1393 | 0.074 | 0.060 | 0.088 | 0.000 | Controls | bmi |
| 1_vldl_pl | 1393 | 0.071 | 0.057 | 0.085 | 0.000 | Controls | bmi |
| 1_vldl_c | 1393 | 0.068 | 0.054 | 0.082 | 0.000 | Controls | bmi |
| 1_vldl_ce | 1393 | 0.064 | 0.050 | 0.078 | 0.000 | Controls | bmi |
| l_vldl_fc | 1392 | 0.070 | 0.056 | 0.084 | 0.000 | Controls | bmi |


| Metabolite | N |  | Beta |  | Lower 95\% CI | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| l_vldl_tg |  | 1393 |  | 0.074 | 0.060 | 0.088 | 0.000 | Controls | bmi |
| m_vldl_p |  | 1393 |  | 0.068 | 0.054 | 0.082 | 0.000 | Controls | bmi |
| m_vldl_1 |  | 1393 |  | 0.070 | 0.056 | 0.084 | 0.000 | Controls | bmi |
| m_vldl_pl |  | 1393 |  | 0.067 | 0.053 | 0.081 | 0.000 | Controls | bmi |
| m_vldl_c |  | 1394 |  | 0.059 | 0.045 | 0.073 | 0.000 | Controls | bmi |
| m_vldl_ce |  | 1394 |  | 0.047 | 0.033 | 0.061 | 0.000 | Controls | bmi |
| m_vldl_fc |  | 1394 |  | 0.068 | 0.054 | 0.082 | 0.000 | Controls | bmi |
| m_vldl_tg |  | 1393 |  | 0.073 | 0.059 | 0.087 | 0.000 | Controls | bmi |
| s_vldl_p |  | 1395 |  | 0.057 | 0.043 | 0.071 | 0.000 | Controls | bmi |
| s_vldl_1 |  | 1395 |  | 0.056 | 0.042 | 0.070 | 0.000 | Controls | bmi |
| s_vldl_pl |  | 1395 |  | 0.055 | 0.041 | 0.069 | 0.000 | Controls | bmi |
| s_vldl_c |  | 1395 |  | 0.025 | 0.011 | 0.039 | 0.001 | Controls | bmi |
| s_vldl_ce |  | 1395 |  | 0.008 | -0.006 | 0.022 | 0.252 | Controls | bmi |
| s_vldl_fc |  | 1395 |  | 0.049 | 0.035 | 0.063 | 0.000 | Controls | bmi |
| s_vldl_tg |  | 1394 |  | 0.069 | 0.055 | 0.083 | 0.000 | Controls | bmi |
| xs_vldl_p |  | 1395 |  | 0.006 | -0.008 | 0.020 | 0.377 | Controls | bmi |
| xs_vldl_1 |  | 1395 |  | -0.001 | -0.015 | 0.013 | 0.906 | Controls | bmi |
| xs_vldl_pl |  | 1395 |  | -0.019 | -0.033 | -0.005 | 0.007 | Controls | bmi |
| xs_vldl_c |  | 1395 |  | -0.018 | -0.032 | -0.004 | 0.012 | Controls | bmi |
| xs_vldl_ce |  | 1395 |  | -0.013 | -0.027 | 0.001 | 0.069 | Controls | bmi |
| xs_vldl_fc |  | 1395 |  | -0.026 | -0.040 | -0.012 | 0.000 | Controls | bmi |
| xs_vldl_tg |  | 1395 |  | 0.054 | 0.040 | 0.068 | 0.000 | Controls | bmi |


| Metabolite | N |  | Beta | Lower $95 \% \text { CI }$ | $\begin{array}{\|l\|} \hline \text { Upper } \\ 95 \% \text { CI } \end{array}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| idl_p |  | 1394 | -0.024 | -0.038 | -0.010 | 0.001 | Controls | bmi |
| idl_1 |  | 1394 | -0.031 | -0.045 | -0.017 | 0.000 | Controls | bmi |
| idl_pl |  | 1395 | -0.037 | -0.051 | -0.023 | 0.000 | Controls | bmi |
| idl_c |  | 1395 | -0.034 | -0.048 | -0.020 | 0.000 | Controls | bmi |
| idl_ce |  | 1395 | -0.026 | -0.039 | -0.012 | 0.000 | Controls | bmi |
| idl_fc |  | 1395 | -0.050 | -0.064 | -0.036 | 0.000 | Controls | bmi |
| idl_tg |  | 1393 | 0.028 | 0.014 | 0.042 | 0.000 | Controls | bmi |
| 1_ldl_p |  | 1395 | -0.026 | -0.040 | -0.012 | 0.000 | Controls | bmi |
| 1_ldl_1 |  | 1395 | -0.031 | -0.046 | -0.017 | 0.000 | Controls | bmi |
| 1_ldl_pl |  | 1395 | -0.029 | -0.043 | -0.015 | 0.000 | Controls | bmi |
| 1_ldl_c |  | 1395 | -0.034 | -0.048 | -0.020 | 0.000 | Controls | bmi |
| 1_ldl_ce |  | 1395 | -0.029 | -0.043 | -0.015 | 0.000 | Controls | bmi |
| 1_ldl_fc |  | 1395 | -0.045 | -0.059 | -0.031 | 0.000 | Controls | bmi |
| 1_ldl_tg |  | 1394 | 0.012 | -0.002 | 0.026 | 0.104 | Controls | bmi |
| m_ldl_p |  | 1395 | -0.025 | -0.039 | -0.011 | 0.001 | Controls | bmi |
| m_ldl_1 |  | 1395 | -0.029 | -0.043 | -0.015 | 0.000 | Controls | bmi |
| m_ldl_pl |  | 1395 | -0.011 | -0.025 | 0.003 | 0.130 | Controls | bmi |
| m_ldl_c |  | 1395 | -0.034 | -0.048 | -0.020 | 0.000 | Controls | bmi |
| m_ldl_ce |  | 1395 | -0.034 | -0.048 | -0.020 | 0.000 | Controls | bmi |
| m_ldl_fc |  | 1395 | -0.033 | -0.046 | -0.019 | 0.000 | Controls | bmi |
| m_ldl_tg |  | 1395 | 0.009 | -0.005 | 0.023 | 0.196 | Controls | bmi |
| s_ldl_p |  | 1395 | -0.025 | -0.039 | -0.011 | 0.000 | Controls | bmi |


| Metabolite | N |  | Beta | Lower 95\% CI | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| s_ldl_1 |  | 1395 | -0.030 | -0.044 | -0.016 | 0.000 | Controls | bmi |
| s_ldl_pl |  | 1395 | -0.014 | -0.028 | 0.000 | 0.054 | Controls | bmi |
| s_ldl_c |  | 1395 | -0.038 | -0.052 | -0.024 | 0.000 | Controls | bmi |
| s_ldl_ce |  | 1395 | -0.037 | -0.051 | -0.023 | 0.000 | Controls | bmi |
| s_ldl_fc |  | 1395 | -0.037 | -0.051 | -0.024 | 0.000 | Controls | bmi |
| s_ldl_tg |  | 1395 | 0.035 | 0.021 | 0.049 | 0.000 | Controls | bmi |
| xl_hdl_p |  | 1395 | -0.081 | -0.095 | -0.068 | 0.000 | Controls | bmi |
| xl_hdl_1 |  | 1395 | -0.082 | -0.096 | -0.069 | 0.000 | Controls | bmi |
| xl_hdl_pl |  | 1395 | -0.085 | -0.098 | -0.072 | 0.000 | Controls | bmi |
| xl_hdl_c |  | 1395 | -0.076 | -0.089 | -0.062 | 0.000 | Controls | bmi |
| xl_hdl_ce |  | 1395 | -0.074 | -0.088 | -0.061 | 0.000 | Controls | bmi |
| xl_hdl_fc |  | 1395 | -0.077 | -0.091 | -0.064 | 0.000 | Controls | bmi |
| xl_hdl_tg |  | 1393 | 0.018 | 0.004 | 0.032 | 0.013 | Controls | bmi |
| 1_hdl_p |  | 1395 | -0.076 | -0.089 | -0.063 | 0.000 | Controls | bmi |
| 1_hdl_1 |  | 1395 | -0.077 | -0.091 | -0.064 | 0.000 | Controls | bmi |
| l_hdl_pl |  | 1395 | -0.075 | -0.088 | -0.061 | 0.000 | Controls | bmi |
| 1_hdl_c |  | 1395 | -0.079 | -0.092 | -0.065 | 0.000 | Controls | bmi |
| l_hdl_ce |  | 1395 | -0.077 | -0.090 | -0.064 | 0.000 | Controls | bmi |
| 1_hdl_fc |  | 1395 | -0.083 | -0.096 | -0.070 | 0.000 | Controls | bmi |
| 1_hdl_tg |  | 1394 | -0.039 | -0.053 | -0.025 | 0.000 | Controls | bmi |
| m_hdl_p |  | 1395 | -0.018 | -0.032 | -0.004 | 0.011 | Controls | bmi |
| m_hdl_l |  | 1395 | -0.022 | -0.036 | -0.008 | 0.003 | Controls | bmi |


| Metabolite | N |  | Beta | Lower 95\% CI | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_hdl_pl |  | 1395 | -0.020 | -0.034 | -0.005 | 0.006 | Controls | bmi |
| m_hdl_c |  | 1395 | -0.032 | -0.047 | -0.018 | 0.000 | Controls | bmi |
| m_hdl_ce |  | 1395 | -0.031 | -0.045 | -0.017 | 0.000 | Controls | bmi |
| m_hdl_fc |  | 1395 | -0.037 | -0.051 | -0.023 | 0.000 | Controls | bmi |
| m_hdl_tg |  | 1394 | 0.063 | 0.049 | 0.077 | 0.000 | Controls | bmi |
| s_hdl_p |  | 1395 | 0.030 | 0.016 | 0.045 | 0.000 | Controls | bmi |
| s_hdl_1 |  | 1395 | 0.024 | 0.010 | 0.039 | 0.001 | Controls | bmi |
| s_hdl_pl |  | 1395 | 0.044 | 0.030 | 0.059 | 0.000 | Controls | bmi |
| s_hdl_c |  | 1392 | -0.031 | -0.045 | -0.017 | 0.000 | Controls | bmi |
| s_hdl_ce |  | 1392 | -0.036 | -0.051 | -0.022 | 0.000 | Controls | bmi |
| s_hdl_fc |  | 1395 | 0.022 | 0.007 | 0.036 | 0.003 | Controls | bmi |
| s_hdl_tg |  | 1395 | 0.079 | 0.065 | 0.092 | 0.000 | Controls | bmi |
| vldl_d |  | 1394 | 0.081 | 0.068 | 0.094 | 0.000 | Controls | bmi |
| ldl_d |  | 1394 | 0.000 | -0.014 | 0.015 | 0.958 | Controls | bmi |
| hdl_d |  | 1394 | -0.087 | -0.100 | -0.074 | 0.000 | Controls | bmi |
| serum_c |  | 1394 | -0.040 | -0.054 | -0.026 | 0.000 | Controls | bmi |
| vldl_c |  | 1394 | 0.045 | 0.031 | 0.058 | 0.000 | Controls | bmi |
| remnant_c |  | 1394 | 0.014 | 0.000 | 0.028 | 0.055 | Controls | bmi |
| ldl_c |  | 1394 | -0.037 | -0.051 | -0.022 | 0.000 | Controls | bmi |
| hdl_c |  | 1394 | -0.074 | -0.088 | -0.061 | 0.000 | Controls | bmi |
| hdl2_c |  | 1394 | -0.073 | -0.087 | -0.060 | 0.000 | Controls | bmi |
| hdl3_c |  | 1394 | -0.063 | -0.077 | -0.049 | 0.000 | Controls | bmi |


| Metabolite | N | Beta | Lower $95 \% \text { CI }$ | $\begin{array}{\|l\|} \hline \text { Upper } \\ 95 \% \text { CI } \end{array}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| estc | 1393 | -0.043 | -0.057 | -0.029 | 0.000 | Controls | bmi |
| freec | 1393 | -0.034 | -0.048 | -0.020 | 0.000 | Controls | bmi |
| serum_tg | 1394 | 0.071 | 0.057 | 0.085 | 0.000 | Controls | bmi |
| vldl_tg | 1394 | 0.074 | 0.061 | 0.088 | 0.000 | Controls | bmi |
| ldl_tg | 1394 | 0.017 | 0.004 | 0.031 | 0.014 | Controls | bmi |
| hdl_tg | 1394 | 0.047 | 0.033 | 0.061 | 0.000 | Controls | bmi |
| totpg | 1393 | -0.016 | -0.030 | -0.002 | 0.023 | Controls | bmi |
| tg_pg | 1392 | 0.083 | 0.069 | 0.096 | 0.000 | Controls | bmi |
| pc | 1393 | -0.025 | -0.039 | -0.011 | 0.000 | Controls | bmi |
| sm | 1393 | -0.045 | -0.059 | -0.032 | 0.000 | Controls | bmi |
| totcho | 1393 | -0.030 | -0.044 | -0.016 | 0.000 | Controls | bmi |
| apoa1 | 1394 | -0.060 | -0.074 | -0.046 | 0.000 | Controls | bmi |
| apob | 1394 | 0.020 | 0.006 | 0.034 | 0.005 | Controls | bmi |
| apob_apoa1 | 1394 | 0.045 | 0.031 | 0.059 | 0.000 | Controls | bmi |
| totfa | 1392 | 0.031 | 0.016 | 0.045 | 0.000 | Controls | bmi |
| unsat | 1392 | -0.064 | -0.078 | -0.051 | 0.000 | Controls | bmi |
| dha | 1392 | 0.005 | -0.010 | 0.019 | 0.521 | Controls | bmi |
| la | 1392 | -0.014 | -0.028 | 0.000 | 0.045 | Controls | bmi |
| faw3 | 1392 | 0.014 | 0.000 | 0.029 | 0.049 | Controls | bmi |
| faw6 | 1392 | -0.014 | -0.028 | 0.000 | 0.054 | Controls | bmi |
| pufa | 1392 | -0.009 | -0.023 | 0.005 | 0.206 | Controls | bmi |
| mufa | 1392 | 0.053 | 0.039 | 0.067 | 0.000 | Controls | bmi |


| Metabolite | N | Beta | Lower $95 \% \mathrm{CI}$ | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| sfa | 1392 | 0.036 | 0.022 | 0.050 | 0.000 | Controls | bmi |
| dha_fa | 1392 | -0.024 | -0.038 | -0.010 | 0.001 | Controls | bmi |
| la_fa | 1392 | -0.081 | -0.095 | -0.067 | 0.000 | Controls | bmi |
| faw3_fa | 1392 | -0.012 | -0.027 | 0.002 | 0.086 | Controls | bmi |
| faw6_fa | 1392 | -0.088 | -0.102 | -0.074 | 0.000 | Controls | bmi |
| pufa_fa | 1392 | -0.084 | -0.098 | -0.070 | 0.000 | Controls | bmi |
| mufa_fa | 1392 | 0.087 | 0.074 | 0.101 | 0.000 | Controls | bmi |
| sfa_fa | 1392 | 0.035 | 0.021 | 0.049 | 0.000 | Controls | bmi |
| glc | 1388 | 0.052 | 0.038 | 0.066 | 0.000 | Controls | bmi |
| lac | 1393 | 0.017 | 0.005 | 0.030 | 0.006 | Controls | bmi |
| pyr | 1392 | 0.032 | 0.021 | 0.043 | 0.000 | Controls | bmi |
| cit | 1392 | -0.007 | -0.021 | 0.008 | 0.351 | Controls | bmi |
| ala | 1393 | 0.023 | 0.009 | 0.037 | 0.001 | Controls | bmi |
| $g \ln$ | 1392 | -0.064 | -0.078 | -0.050 | 0.000 | Controls | bmi |
| gly | 1390 | -0.058 | -0.072 | -0.044 | 0.000 | Controls | bmi |
| his | 1388 | -0.015 | -0.029 | 0.000 | 0.046 | Controls | bmi |
| ile | 1393 | 0.074 | 0.060 | 0.088 | 0.000 | Controls | bmi |
| leu | 1393 | 0.070 | 0.056 | 0.084 | 0.000 | Controls | bmi |
| val | 1392 | 0.071 | 0.057 | 0.085 | 0.000 | Controls | bmi |
| phe | 1393 | 0.060 | 0.046 | 0.074 | 0.000 | Controls | bmi |
| tyr | 1388 | 0.055 | 0.041 | 0.069 | 0.000 | Controls | bmi |
| ace | 1393 | -0.005 | -0.018 | 0.008 | 0.488 | Controls | bmi |


| Metabolite | N |  | Beta |  | Lower $95 \% \text { CI }$ | $\begin{aligned} & \hline \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| acace |  | 1393 |  | 0.040 | 0.026 | 0.053 | 0.000 | Controls | bmi |
| bohbut |  | 1357 |  | 0.006 | -0.008 | 0.020 | 0.409 | Controls | bmi |
| crea |  | 1388 |  | 0.022 | 0.008 | 0.036 | 0.002 | Controls | bmi |
| alb |  | 1394 |  | 0.010 | -0.004 | 0.025 | 0.153 | Controls | bmi |
| gp |  | 1393 |  | 0.070 | 0.056 | 0.084 | 0.000 | Controls | bmi |
| xxl_vldl_p |  | 2820 |  | 0.058 | 0.048 | 0.067 | 0.000 | Pooled | bmi |
| xxl_vldl_1 |  | 2811 |  | 0.062 | 0.052 | 0.072 | 0.000 | Pooled | bmi |
| xxl_vldl_pl |  | 2811 |  | 0.061 | 0.052 | 0.071 | 0.000 | Pooled | bmi |
| xxl_vldl_c |  | 2814 |  | 0.057 | 0.047 | 0.067 | 0.000 | Pooled | bmi |
| xxl_vldl_ce |  | 2814 |  | 0.050 | 0.040 | 0.060 | 0.000 | Pooled | bmi |
| xxl_vldl_fc |  | 2812 |  | 0.063 | 0.053 | 0.073 | 0.000 | Pooled | bmi |
| xxl_vldl_tg |  | 2812 |  | 0.063 | 0.054 | 0.073 | 0.000 | Pooled | bmi |
| xl_vldl_p |  | 2813 |  | 0.067 | 0.057 | 0.076 | 0.000 | Pooled | bmi |
| xl_vldl_1 |  | 2813 |  | 0.068 | 0.058 | 0.077 | 0.000 | Pooled | bmi |
| xl_vldl_pl |  | 2812 |  | 0.063 | 0.054 | 0.073 | 0.000 | Pooled | bmi |
| xl_vldl_c |  | 2814 |  | 0.061 | 0.051 | 0.071 | 0.000 | Pooled | bmi |
| xl_vldl_ce |  | 2814 |  | 0.060 | 0.051 | 0.070 | 0.000 | Pooled | bmi |
| xl_vldl_fc |  | 2814 |  | 0.061 | 0.051 | 0.071 | 0.000 | Pooled | bmi |
| xl_vldl_tg |  | 2813 |  | 0.069 | 0.060 | 0.079 | 0.000 | Pooled | bmi |
| 1_vldl_p |  | 2815 |  | 0.069 | 0.059 | 0.078 | 0.000 | Pooled | bmi |
| 1_vldl_1 |  | 2815 |  | 0.072 | 0.063 | 0.082 | 0.000 | Pooled | bmi |
| 1_vldl_pl |  | 2815 |  | 0.069 | 0.060 | 0.079 | 0.000 | Pooled | bmi |


| Metabolite | N |  | Beta |  | Lower 95\% CI | $\begin{aligned} & \hline \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1_vldl_c |  | 2814 |  | 0.066 | 0.056 | 0.075 | 0.000 | Pooled | bmi |
| 1_vldl_ce |  | 2814 |  | 0.062 | 0.052 | 0.072 | 0.000 | Pooled | bmi |
| 1_vldl_fc |  | 2814 |  | 0.068 | 0.058 | 0.077 | 0.000 | Pooled | bmi |
| 1_vldl_tg |  | 2815 |  | 0.073 | 0.063 | 0.082 | 0.000 | Pooled | bmi |
| m_vldl_p |  | 2814 |  | 0.066 | 0.056 | 0.075 | 0.000 | Pooled | bmi |
| m_vldl_1 |  | 2814 |  | 0.068 | 0.058 | 0.078 | 0.000 | Pooled | bmi |
| m_vldl_pl |  | 2815 |  | 0.065 | 0.055 | 0.074 | 0.000 | Pooled | bmi |
| m_vldl_c |  | 2816 |  | 0.056 | 0.046 | 0.065 | 0.000 | Pooled | bmi |
| m_vldl_ce |  | 2817 |  | 0.043 | 0.033 | 0.053 | 0.000 | Pooled | bmi |
| m_vldl_fc |  | 2815 |  | 0.065 | 0.056 | 0.075 | 0.000 | Pooled | bmi |
| m_vldl_tg |  | 2815 |  | 0.071 | 0.062 | 0.081 | 0.000 | Pooled | bmi |
| s_vldl_p |  | 2818 |  | 0.052 | 0.042 | 0.062 | 0.000 | Pooled | bmi |
| s_vldl_l |  | 2819 |  | 0.051 | 0.041 | 0.061 | 0.000 | Pooled | bmi |
| s_vldl_pl |  | 2819 |  | 0.049 | 0.039 | 0.059 | 0.000 | Pooled | bmi |
| s_vldl_c |  | 2819 |  | 0.020 | 0.010 | 0.030 | 0.000 | Pooled | bmi |
| s_vldl_ce |  | 2819 |  | 0.004 | -0.006 | 0.014 | 0.428 | Pooled | bmi |
| s_vldl_fc |  | 2819 |  | 0.043 | 0.033 | 0.053 | 0.000 | Pooled | bmi |
| s_vldl_tg |  | 2816 |  | 0.065 | 0.056 | 0.075 | 0.000 | Pooled | bmi |
| xs_vldl_p |  | 2819 |  | 0.001 | -0.009 | 0.011 | 0.866 | Pooled | bmi |
| xs_vldl_1 |  | 2819 |  | -0.006 | -0.016 | 0.004 | 0.255 | Pooled | bmi |
| xs_vldl_pl |  | 2820 |  | -0.024 | -0.034 | -0.014 | 0.000 | Pooled | bmi |
| xs_vldl_c |  | 2820 |  | -0.021 | -0.031 | -0.011 | 0.000 | Pooled | bmi |


| Metabolite | N |  | Beta |  | Lower 95\% CI | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| xs_vldl_ce |  | 2820 |  | -0.016 | -0.026 | -0.006 | 0.002 | Pooled | bmi |
| xs_vldl_fc |  | 2819 |  | -0.031 | -0.040 | -0.021 | 0.000 | Pooled | bmi |
| xs_vldl_tg |  | 2819 |  | 0.048 | 0.038 | 0.058 | 0.000 | Pooled | bmi |
| idl_p |  | 2819 |  | -0.027 | -0.037 | -0.017 | 0.000 | Pooled | bmi |
| idl_1 |  | 2819 |  | -0.033 | -0.043 | -0.023 | 0.000 | Pooled | bmi |
| idl_pl |  | 2820 |  | -0.039 | -0.049 | -0.029 | 0.000 | Pooled | bmi |
| idl_c |  | 2820 |  | -0.035 | -0.045 | -0.025 | 0.000 | Pooled | bmi |
| idl_ce |  | 2820 |  | -0.027 | -0.037 | -0.017 | 0.000 | Pooled | bmi |
| idl_fc |  | 2820 |  | -0.052 | -0.061 | -0.042 | 0.000 | Pooled | bmi |
| idl_tg |  | 2817 |  | 0.021 | 0.011 | 0.031 | 0.000 | Pooled | bmi |
| 1_ldl_p |  | 2820 |  | -0.028 | -0.038 | -0.018 | 0.000 | Pooled | bmi |
| 1_ldl_1 |  | 2820 |  | -0.033 | -0.043 | -0.023 | 0.000 | Pooled | bmi |
| 1_ldl_pl |  | 2820 |  | -0.031 | -0.040 | -0.021 | 0.000 | Pooled | bmi |
| 1_ldl_c |  | 2820 |  | -0.035 | -0.045 | -0.025 | 0.000 | Pooled | bmi |
| 1_ldl_ce |  | 2820 |  | -0.031 | -0.040 | -0.021 | 0.000 | Pooled | bmi |
| 1_ldl_fc |  | 2820 |  | -0.046 | -0.055 | -0.036 | 0.000 | Pooled | bmi |
| 1_ldl_tg |  | 2818 |  | 0.005 | -0.005 | 0.015 | 0.344 | Pooled | bmi |
| m_ldl_p |  | 2820 |  | -0.027 | -0.037 | -0.017 | 0.000 | Pooled | bmi |
| m_ldl_1 |  | 2820 |  | -0.030 | -0.040 | -0.020 | 0.000 | Pooled | bmi |
| m_ldl_pl |  | 2820 |  | -0.013 | -0.023 | -0.003 | 0.012 | Pooled | bmi |
| m_ldl_c |  | 2820 |  | -0.035 | -0.045 | -0.025 | 0.000 | Pooled | bmi |
| m_ldl_ce |  | 2820 |  | -0.035 | -0.045 | -0.025 | 0.000 | Pooled | bmi |


| Metabolite | N |  | Beta |  | Lower $95 \% \text { CI }$ | $\begin{array}{\|l\|} \hline \text { Upper } \\ 95 \% \text { CI } \end{array}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| m_ldl_fc |  | 2820 |  | -0.033 | -0.043 | -0.023 | 0.000 | Pooled | bmi |
| m_ldl_tg |  | 2819 |  | 0.001 | -0.009 | 0.011 | 0.860 | Pooled | bmi |
| s_ldl_p |  | 2820 |  | -0.027 | -0.037 | -0.018 | 0.000 | Pooled | bmi |
| s_ldl_1 |  | 2820 |  | -0.031 | -0.041 | -0.021 | 0.000 | Pooled | bmi |
| s_ldl_pl |  | 2820 |  | -0.015 | -0.025 | -0.005 | 0.003 | Pooled | bmi |
| s_ldl_c |  | 2820 |  | -0.038 | -0.048 | -0.029 | 0.000 | Pooled | bmi |
| s_ldl_ce |  | 2820 |  | -0.038 | -0.048 | -0.028 | 0.000 | Pooled | bmi |
| s_ldl_fc |  | 2820 |  | -0.037 | -0.047 | -0.027 | 0.000 | Pooled | bmi |
| s_ldl_tg |  | 2819 |  | 0.029 | 0.019 | 0.038 | 0.000 | Pooled | bmi |
| xl_hdl_p |  | 2818 |  | -0.079 | -0.088 | -0.069 | 0.000 | Pooled | bmi |
| xl_hdl_1 |  | 2818 |  | -0.079 | -0.089 | -0.069 | 0.000 | Pooled | bmi |
| xl_hdl_pl |  | 2818 |  | -0.084 | -0.093 | -0.074 | 0.000 | Pooled | bmi |
| xl_hdl_c |  | 2818 |  | -0.071 | -0.080 | -0.061 | 0.000 | Pooled | bmi |
| xl_hdl_ce |  | 2818 |  | -0.069 | -0.079 | -0.059 | 0.000 | Pooled | bmi |
| xl_hdl_fc |  | 2818 |  | -0.073 | -0.083 | -0.063 | 0.000 | Pooled | bmi |
| xl_hdl_tg |  | 2815 |  | 0.016 | 0.007 | 0.026 | 0.001 | Pooled | bmi |
| 1_hdl_p |  | 2819 |  | -0.077 | -0.087 | -0.068 | 0.000 | Pooled | bmi |
| 1_hdl_1 |  | 2819 |  | -0.078 | -0.088 | -0.069 | 0.000 | Pooled | bmi |
| 1_hdl_pl |  | 2820 |  | -0.076 | -0.086 | -0.067 | 0.000 | Pooled | bmi |
| 1_hdl_c |  | 2819 |  | -0.079 | -0.088 | -0.069 | 0.000 | Pooled | bmi |
| 1_hdl_ce |  | 2819 |  | -0.077 | -0.087 | -0.068 | 0.000 | Pooled | bmi |
| 1_hdl_fc |  | 2819 |  | -0.083 | -0.092 | -0.073 | 0.000 | Pooled | bmi |


| Metabolite | N |  | Beta |  | Lower 95\% CI | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1_hdl_tg |  | 2817 |  | -0.043 | -0.053 | -0.033 | 0.000 | Pooled | bmi |
| m_hdl_p |  | 2820 |  | -0.023 | -0.033 | -0.013 | 0.000 | Pooled | bmi |
| m_hdl_1 |  | 2820 |  | -0.026 | -0.036 | -0.016 | 0.000 | Pooled | bmi |
| m_hdl_pl |  | 2820 |  | -0.024 | -0.034 | -0.014 | 0.000 | Pooled | bmi |
| m_hdl_c |  | 2820 |  | -0.036 | -0.046 | -0.026 | 0.000 | Pooled | bmi |
| m_hdl_ce |  | 2820 |  | -0.034 | -0.044 | -0.024 | 0.000 | Pooled | bmi |
| m_hdl_fc |  | 2820 |  | -0.041 | -0.051 | -0.031 | 0.000 | Pooled | bmi |
| m_hdl_tg |  | 2818 |  | 0.056 | 0.047 | 0.066 | 0.000 | Pooled | bmi |
| s_hdl_p |  | 2820 |  | 0.026 | 0.016 | 0.036 | 0.000 | Pooled | bmi |
| s_hdl_1 |  | 2820 |  | 0.020 | 0.010 | 0.030 | 0.000 | Pooled | bmi |
| s_hdl_pl |  | 2820 |  | 0.041 | 0.032 | 0.051 | 0.000 | Pooled | bmi |
| s_hdl_c |  | 2814 |  | -0.033 | -0.043 | -0.023 | 0.000 | Pooled | bmi |
| s_hdl_ce |  | 2814 |  | -0.038 | -0.048 | -0.028 | 0.000 | Pooled | bmi |
| s_hdl_fc |  | 2820 |  | 0.021 | 0.011 | 0.031 | 0.000 | Pooled | bmi |
| s_hdl_tg |  | 2818 |  | 0.075 | 0.066 | 0.085 | 0.000 | Pooled | bmi |
| vldl_d |  | 2819 |  | 0.082 | 0.072 | 0.091 | 0.000 | Pooled | bmi |
| ldl_d |  | 2819 |  | -0.004 | -0.014 | 0.006 | 0.487 | Pooled | bmi |
| hdl_d |  | 2819 |  | -0.087 | -0.096 | -0.077 | 0.000 | Pooled | bmi |
| serum_c |  | 2819 |  | -0.042 | -0.052 | -0.032 | 0.000 | Pooled | bmi |
| vldl_c |  | 2819 |  | 0.041 | 0.031 | 0.050 | 0.000 | Pooled | bmi |
| remnant_c |  | 2819 |  | 0.010 | 0.000 | 0.020 | 0.041 | Pooled | bmi |
| ldl_c |  | 2819 |  | -0.037 | -0.047 | -0.027 | 0.000 | Pooled | bmi |


| Metabolite | N | Beta | Lower 95\% CI | $\begin{array}{\|l\|} \hline \text { Upper } \\ 95 \% \text { CI } \end{array}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| hdl_c | 2819 | -0.075 | -0.084 | -0.065 | 0.000 | Pooled | bmi |
| hdl2_c | 2819 | -0.074 | -0.084 | -0.064 | 0.000 | Pooled | bmi |
| hdl3_c | 2819 | -0.063 | -0.072 | -0.053 | 0.000 | Pooled | bmi |
| estc | 2817 | -0.044 | -0.054 | -0.034 | 0.000 | Pooled | bmi |
| freec | 2817 | -0.037 | -0.047 | -0.027 | 0.000 | Pooled | bmi |
| serum_tg | 2819 | 0.068 | 0.058 | 0.077 | 0.000 | Pooled | bmi |
| vldl_tg | 2819 | 0.072 | 0.062 | 0.081 | 0.000 | Pooled | bmi |
| ldl_tg | 2819 | 0.010 | 0.000 | 0.020 | 0.056 | Pooled | bmi |
| hdl_tg | 2819 | 0.041 | 0.031 | 0.051 | 0.000 | Pooled | bmi |
| totpg | 2817 | -0.019 | -0.029 | -0.009 | 0.000 | Pooled | bmi |
| tg_pg | 2816 | 0.080 | 0.070 | 0.089 | 0.000 | Pooled | bmi |
| pc | 2817 | -0.028 | -0.038 | -0.018 | 0.000 | Pooled | bmi |
| sm | 2817 | -0.043 | -0.053 | -0.033 | 0.000 | Pooled | bmi |
| totcho | 2817 | -0.032 | -0.041 | -0.022 | 0.000 | Pooled | bmi |
| apoa1 | 2819 | -0.063 | -0.072 | -0.053 | 0.000 | Pooled | bmi |
| apob | 2818 | 0.017 | 0.007 | 0.027 | 0.001 | Pooled | bmi |
| apob_apoa1 | 2818 | 0.043 | 0.033 | 0.053 | 0.000 | Pooled | bmi |
| totfa | 2815 | 0.026 | 0.016 | 0.036 | 0.000 | Pooled | bmi |
| unsat | 2815 | -0.067 | -0.077 | -0.057 | 0.000 | Pooled | bmi |
| dha | 2815 | -0.001 | -0.011 | 0.009 | 0.828 | Pooled | bmi |
| la | 2815 | -0.015 | -0.025 | -0.005 | 0.003 | Pooled | bmi |
| faw3 | 2815 | 0.009 | -0.001 | 0.020 | 0.065 | Pooled | bmi |


| Metabolite | N |  | Beta |  | Lower 95\% CI | $\begin{aligned} & \text { Upper } \\ & 95 \% \text { CI } \end{aligned}$ | p-value | Model | Outcome |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| faw6 |  | 2815 |  | -0.016 | -0.026 | -0.006 | 0.001 | Pooled | bmi |
| pufa |  | 2815 |  | -0.012 | -0.022 | -0.002 | 0.017 | Pooled | bmi |
| mufa |  | 2815 |  | 0.047 | 0.038 | 0.057 | 0.000 | Pooled | bmi |
| sfa |  | 2815 |  | 0.033 | 0.023 | 0.043 | 0.000 | Pooled | bmi |
| dha_fa |  | 2815 |  | -0.027 | -0.037 | -0.017 | 0.000 | Pooled | bmi |
| la_fa |  | 2815 |  | -0.073 | -0.083 | -0.063 | 0.000 | Pooled | bmi |
| faw3_fa |  | 2815 |  | -0.015 | -0.025 | -0.005 | 0.004 | Pooled | bmi |
| faw6_fa |  | 2815 |  | -0.084 | -0.093 | -0.074 | 0.000 | Pooled | bmi |
| pufa_fa |  | 2815 |  | -0.081 | -0.090 | -0.071 | 0.000 | Pooled | bmi |
| mufa_fa |  | 2815 |  | 0.080 | 0.071 | 0.090 | 0.000 | Pooled | bmi |
| sfa_fa |  | 2815 |  | 0.041 | 0.031 | 0.051 | 0.000 | Pooled | bmi |
| glc |  | 2809 |  | 0.053 | 0.044 | 0.063 | 0.000 | Pooled | bmi |
| lac |  | 2818 |  | 0.016 | 0.007 | 0.025 | 0.000 | Pooled | bmi |
| pyr |  | 2815 |  | 0.033 | 0.025 | 0.041 | 0.000 | Pooled | bmi |
| cit |  | 2816 |  | -0.009 | -0.019 | 0.002 | 0.101 | Pooled | bmi |
| ala |  | 2818 |  | 0.025 | 0.015 | 0.035 | 0.000 | Pooled | bmi |
| $g \mathrm{ln}$ |  | 2816 |  | -0.061 | -0.070 | -0.051 | 0.000 | Pooled | bmi |
| gly |  | 2806 |  | -0.050 | -0.059 | -0.040 | 0.000 | Pooled | bmi |
| his |  | 2809 |  | -0.006 | -0.016 | 0.005 | 0.284 | Pooled | bmi |
| ile |  | 2817 |  | 0.078 | 0.068 | 0.087 | 0.000 | Pooled | bmi |
| leu |  | 2818 |  | 0.074 | 0.064 | 0.084 | 0.000 | Pooled | bmi |
| val |  | 2817 |  | 0.076 | 0.066 | 0.086 | 0.000 | Pooled | bmi |


| Metabolite | N | Beta | Lower <br> $\mathbf{9 5 \%} \mathbf{C I}$ | Upper <br> $\mathbf{9 5 \%} \mathbf{C I}$ | p-value | Model | Outcome |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| phe | 2818 | 0.060 | 0.051 | 0.070 | 0.000 | Pooled | bmi |
| tyr | 2809 | 0.058 | 0.048 | 0.068 | 0.000 | Pooled | bmi |
| ace | 2818 | -0.001 | -0.011 | 0.008 | 0.823 | Pooled | bmi |
| acace | 2818 | 0.038 | 0.028 | 0.048 | 0.000 | Pooled | bmi |
| bohbut | 2744 | 0.005 | -0.006 | 0.015 | 0.376 | Pooled | bmi |
| crea | 2809 | 0.015 | 0.005 | 0.025 | 0.002 | Pooled | bmi |
| alb | 2819 | 0.002 | -0.009 | 0.012 | 0.756 | Pooled | bmi |
| gp | 2818 | 0.065 | 0.056 | 0.075 | 0.000 | Pooled | bmi |

Appendix D Figure D 1: Hazard ratios (per 1SD) for the associations between individual biomarkers and all-cause mortality in controls, cases and pooled cases and controls) with follow-up data in the ProtecT trial ( $\mathrm{N}=4,260$ participants ( 2,167 controls and 2,093 cases) with 436 deaths)

## Lipoprotein subclasses



## Lipoprotein particle size

VLDL particles size
LDL particles size
HDL particles size


## Cholesterol



Glycerides and phospholipids


## Apolipoproteins



## Fatty acids



## Fatty acids ratios



Glycolysis related metabolites


## Amino acids



Branched-chain amino acids


Aromatic amino acids


## Ketone bodies



Fluid balance


Appendix D Figure D 2: Beta coefficients (per 1SD) for the associations between individual biomarkers and age in cases, controls and pooled cases and controls with follow-up data in the CAP trial (4,260 participants: 2,167 controls, 2,093 cases)


Appendix D Figure D 3: Beta coefficients for the associations between individual biomarkers and BMI in cases, controls and pooled cases and controls with follow-up data in the CAP trial ( $\mathrm{N}=2,820$ participants: 1,393 controls and 1,425 cases)



[^0]:    - Your contact details
    -Bibliographic details for the item, including a URL
    -An outline nature of the complaint

[^1]:    Credit: Cancer Research UK(2)

[^2]:    $\rightarrow$ UKBB $\rightarrow-$ ProtecT
    $V L D L=V e r y$ low-density lipoprotein; $L D L=$ low-density lipoprotein; $H D L=h i g h$-density lipoprotein; $C=$ Cholesterol; $F A=$ fatty acids;
    DHA = docosahexaenoic acid; MUFA=monounsaturated fatty acids; PUFA=polyunsaturated fatty acids; n3=omega; n6=omega 6

[^3]:    PCa=Prostate Cancer

[^4]:    *Main regression: adjusted for age and centre

