Pre-diagnostic circulating untargeted metabolomics and risk of overall and clinically significant prostate cancer: A systematic review meta-analysis

Harriett Fuller[1*], Orietta P. Agasaro[1,2], Jim M. Guevera[1], Burcu F. Darst[1,2*]

[1] Public Health Sciences, Fred Hutchinson Cancer Center, Seattle, Washington, USA

[2] Department of Epidemiology, University of Washington, Seattle, Washington, USA

*Corresponding Authors

Burcu F. Darst

1100 Fairview Ave N

Seattle, WA, 98109 USA

1-206-667-1036

bdarst@fredhutch.org

Harriett Fuller

1100 Fairview Ave N

Seattle, WA, 98109 USA

1-206-667-6671

hfuller@fredhutch.org

Abstract (200 word max)

Background: Metabolomic dysregulation contributes to prostate cancer (PCa) pathogenesis, and studies suggest that circulating metabolites have strong potential to act as clinical biomarkers. However, evidence of associations between circulating metabolites with overall and clinically significant PCa risk has not been quantitively aggregated. **Methods:** We performed a systematic review and meta-analysis of untargeted pre-diagnostic circulating metabolomic studies across four clinically distinct outcomes: overall, low- to intermediate-risk, high- to very high-risk, and lethal PCa, each compared to controls. **Results:** Twelve studies were identified in the systematic review, and up to 408 metabolites were meta-analyzed across the four PCa outcomes. Three, eleven and nineteen metabolites were significantly associated with risk of overall, high- to very high-risk and lethal PCa, respectively. Metabolites associated with high- to very high-risk PCa were significantly enriched for lipids. Limited evidence of correlation between metabolite effects across outcomes was identified, highlighting potentially unique metabolite drivers of high-risk and lethal PCa. In follow-up analyses, 13 of the significant metabolites were found to be drug and/or dietary modifiable. **Conclusions:** These findings suggest the strong potential for metabolites to inform risk of lethal PCa, which could inform risk-stratified screening strategies and facilitate the identification of targets for PCa prevention.

Introduction

Globally, prostate cancer (PCa) is the fifth leading cause of cancer mortality and second most frequently diagnosed cancer in men[1,2]. Between 2013-2019, in the US, survival rates were 34% for distant PCa compared to 97% for any PCa[3], highlighting the need to improve our knowledge of aggressive PCa risk factors and pathogenesis. One promising avenue of PCa biomarker identification is metabolomics. Due to the increased energy demands of cellular malignancy[4,5], metabolic dysregulation is expected to contribute to PCa pathogenesis. As such, metabolomic investigations could elucidate biological mechanisms that drive tumor development and progression and ultimately improve screening strategies and the identification of lethal PCa. Untargeted metabolomic investigations, which aim to characterize all metabolites in a sample in an unbiased manner, offer the greatest potential for discovery[6-8].

Although metabolomic epidemiology investigations of PCa risk have been conducted, replication of metabolites identified has been challenging due to limited sample sizes and between-study heterogeneity, particularly regarding metabolomic platforms, data processing, and analysis procedures. Accordingly, three previous systematic reviews reported limited reproducibility across circulating metabolomic studies of PCa, but highlighted the potential role of select amino acids and lipids in PCa risk assessment and distinguishing aggressive from non-aggressive PCa at diagnosis[9-11]. However, metabolomic evidence has not been quantitatively aggregated across studies through meta-analyses, limiting the interpretability of findings, ability to assess heterogeneity, and strength of conclusions. Moreover, previous systematic reviews have not compared metabolite findings across clinically distinct PCa outcomes;

identifying pre-diagnostic metabolites that are specific to high-risk or lethal PCa could have notable clinical implications for screening and preventive practices.

We conducted a systematic review and meta-analysis of untargeted pre-diagnostic circulating metabolomic studies to quantitatively evaluate current evidence and robustly identify metabolites associated with risk of overall, low- to intermediate-risk, high- to very high-risk and lethal PCa and to compare risk profiles across these clinically distinct PCa outcomes. Furthermore, we conducted a comprehensive bias assessment to investigate potential causes of heterogeneity present in PCa molecular epidemiological studies.

Results

Study Identification

Upon conducting a systematic review using PubMed and Embase through October 17th, 2024, 1,282 publications were identified for abstract screening (**Figure 1**, **Methods**). After limiting to prospective investigations with ≥50 metabolites measured in pre-diagnostic blood, 12 publications were eligible for inclusion, all of which were nested case-control studies. A total of 15,382 predominantly (>90%) European descent participants, including 7,643 cases and 7,739 controls, were included across 7 cohorts (Alpha-Tocopherol, Beta-Carotene Cancer Prevention [ATBC][12-15], European Prospective Investigation into Cancer and Nutrition [EPIC][16,17], Supplémentation en Vitamines et Minéraux Antioxydants [SU.VI.MAX][18,19], Health Professionals Follow-up Study [HPFS][20], Physicians' Health Study [PHS][20], Northern Sweden Health and Disease Study [NSHDS][21], and Cancer Prevention Study-II Nutrition Cohort [CPS-11][22];

Table 1, **Supplementary Tables 1-2**). Ten studies quantified metabolite levels via MS and two used NMR, with one study using both. The average number of metabolites reported across MS studies was 478 (range 67-1,100) and across NMR studies was 208 (range 159-277).

We evaluated four PCa outcomes, informed by the National Comprehensive Cancer Network (NCCN)[23]: overall PCa, low- to intermediate-risk PCa, high- to very high-risk PCa and lethal PCa (defined in **Methods**), each compared to PCa-free controls (**Supplementary Tables 3**). After harmonizing metabolites (**Methods**), 595 associations were evaluated across the four PCa outcomes, representing 408 unique metabolites (101 amino acids, 8 carbohydrates, 13 cofactors and vitamins, 6 energy-related metabolites, 194 lipids, 20 nucleotides, 13 peptides and 53 xenobiotics) (**Supplementary Figure 1**). Of these associations, 149 (25%) were evaluated in ≥3 studies.

Overall PCa

Nine studies (6,431 cases, 6,674 controls) reported on associations between 727 metabolite biomarkers and overall PCa risk (**Supplementary Table 3**), with 150 metabolites tested in ≥2 studies and therefore meta-analyzed (**Methods**). After applying an FDR (α=0.05) correction for multiple testing, three metabolites, amino acids serotonin and tiglylcarnitine (C5:1-DC) and the lipid sphinganine, were significantly associated with decreased overall PCa risk (**Figure 2, Supplementary Figure 2, Supplementary Table 4**). Nominal association evidence (unadjusted P<0.05) was observed for 25 metabolites (**Figures 2-3**, **Supplementary Table 4**). No metabolite

pathway was significantly enriched among metabolites significantly or nominally associated with overall PCa risk (**Supplementary Table 5**, **Methods**).

Low- to Intermediate-risk PCa

Four studies (3,454 cases, 4,505 controls) reported on associations between 186 metabolite biomarkers and risk of low- to intermediate-risk PCa (**Supplementary Table 3**), with six metabolites (all amino acids) tested in ≥2 studies and therefore meta-analyzed. No metabolite was significantly associated with risk of low- to intermediate-risk PCa (**Figure 2**, **Supplementary Table 6**). One amino acid, 3-methyl histidine, was nominally associated with increased risk of low- to intermediate-risk PCa (**Figures 2-3, Supplementary Table 6**). As all six metabolites were amino acids, enrichment analyses were not conducted.

High- to Very High-risk PCa

Eight studies (1,470 cases, 6,679 controls) reported on associations between 1,001 metabolite biomarkers and risk of high- to very high-risk PCa (**Supplementary Table 3**), with 391 metabolites tested in ≥2 studies and therefore meta-analyzed. Eleven metabolites were significantly associated with decreased risk of high- to very high-risk PCa, including 10 lipids (9 phosphatidylcholines and 1 carnitine (C18:2 carnitine)) and 1 energy metabolite (alpha-ketoglutarate; **Figure 2**, **Supplementary Figure 2**, **Supplementary Table 7**). Fifty-five metabolites had nominal association evidence (**Figures 2-3**, **Supplementary Table 7**). Metabolites with significant and

nominal association evidence were enriched for lipids (P_{adj}=0.03 and P_{adj}=5.71E-09, respectively; **Supplementary Table 5**).

Lethal PCa

Four studies (1,059 cases, 4,999 controls) reported on associations between 806 metabolites and lethal PCa risk (**Supplementary Table 3**), with 48 metabolites tested in ≥2 studies and therefore meta-analyzed. Nineteen metabolites were significantly associated with lethal PCa risk, including positive associations with 14 metabolites (4 amino acids: C-glycosyltryptophan, taurine, glutamate and N-acetylserine; 3 lipids: 3-hydroxybutyrate, glycerol and androstenediol (3beta,17beta) disulfate (2); 3 nucleotides: 2'-O-methyluridine, 5,6-dihydrouridine and pseudouridine; 3 peptides: leucylglycine, gamma-glutamylvaline and gamma-glutamylglycine; and 1 xenobiotic: oleoyl ethanolamide) and inverse associations with 5 metabolites (amino acids cysteine and 4-hydroxyphenylpyruvate, cofactor/vitamin oxalate, phospholipid 1-linoleoyl-GPC and nucleotide dihydroorotate) (**Figure 2**, **Supplementary Figure 2, Supplementary Table 8**). Twenty-four metabolites had nominal association evidence (**Figures 2-3**, **Supplementary Table 8**). No pathway was significantly enriched among metabolites significantly or nominally associated with lethal PCa risk (**Supplementary Table 5**).

Evidence of Effect Heterogeneity

Of the 595 associations tested, 280 (47.1%) had an I^2=0, indicating the absence of between-study heterogeneity. The median I^2 ranged from 0 for lethal PCa to 35.04 for low- to intermediate-risk PCa (**Figure 4**), with heterogeneity estimates largely

comparable across metabolite pathways (**Supplementary Table 9**). In total, 94 (15.8%) associations had a Q-test P-value<0.05 and an $I^2 \geq 40\%$, indicating between-study heterogeneity, with a lack of heterogeneity evidence predominantly observed for significantly (28/33, 84.85%) and nominally (81/105, 77.14%) associated metabolites (**Figure 4**).

Of the 150 metabolites evaluated with overall PCa, 28 were quantified with both MS and NMR, and four demonstrated between-platform heterogeneity (Q-test P-value<0.05). Of the 391 metabolites evaluated with high- to very high-risk PCa, 23 were quantified with NMR and MS, and one demonstrated between-platform heterogeneity. None of these five metabolites were associated with PCa outcomes **(Supplementary Table 10**). All metabolites evaluated with low- to intermediate-risk and lethal PCa were quantified with MS.

Comparison of Evidence Between PCa Outcomes

No metabolite was significantly associated with >1 outcome; however, 13 metabolites were nominally associated with >1 outcome (**Figure 4**), all of which had consistent effect directions across outcomes. We found limited evidence of metabolite effects being correlated between outcomes, except for 11 metabolites nominally associated with both overall and high- to very high-risk PCa ($R=0.91$, $P<1.1\times10^{-4}$; **Figure 5**). Low- to intermediate-risk PCa was not considered in correlations, as only 6 metabolites were meta-analyzed for this outcome.

Risk of Bias (ROB)

We assessed risk of bias (ROB) in our systematic review using a published tool that evaluates the quality of each included study based on six bias domains[24] (**Methods**, **Supplementary Table 11**). Four of the 12 studies had low reporting quality (i.e., ROB score≤3), with bias most notably identified in the statistical analysis domain, primarily due to the selective reporting of significant findings in four studies (**Supplementary Table 12**). Bias in the exposure assessment domain, observed to some degree in six studies, was largely due to four studies not requiring fasting samples and one study reporting on m/z values rather than metabolite names, limiting the ability to interpret findings. Bias in the outcome domain occurred in six studies due to reporting the range/threshold of times between sample collection and PCa diagnosis rather than the median or mean. Additionally, seven studies were limited to participants that may not represent the general population, such as smokers or healthcare professionals. Participants predominantly represented European ancestries. To address these sources of potential bias, improve metabolite harmonization, and facilitate further metabolomic epidemiology meta-analyses, we have summarized key recommendations that should be considered in future metabolomic epidemiology studies (**Figure 7**).

Associations Between Identified Metabolites and Other Cancers and Traits

We investigated whether any of the 33 total significant metabolites have been previously associated with other cancers and traits. Of these, 24 metabolites were identified in the Human Metabolite Database (HMDB)[25], 17 of which had been previously associated with traits ranging from neurological conditions such as schizophrenia to diabetes and obesity, with 13 metabolites associated with ≥1 cancer

type (colorectal, pancreatic, lung, breast and leukemia), although none were previously associated with PCa (**Supplementary Table 13**).

Drug and Diet Targets

To explore whether the 33 PCa-associated metabolites were modifiable by diet, the FooDB and MetaboFood databases were searched. Twenty-one metabolites were identified in FoodDB, 8 of which were previously quantified in food, with the other 13 expected but yet to be quantified. Dairy products were the most common food identified, with 5 metabolites previously quantified in dairy products (tiglycarnitine, alpha-ketoglutarate, glutamate, pseudouridine, and taurine). One additional metabolite (oxalate, found in oranges) was identified in MetaboFood (**Supplementary Table 14**, **Figure 6**).

The DrugBank pharmaco-metabolomics database was searched to determine whether the PCa-associated metabolites were drug modifiable. Two metabolites associated with overall PCa risk were altered by drugs (sphinganine and serotonin), while 7 metabolites associated with lethal PCa risk were altered by drugs (taurine, C-glycosyltryptophan, pseudouridine, 3-hydroxybutyrate, glutamate, glycerol, and cysteine; **Supplementary Table 15, Figure 6**).

Discussion

This large-scale systematic review and meta-analysis identified 33 circulating metabolites that were associated with PCa risk, with 3 associated with overall PCa, 11 associated with high- to very high-risk PCa, and 19 associated with lethal PCa risk.

Metabolite risk profiles were unique to each PCa outcome, with no metabolite significantly associated with >1 outcome and metabolite effects not correlated between PCa outcomes. This suggests that distinct metabolite drivers may contribute to the development of more aggressive PCa, which could have notable implications for tailored risk stratification to discern risk of clinically significant PCa from indolent PCa.

Changes in tumor metabolism during cancer progression are well documented and result in metabolomic heterogeneity between cancer stages, providing an opportunity to identify distinct markers of advanced and lethal disease[26]. A metabolomic study of PCa biopsy tissue identified seven metabolites that significantly differed when comparing tumors with aggressive features (low Ki67, a cellular proliferation marker, and high PSA (>8ng/mL)) from less aggressive tumors (high Ki67 and low PSA), including the amino acid taurine and the ratio of glutamate/glutamine[27]. This corroborates our finding that taurine was associated with increased lethal PCa risk. Further, the long noncoding RNA taurine-upregulated 1 gene (lncRNA *TUG1*), a gene expression regulator that can be upregulated by taurine, was shown to be upregulated in prostate tumor tissue, with higher *TUG1* associated with more aggressive PCa, increased cell proliferation and poorer survival[28], providing mechanistic insights into the role of taurine in lethal PCa.

We also found that glutamate was exclusively associated with lethal PCa risk, as suggested by previous systematic reviews[9,11]. Accordingly, a Mendelian randomization study found no evidence of association between glutamate and overall PCa risk[29]; however, higher serum glutamate has been reported in White and African American PCa patients with Gleason scores≥8 versus Gleason scores≤7 (measured post-

diagnosis)[30]. Glutamate can be produced by the glutaminolysis pathway, converted into alpha-ketoglutarate, and then utilized in the tricarboxylic acid (TCA) cycle, increasing energy production in cancer cells and facilitating cellular proliferation[31]. Interestingly, we also identified a protective effect of alpha-ketoglutarate on high- to very high-risk PCa, which has been previously suggested[11]. Recent work has shown that sine oculis homeobox homolog 1 (SIX1), a newly identified PCa driver, contributes to migration and proliferation by acting as an upstream regulator of the enzyme glutamate-pyruvate transaminase 2 (GPT2), increasing cellular alpha-ketoglutarate production[32]. Future mechanistic work should further explore the role of glytaminolysis and the TCA cycle in PCa progression.

Amino acids are expected to contribute to PCa due to their essential role in maintaining redox balance, biosynthesis and homeostatic regulation[33]. Alongside taurine and glutamate, we found that amino acids C-glycosyltryptophan and N-acetylserine were associated with increased lethal PCa risk, while cysteine and 4-hydroxyphenolpyruvate (4-HPPA) were associated with decreased lethal PCa risk. C-glycosyltryptophan has been associated with increased all-cause and cardiovascular disease mortality[34]. Pre-diagnostic serum N-acetylserine has been associated with increased colorectal cancer risk[35], and increased urinary N-acetylserine levels have been observed in endometrial carcinoma patients[36]. Cysteine, along with glutamate, is involved in the synthesis of glutathione (GSH), an antioxidant that helps mitigate oxidative stress from the accumulation of reactive oxygen species in cancer cells. *In vitro* and *in vivo* experiments found that intracellular L-cysteine depletion through the introduction of the engineered human enzyme Cyst(e)inase nearly depleted intracellular

GSH leading to cancer cell death, reduced tumor growth in prostate and breast cancer xenografts, and increased DNA damage in PCa cells[37,38], positioning Cyst(e)inase as a potential therapeutic agent. 4-HPPA is involved in tyrosine catabolism and is thought to reduce free radical levels, in turn reducing colon cancer risk[39]. As studies investigating the roles of these amino acids in PCa development are limited, future research may provide mechanistic insights into the potential for these to serve as preventive or therapeutic targets for clinically significant PCa.

Dysregulation of lipid metabolism is a characterizing feature of cancer[40]. Accordingly, we identified 1, 10 and 4 lipids that were significantly associated with risk of overall, high- to very high-risk and lethal PCa, respectively, with metabolites associated with high- to very high-risk PCa enriched for lipids. Nine of the 10 lipids associated with high- to very high-risk PCa were phosphatidylcholines (PCs), all of which were associated with reduced risk, with two also nominally associated with reduced risk of overall PCa. PCs are the most abundant membrane phospholipid and are precursors to other major membrane phospholipids (sphingomyelins and phosphatidylethanolamines) that can act as secondary messengers[41]. Due to this role in signaling and the generation of lipid mediators, PCs have been associated with both increased and decreased cancer risk[42], with limited studies exploring how PCs link to PCa. We also found that glycerophospholipid 1-linoleoyl-GPC was protective against lethal PCa risk, a lipid that was previously associated with reduced kidney cancer risk[43].

Three additional lipids, androstenediol (3beta,17beta) disulfate (2), 3-hydroxybutyrate (BHBA) and glycerol, were significantly positively associated with lethal PCa risk. Androstenediols are androgen steroid hormones that promote the growth of

both normal and cancerous prostate cells through androgen receptor binding, and androgen deprivation therapy (ADT) is commonly utilized to treat metastatic hormone-sensitive PCa[44]. Unfortunately, most patients eventually develop ADT resistance, resulting in metastatic hormone-resistant PCa, which is responsible for most PCa deaths[45]. Hence, understanding the role of androgens in PCa development and progression is a major area of research. Androstenediol (3beta,17beta) disulfate (2) is a disulfated androgen, and although the role of steroid disulfates is not fully understood, the addition of two steroid groups is thought to reduce membrane transferability and receptor binding[46]. To understand if these could be targets for PCa prevention, future research is needed to understand the role of disulfated androgens in androgen signaling pathways that influence PCa progression through increased cellular proliferation, migration and decreased apoptosis[47].

BHBA has been associated with increased and decreased cancer risk, likely due the 'butyrate paradox' in which some cancers preferentially utilize glucose as fuel and others preferentially oxidize BHBA[48]. BHBA acts as an inhibitor of cell proliferation via histone acetylation; therefore, in cancer types that preferentially oxidize BHBA, lower levels of BHBA are present and cell proliferation is not inhibited[48,49]. Indeed, in a cross-cancer meta-analysis of 1,900 metabolites, BHBA was one of the most upregulated blood metabolites across cancers[50]. In agreement with our findings, in a targeted multivariate serum metabolomics study, BHBA differentiated PCa cases with and without bone metastasis[51]. Furthermore, BHBA has been shown to increase proliferation and metastasis of colorectal cancer through the regulation of acetyl-CoA acetyltransferase (ACAT1)[52], the expression of which has been postulated as a cancer

therapeutic target and marker for PCa that may distinguish non-aggressive and aggressive disease[53,54].

Glycerol is commonly used as food additive to introduce sweetness to foods. Although considered safe, our positive association with lethal PCa risk aligns with a study reporting that the rate of tumor growth of human PC3 hormone-resistant PCa cells was doubled in mice receiving glycerol compared to control mice receiving a saline solution, which was speculated to be a result of a reduction in oxidative damage[55]. Furthermore, glycerol is a by-product of glycerol-3-phosphate phosphatase (G3PP), an enzyme involved in glucose and lipid metabolism and the maintenance of redox balance. Elevated G3PP expression in PCa epithelial cells has been associated with increased risk of biochemical recurrence and bone metastasis, potentially due to G3PP alleviating metabolic stress in PCa cells[56], demonstrating the potential prognostic utility of G3PP and the mechanistic relevance of glycerol to lethal PCa.

Four nucleotides were significantly associated with lethal PCa risk: pseudouridine, 2'-O-methyluridine, 5,6-dihydrouridine and dihydroorotate, with the latter three having positive associations and being related to pyrimidine metabolism. Limited studies have investigated the role of these nucleotides on PCa development; however, dihydroorotate is essential in the de-novo synthesis of pyrimidine and is the substrate for the enzyme dihydroorotate dehydrogenase (DHDOH), the inhibition of which is a potential target for cancer therapy due to the role of pyrimidines in RNA/DNA synthesis and cell proliferation[57]. Supporting the potential role of pyrimidine metabolism in lethal PCa, a previous systematic review[11] and a prospective investigation that was ineligible for our meta-analysis (as it did not report risk estimates on a continuous scale)

suggested a positive association between the pyrimidine nucleoside 2'-deoxyuridine and aggressive PCa risk[58]—consistent with our nominal evidence of association between 2'-deoxyuridine and risk of high- to very high-risk PCa. Future work is warranted to better understand the mechanistic role of pyrimidine nucleotides in lethal PCa and as potential biomarkers of clinically significant PCa.

Finally, three peptides were associated with increased lethal PCa risk: gamma-glutamylvaline, gamma-glutamylglycine and leucylglycine, with gamma-glutamylvaline also nominally associated with increased risk of high- to very high-risk PCa. Pre-diagnostic serum levels of these and several other gamma-glutamyl amino acids and dipeptides were associated with increased PCa-specific mortality in a survival case-only metabolomics investigation (that was ineligible for the current investigation)[59]. The enzyme responsible for the addition of a gamma-glutamyl group onto peptides, gamma-glutamyl transpeptidase (GGT), is expected to protect cancer cells against oxidative stress, leading to cancer cell growth and survival[60]. GGT has been shown to be upregulated in numerous cancers, including PCa, with greater levels observed in androgen-independent C4-2 and bone metastatic C4-2B cells compared to androgen-dependent LNCaP and partially androgen-independent C4 cells[61]. Further, expression of GGT was shown to increase tumor growth and chemoresistance in murine models[62]. Collectively, these findings demonstrate the importance of GGT in PCa progression and warrant additional research to understand how these peptides may act as biomarkers for lethal disease.

Our ROB assessment suggested that the publications included in this review were of good quality; however, key sources of bias were identified. One notable issue

was the selective reporting of only significant results in three (25%) studies, which could contribute to publication bias and introduce false positive and false negative associations in our findings. In addition, inconsistent naming of metabolites and the reporting of associations based on spectroscopy measures (e.g., m/z values) could further bias our results due to the inability to accurately aggregate all reported associations. Inconsistent adjustment of covariates across studies was observed (**Supplementary Table 2**) and may have contributed to evidence of heterogeneity. The inconsistent reporting of findings stratified by time-to-diagnosis limited our ability to investigate temporal variability in metabolite-PCa associations, which has often been found to impact metabolite associations and will be important to investigate in follow-up studies of the metabolites reported here. Bias was also attributed to features of the primary cohort, including sample collection methods (e.g., collecting non-fasting samples, which could impact certain metabolite levels) and the generalizability of studies. Most notably, all identified studies were conducted in predominantly European descent individuals; therefore, our results may not be generalizable across populations, which is especially important given the stark health disparities of PCa, with Black and African American men having significantly higher incidence and mortality rates compared to White men[63,64]. This lack of diversity has also been noted in targeted metabolomic investigations of PCa[11], further restricting the ability to validate metabolite findings across populations.

This is the first study to quantitively aggregate PCa metabolomic investigations to assess the relationship between untargeted circulating metabolites and clinically relevant PCa outcomes, combining the power of previous investigations into a large

meta-analysis of ~600 metabolites. With this approach, across the tested metabolites, we had 80% power to detect average relative risks of ≥1.23, 1.25, 1.51, and 1.29 for overall, low- to intermediate risk, high- to very high-risk, and lethal PCa, respectively, in fixed effects models (**Supplementary Table 16**). As such, this represents the most comprehensive understanding of metabolomic risk of PCa to date. Follow-up analyses of significant metabolites using drug and dietary databases also provides clinical and public health contexts into whether the identified metabolites could serve as dietary or drug targets.

However, this study has limitations. Due to the limited number of studies identified and overlap of metabolites between studies, we were unable to investigate potentially confounding factors (e.g., age or follow up time) via meta-regression or to formally assess publication bias. (e.g., funnel plot inspection or the harbord test). This also resulted in 75% of metabolites being tested in only two studies, restricting our ability to perform random effect analyses and more thoroughly assess heterogeneity. It will be important to further validate the PCa-associated metabolites only tested in two studies. Despite our extensive efforts to harmonize metabolites, naming discrepancies likely reduced the number of metabolites that could be successfully harmonized and meta-analyzed, particularly for poorly characterized metabolites and metabolites with multiple naming conventions (e.g., fatty acyls). Challenges in metabolite harmonization due to discrepancies in metabolite quantification, identification and annotations, along with the selective reporting of significant metabolite findings, have been similarly described in other metabolomic meta-analyses[50,65] and will be important for the field to address to facilitate evidence synthesis from metabolomic studies and better

understand the contribution of metabolites to disease risk. Towards this goal, we propose a set of guidelines for metabolomic epidemiology studies that could mitigate such challenges.

In conclusion, this study has identified a range of metabolites significantly associated with risk of clinically significant PCa. Metabolites associated with risk of high- to very high-risk and lethal PCa, such as glycerophospholipids, amino acids and pyrimidine nucleotides, may have clinical utility as biomarkers for early detection of PCa with poor prognosis. Future prospective investigations are warranted to assess the generalizability of these findings across diverse populations, particularly high-risk Black and African American populations. It also will be crucial to perform untargeted metabolomic studies across diverse populations to identify novel metabolite associations that may have been missed by previous investigations limited to predominantly European descent men. Further, conducting *in vivo* and/or *in vitro* experiments will be important next steps to determine the biological mechanisms underlying the metabolite associations reported here and better understand the potential for these metabolites to serve as biomarkers for PCa prevention, screening, and disease management.

Methods

Search Strategy

This systematic review is registered in the International Prospective Register of Systematic Reviews (PROSPERO; ID CRD42023462809). PubMed and Embase databases were systematically searched by one author (H.F.) for literature published

through October 17th, 2024, using the following strategy: ('prostate cancer' or 'prostate carcinoma') AND (metabolite* or metabolomics or metabolomic) AND (serum or blood or plasma or circulating). Mesh indexing, human studies, English language, and original research filters were applied in each database, as applicable. Titles and abstracts were screened in duplicate by two authors (H.F. and O.P.A) using Abstrackr[66], and disagreements were mediated by a third author (B.F.D.).

Inclusion criteria for the review were published original prospective human research studies written in English that reported on the association between PCa risk and ≥50 circulating metabolites (taken as an indication of an untargeted study) measured prior to prostate cancer diagnosis (i.e., cohort, nested case-control, or case-cohort). Randomized control trials (RCTs) were excluded due to their targeted nature. Given the accessibility of blood samples, which is ideal for biomarker development, and the commonality of blood-based metabolomic studies[4], we limited studies to those conducted in blood samples. Studies that did not provide risk estimates for metabolites individually (rather they provided risk estimates for metabolite scores, conducted multivariate analyses, or evaluated prediction models) were also excluded along with studies that did not report metabolites on a continuous scale. Citation lists of included studies were searched until no additional studies were identified.

Data Extraction and Synthesis

The following variables were extracted from included studies: first author, year, study name, study design, study population/country, case and control sample size, age at sample draw, years between sample draw and diagnosis, PCa definitions, sample

type, fasting status, quantification technique, metabolomics platform, number of metabolites evaluated, main method of data analysis, model covariates, all metabolite identifiers, metabolite unit, effect estimates and standard errors. Odds ratios (ORs), risk ratios (RR), and hazard ratios (HRs) were interpreted as relative risks. Study authors were contacted to attempt to obtain missing results as needed.

When ≥1 publication was identified from the same cohort, the publication with the largest number of cases was selected unless it was stated that included individuals were independent from the previous publication (**Supplementary Figure 3**). When a study presented multiple effect estimates for a given metabolite, effect estimates adjusting for the most covariates were utilized. When effect estimates were only provided for specific strata (i.e., tumor stage, PTEN status, or age categories), the category most comparable to other studies included for the given outcome was selected to minimize heterogeneity.

When a study quantified a metabolite on both MS and NMR platforms, MS estimates were used in meta-analyses given the larger number of MS studies.

PCa Outcome Definitions

Based on the case characteristics of included studies, four distinct PCa outcomes were assessed, based on NCCN definitions[23]: overall PCa, low- to intermediate-risk PCa, high- to very high risk PCa and lethal PCa, each compared to PCa-free controls (**Supplementary Table 4**). Overall PCa was based on national registries, cancer registries, and electronic health records. Due to limited sample sizes, “low- to intermediate-risk PCa” combined low-risk PCa (patients with grade group 1,

cT1-cT2a staging and PSA<10ng/mL) and intermediate-risk PCa (patients with no high-risk features who had grade group 2-3, PSA of 10-20ng/mL and/or cT2b-cT2c staging). Likewise, “high- to very high-risk PCa” combined high-risk and very high-risk PCa (patients with cancer that had grown outside the prostate (cT3a or above), grade group of 4-5 and/or PSA>20ng/mL). Lethal PCa was defined as metastatic PCa or PCa-specific death. Studies defining PCa outcomes with other definitions were excluded to minimize heterogeneity.

Metabolite Harmonization

Metabolites were harmonized based on COMP ID (when available for metabolites quantified by Metabolon), followed by HMDB IDs[25] accessed online and via the R package hmdbQuery,[67] PubChem IDs,[68] and metabolite names. When metabolite names were used for harmonization, alternate metabolite names reported in HMDB and PubChem along with metabolite structures were also considered. Metabolite harmonization was verified using all available IDs and by manual check. Metabolites were mapped to biological pathways based on information provided in identified studies and HMDB.

Statistical Analysis

Metabolites were meta-analyzed separately for each outcome using a restricted maximum likelihood (REML) approach for metabolites identified in ≥2 studies to reduce error resulting from meta-analyzing studies with small sample sizes[69,70]. Fixed effect meta-analysis results were presented when heterogeneity was limited (I^2≤40%) or if

there were two studies, with random effect meta-analysis results presented otherwise. Results were presented stratified by MS and NMR platforms to assess heterogeneity. Between-platform and between-study heterogeneity was assessed via the I^2 statistic and a chi-squared test for subgroup differences (Q-test P-value <0.05). An FDR (α=0.05) correction was implemented to account for multiple testing separately for each cancer outcome. Nominal associations were defined as those with an unadjusted P<0.05.

To determine whether PCa-associated metabolites were enriched for any biological pathways, enrichment analyses were performed using a one-sided Fisher's exact hypergeometric test with the R package bc3net[71]. Enrichment analyses were conducted on significant and nominally significant metabolites separately for each PCa outcome, with the reference panel including all metabolite meta-analyzed for a given outcome and pathway, comparing the number of observed to expected metabolites for a given pathway. Significance was defined using an FDR α=0.05 to account for multiple testing separately for each outcome.

We used Pearson correlation to determine how comparable effect estimates were between PCa outcomes for overlapping metabolites. This was conducted across all overlapping metabolites, metabolites with an unadjusted P<0.05 in both outcomes, and those with an unadjusted P<0.05 in one or the other outcome. Correlations were not assessed for outcomes with <15 metabolites due to insufficient data.

Results were reported following the Meta-analysis of Observational Studies in Epidemiology (MOOSE) and the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines[72,73].

Associations Between Identified Metabolites and Other Cancers and Traits

To evaluate whether any of the PCa-associated metabolites were previously associated with cancer or other traits, we searched the HMDB database on December 12th, 2024. Conditions listed under 'associated disorders and diseases' and the 'health effects' subcategory of physiological effects were extracted.

Drug and Diet Metabolite Targets

To explore potential environmental modifiers of the identified significant metabolites, the FooDB and MetaboFood databases were searched to determine if these metabolites have previously been quantified in foodstuffs, and therefore, if any identified metabolites could potentially be dietary modifiable. Metabolites were searched using metabolite names (FooDB and MetaboFood), HMDB IDs (FooDB) and chemical structure (MetaboFood). To determine whether the PCa-associated metabolites could be drug modifiable, we searched the DrugBank pharmaco-metabolomics database[74], utilizing all commonly used metabolite synonyms. Searches were conducted on July 29th, 2025.

Risk of Bias (ROB) Assessment

Quality of the included studies was assessed with a published scoring tool[24], previously utilized in other metabolomic epidemiology systemic reviews[70], which evaluates six potential sources of bias: study participation, study attrition, exposure assessment, outcome assessment, evaluation of confounders, and the appropriateness

of statistical analysis. Each domain was assigned a positive ('yes', scoring 1), neutral ('somewhat', scoring 0.5) or negative ('no', scoring 0) score to indicate whether the study avoided biases in each domain, based upon predefined criteria adopted for this review (**Supplementary Table 13**). An overall score was summated for each study, and a score ≤3 was considered low quality[70]. Question two on study attrition was not considered for nested case-control studies, as information on study attrition details were expected to be reported in the overall cohort study rather than in the nested case-control study. ROB was assessed by two reviewers (H.F and O.P.A).

Additional Information

Authors' contributions

H.F. and O.P.A. performed the systematic review and B.F.D. mediated disagreements. H.F. and O.P.A. assessed the risk of bias criteria for each study. H.F. conducted all statistical analyses. J.M.G. assisted with public database searches and network figures. H.F. and B.F.D. drafted the manuscript. O.P.A. and J.M.G. provided critical review of the manuscript. H.F. and B.F.D contributed to study conception and design. B.F.D. supervised the investigation.

Competing interests

The authors declare no conflict of interest.

Funding information

This work was supported by the National Institutes of Health (R00 CA246063, Pacific Northwest Prostate Cancer SPORE P50 CA097186, Cancer Center Support Grant P30 CA015704, R01CA258808, U54HG013243), an award from the Andy Hill Cancer Research Endowment Distinguished Researchers Program, a Fred Hutch/University of Washington SPORE Career Enhancement Program award, the Prostate Cancer Foundation (21YOUN11), and the Institute for Prostate Cancer Research.

Figures and Figure Legends

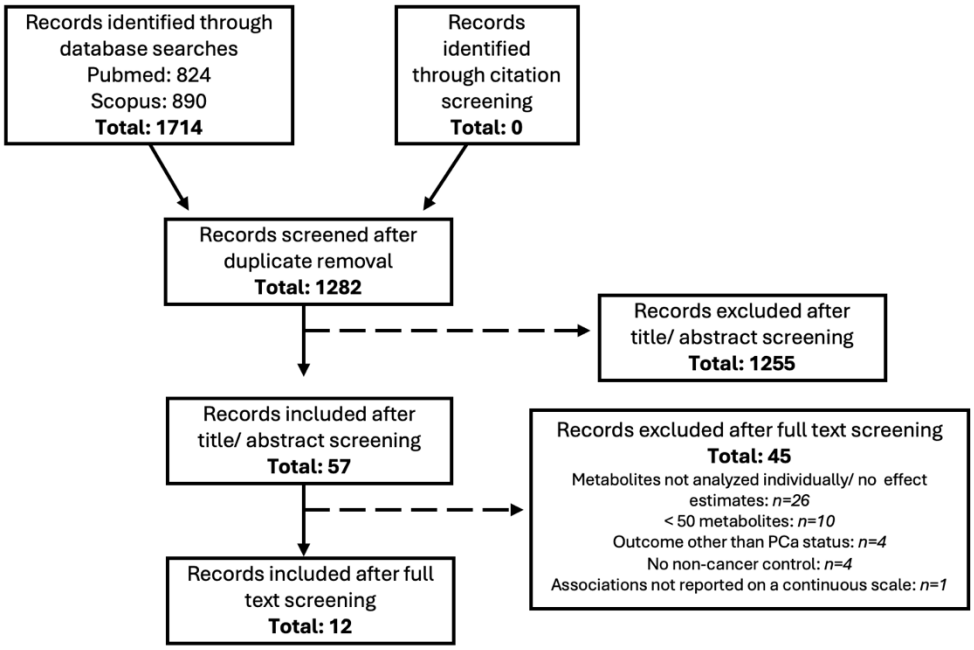


Figure 1: PRISMA flow diagram for systematic review of untargeted metabolomic studies and prostate cancer risk

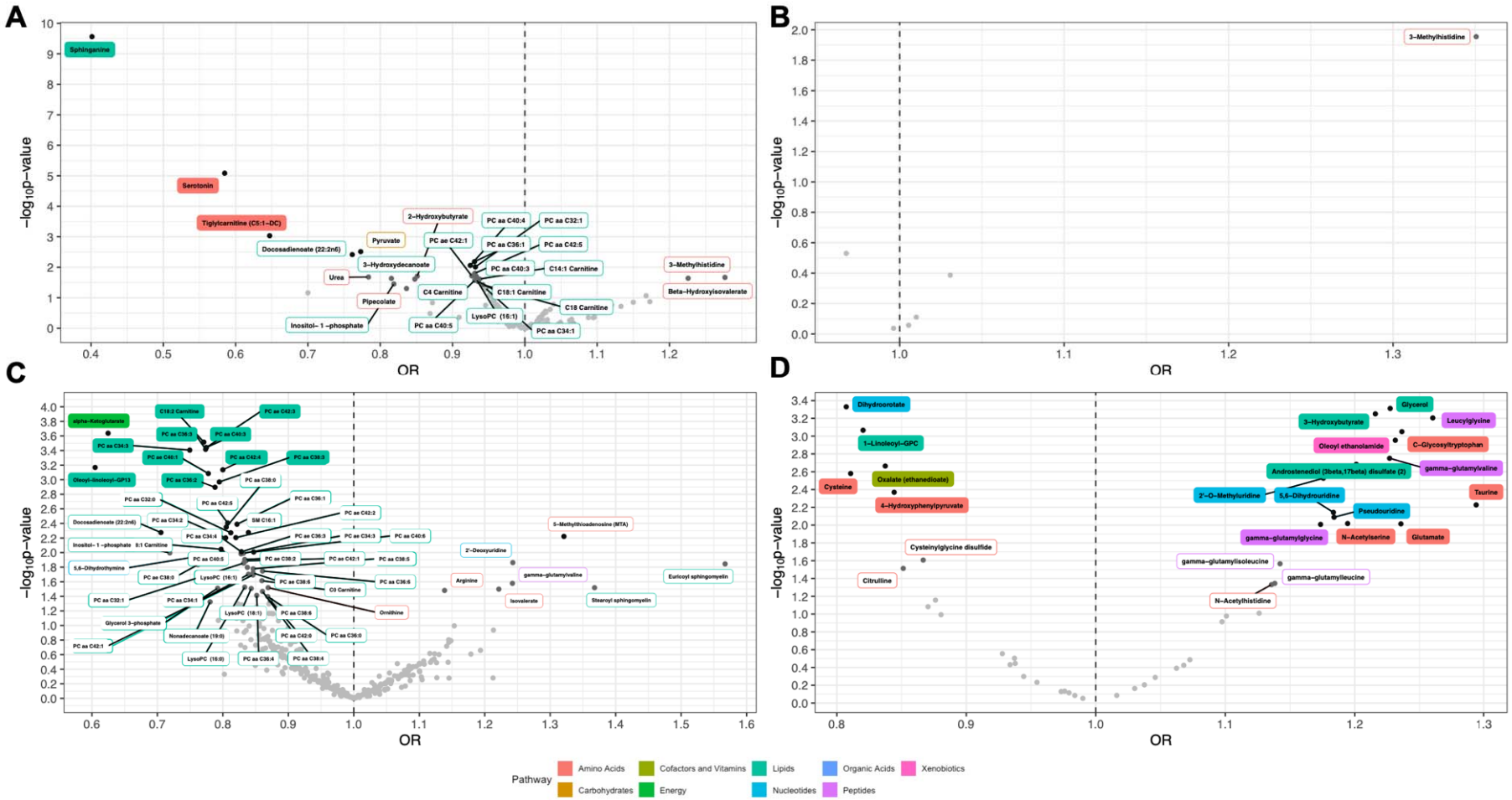


Figure 2: Meta-analysis association results between metabolites and PCa outcomes. Results are shown for A) overall PCa, B) low- to intermediate-risk PCa, C) high- to very high-risk PCa, and D) lethal PCa. Named metabolites were nominally associated (unadjusted P<0.05) with the respective PCa outcome. Shaded labels show FDR significant

metabolites. Shading and label outline color represent metabolite class. Light grey point: $P \geq 0.05$. Dark grey point: $P < 0.05$. Black point: $P < 0.01$.

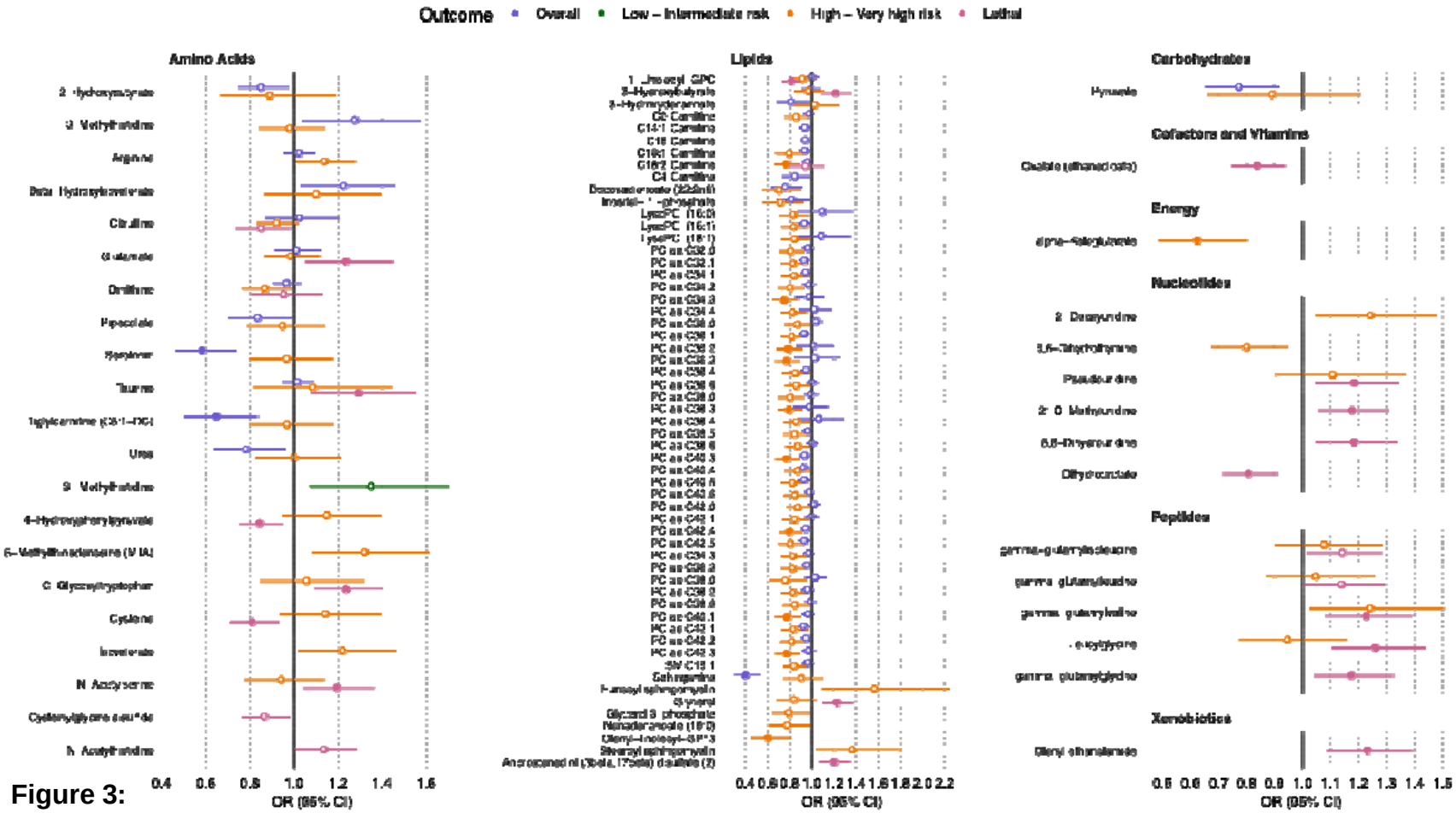
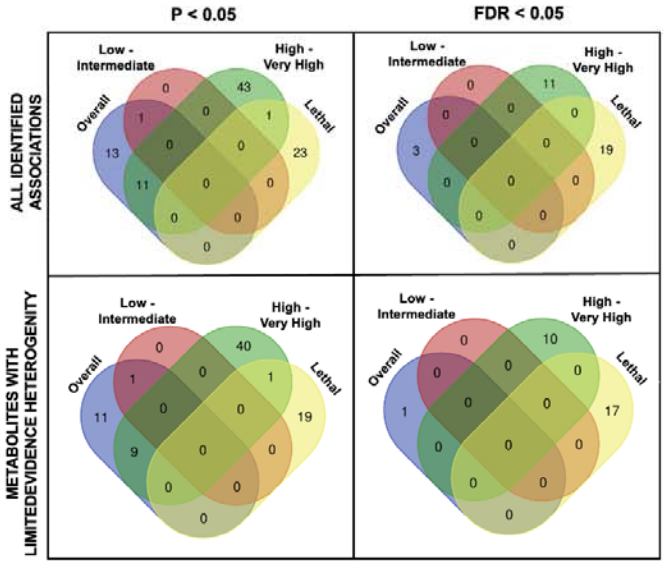


Figure 3:

Association results for metabolites nominally associated with ≥ 1 PCa outcome. All tested outcomes are shown for each metabolite. Shaded circles indicate FDR significant associations.

A



B

Outcome	Nmets	Median I^2	Nmets I^2=0 (%)	Nmets Q.value <0.05 (%)
All tested metabolites				
Overall	150	34.01	63 (42)	31 (21)
Low – Intermediate	6	35.04	2 (33.3)	0 (0)
High – Very High	391	6.11	190 (49)	52 (13)
Lethal	48	0	25 (52)	11 (23)
Nominally significant metabolites				
Overall	25	0	16 (64)	4 (16)
Low – Intermediate	1	0	1 (100)	0 (0)
High – Very High	55	12.44	23 (41.82)	5 (9.09)
Lethal	24	0	14 (58.3)	4 (16.67)
FDR significant metabolites				
Overall	3	91.72	0	2 (66%)
Low – Intermediate	0	-	-	-
High – Very High	11	33.39	2 (18.18)	1 (9.09)
Lethal	19	0	13 (68.42)	2 (10.53)

Figure 4: Summary of metabolite effect estimate heterogeneity by PCa outcome and significance threshold. A: Overlap of FDR significant and nominally significant metabolites before and after the removal of metabolites where N=2 some evidence of heterogeneity. **B:** Heterogeneity measures across all outcomes. Nmets: Number of metabolites.

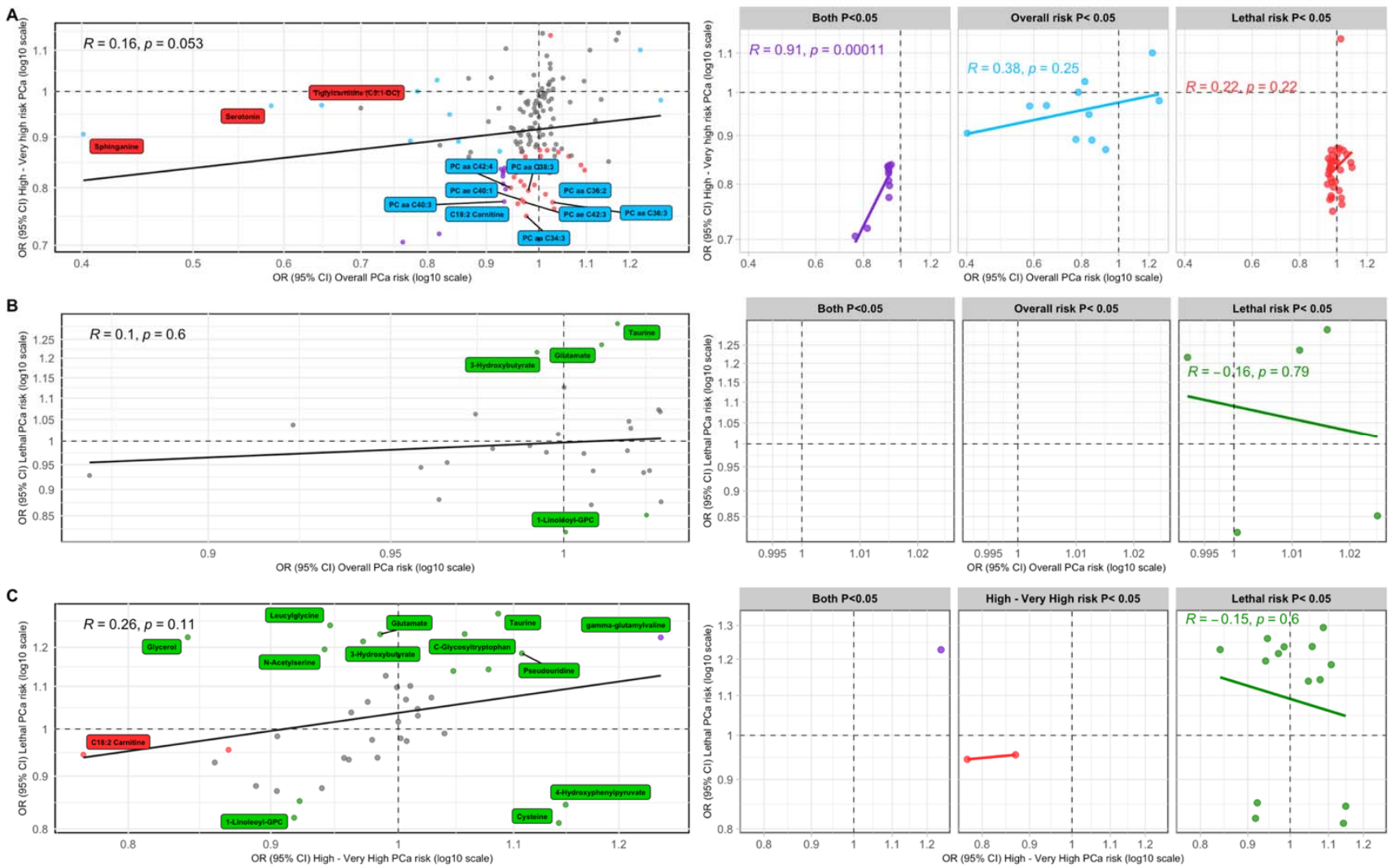


Figure 5: Pearson correlation of metabolite effect estimates between PCa outcomes. **A**: Overall vs high- to very high-risk PCa (N_{mets}=143), **B**: Overall vs lethal PCa (N_{mets}=27), **C**: High- to very high-risk vs lethal PCa (N_{mets}=39). For

each comparison, the three panels on the right indicate correlations between metabolites associated with an unadjusted $P<0.05$ in both traits or in one or the other trait. Purple points indicate results that were associated with both outcomes with an unadjusted $P<0.05$. Blue points indicate results that were associated with overall PCa with an unadjusted $P<0.05$. Red points indicate results that were associated with high- to very high-risk PCa with an unadjusted $P<0.05$. Green points indicate results that were associated with lethal PCa with an unadjusted $P<0.05$. Grey points indicate results that were associated with both outcomes with $P\geq 0.05$. Named metabolites were significantly associated (FDR<0.05) with the outcome of the given color. OR plotted on the natural log scale for clarity.

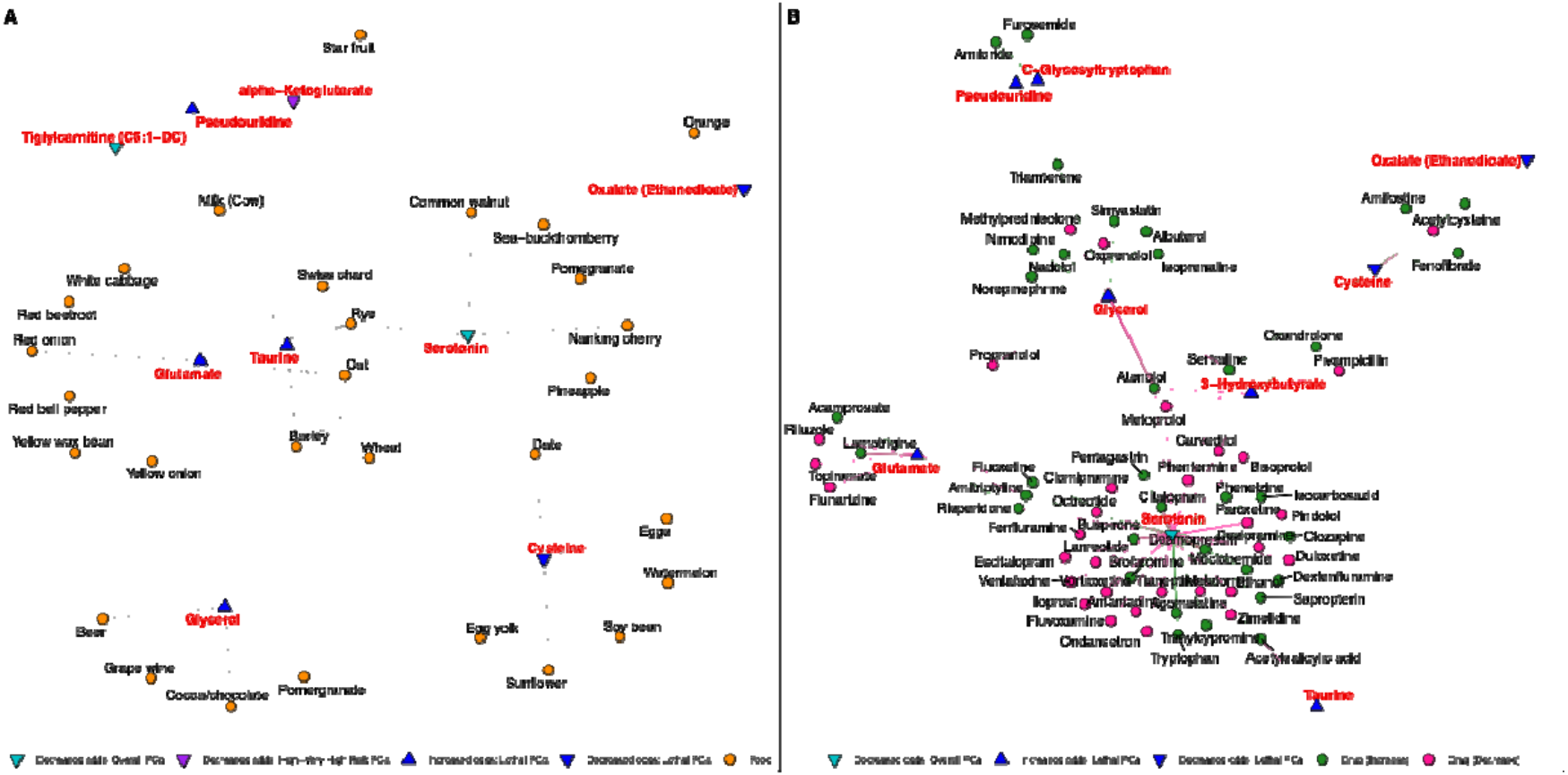


Figure 6: Identified dietary and drug modifiers of metabolites significantly associated with PCa outcomes. A: Foods that the PCa-associated metabolites have been previously quantified in according to FooDB and MetaboFood. Food names are harmonized for simplicity, with details provided in Supplementary Table 14. **B**: Drug modifiers of PCa-associated metabolites in the DrugBank pharmaco-metabolomics database; details are provided in Supplementary Table 15. PCa outcome associated with each metabolite is indicated by the color of the triangle, with upward triangles reflecting

positive associations and downward triangles reflecting negative associations. In A, orange circles indicate foods, while in B, pink circles and lines indicate drugs that decrease metabolite levels and green circles and lines indicate drugs that increase metabolite levels.

❑ **Methods and Statistical Approaches**

- Detail scaling, standardization and transformation procedures applied.

❑ **Sample Information**

- Provide information on fasting status of participants.
- Report on the time of sample collection in relation to outcome.

❑ **Metabolite Descriptions**

- Report commonly used metabolite identifiers (HMDB, PubChem, KEGG, platform specific IDs).
- Use conventional spelling for metabolite names and conventional formatting for abbreviations, including structural abbreviations within metabolite names.
- State the meaning of used abbreviations.
- State clearly if composite metabolite biomarkers are utilized and which metabolites are considered in these measures.
- Explain how metabolite pathways were assigned.

❑ **Outcome Measurement**

- Provide quantitative measures of time until diagnosis, including a measure of average time until diagnosis (eg., mean, median and measures of spread (eg., range, interquartile range).

❑ **Reporting of Findings**

- Provide all association results, including non-significant findings.
- Provide association results on a continuous scale that is comparable across populations.
- Report unadjusted and adjusted results, including information on all adjusted variables.
- Clearly state any cohort characteristics that may result in findings being less generalizable across populations.

Figure 7: Recommended guidelines to improve reproducibility and facilitate meta-analyses of untargeted metabolomic epidemiology studies.

Table 1: Description of studies included in the systematic review and meta-analysis

Reference	Cohort	Country	Quantification Method	Time from sample collection to diagnosis or PCa death	Number of PCa cases and PCa cases by investigated outcome
Feng et al, 2021[20]	HPFS/ PHS	US	LC-MS Broad	Mean time to diagnosis: 5.6-5.9 years	All reported study cases: 488 Overall PCa: 177 Controls: 294
Huang et al, 2017[12]	ATBC	Finland	UHPLC-MS and GC-MS Metabolon	Median time to diagnosis: 10 years (range 1-20 years)	All reported study cases: 137 Low-Intermediate risk PCa: 71 High-Very high risk PCa: 66 Controls: 200
Huang et al, 2019[13]	ATBC	Finland	UHPLC-MS/ tandem MS Metabolon	Mean time to PCa death: 18 years	All reported study cases : 523 Lethal PCa: 523 Controls: 523
Lécuyer et al, 2021[19]	SU.VI.MAX	France	Bruker AVANCE III ^{1}H NMR†	Up to 13 years	All reported study cases: 171 Overall PCa: 171 Low-Intermediate risk PCa: 88 Controls: 171
Lin et al, 2021[18]	SU.VI.MAX	France	UHPLC-HRMS	Mean time to diagnosis: 8.3 years (up to 13 years)	All reported study cases: 146 Overall PCa: 146 Controls: 272
Mondal et al, 2014[15]	ATBC	Finland	UHPLC-MS and GC-MS Metabolon	Range 1-23 years to diagnosis	All reported study cases: 74 Overall PCa: 74 High-Very high risk PCa : 53 Controls: 74
Mondul et al, 2015[14]	ATBC	Finland	UHPLC-MS and GC-MS Metabolon	Range 10-20 years to diagnosis	All reported study cases: 200 Overall PCa: 200 High-Very high risk PCa : 100 Controls: 200
Östman et al, 2022[75]	NSHDS	Sweden	MS	Follow up time >5 years	All reported study cases: 752 Overall PCa: 752 High-Very high risk PCa : 165 Controls: 752
Röhnisch et al, 2020[21]	NSHDS	Sweden	Triple quadrupole MS BIOCRATES NMR	Median time to diagnosis: 10.3 years (range 5-19.9 years)	All reported study cases: 777 Overall PCa: 777 High-Very high risk PCa: 169 Controls: 777
Schmidt et al, 2017[17]	EPIC	European Countries*	Triple quadrupole MS BIOCRATES	39.7% diagnosed <5 years before sample collection 29.3% diagnosed ≥10 years after sample collection	All reported study cases: 1077 Overall PCa: 1077 Low-Intermediate risk PCa: 778 High-Very high risk PCa : 208 Lethal PCa: 127

					Controls: 1077
Schmidt et al, 2020[16]	EPIC	European Countries*	Triple quadrupole MS BIOCRATES	Mean time to diagnosis: 9.4 years (4.2 SD)	All reported study cases: 3057 Overall PCa: 3057 Low-Intermediate risk PCa: 2517 High-Very high risk PCa: 580 Lethal PCa: 297 Controls: 3057
Wang et al, 2021[22]	CPS-II	USA	UHPLC-MS/MS Metabolon	44% cases died < 10 years of follow up, 56% died ≥ 10 years of follow up	All reported study cases: 241 High-Very high risk PCa: 129 Lethal PCa: 112 Controls: 342

ATBC: Alpha-Tocopherol, Beta-Carotene Cancer Prevention; CPS-II: Cancer Prevention Study-II Nutrition Cohort; EPIC: European Prospective Investigation into Cancer and Nutrition; GC: Gas Chromatography; LC: Liquid Chromatography; HPFS: Health Professionals Follow-up Study; MS: Mass Spectroscopy; NSHDS: Northern Sweden Health and Disease Study; PHS: Physician's Health Study; SD: Standard deviation; SU.VI.MAX: Supplémentation en Vitamines et Minéraux Antioxydants; UHPLC-MS: Ultra Performance Liquid Chromatography Mass Spectroscopy. UHPLC-HRMS: Ultra Performance Liquid Chromatography High Resolution Mass Spectroscopy. *19 sites across Denmark, Germany, Greece, Italy, the Netherlands, Spain, Sweden, and the UK. † NOESY1D (NOESY1dgppr sequence) samples and not CPMG (Carr-Purcell-Meiboom-Gill) utilized as larger sample size.

References

1 Gandaglia, G. *et al.* Epidemiology and Prevention of Prostate Cancer. *European Urology Oncology* **4**, 877-892 (2021).

2 Wang, L. *et al.* Prostate Cancer Incidence and Mortality: Global Status and Temporal Trends in 89 Countries From 2000 to 2019. *Front Public Health* **10**, 811044 (2022).

3 American Cancer Society. Cancer Facts & Figures 2024., (American Cancer Society,, 2024).

4 Kelly, R. S., Vander Heiden, M. G., Giovannucci, E. & Mucci, L. A. Metabolomic Biomarkers of Prostate Cancer: Prediction, Diagnosis, Progression, Prognosis, and Recurrence. *Cancer Epidemiol Biomarkers Prev* **25**, 887-906 (2016).

5 Singh, R. & Mills, I. G. The Interplay Between Prostate Cancer Genomics, Metabolism, and the Epigenome: Perspectives and Future Prospects. *Front Oncol* **11**, 704353 (2021).

6 Di Minno, A., Gelzo, M., Stornaiuolo, M., Ruoppolo, M. & Castaldo, G. The evolving landscape of untargeted metabolomics. *Nutrition, Metabolism and Cardiovascular Diseases* **31**, 1645-1652 (2021).

7 Schrimpe-Rutledge, A. C., Codreanu, S. G., Sherrod, S. D. & McLean, J. A. Untargeted Metabolomics Strategies-Challenges and Emerging Directions. *J Am Soc Mass Spectrom* **27**, 1897-1905 (2016).

8 Fuller, H. *et al.* Metabolomic epidemiology offers insights into disease aetiology. *Nat Metab* **5**, 1656-1672 (2023).

9 Kdadra, M., Höckner, S., Leung, H., Kremer, W. & Schiffer, E. Metabolomics Biomarkers of Prostate Cancer: A Systematic Review. *Diagnostics (Basel)* **9** (2019).

10 Bansal, N., Kumar, M., Sankhwar, S. N. & Gupta, A. Relevance of emerging metabolomics-based biomarkers of prostate cancer: a systematic review. *Expert Reviews in Molecular Medicine* **24**, e25 (2022).

11 López-Hernández, Y. *et al.* Prostate cancer risk biomarkers from large cohort and prospective metabolomics studies: A systematic review. *Translational Oncology* **51**, 102196 (2025).

12 Huang, J. *et al.* Prospective serum metabolomic profile of prostate cancer by size and extent of primary tumor. *Oncotarget* **8**, 45190-45199 (2017).

13 Huang, J. *et al.* Prospective serum metabolomic profiling of lethal prostate cancer. *Int J Cancer* **145**, 3231-3243 (2019).

14 Mondul, A. M. *et al.* Metabolomic analysis of prostate cancer risk in a prospective cohort: The alpha-tocolpherol, beta-carotene cancer prevention (ATBC) study. *Int J Cancer* **137**, 2124-2132 (2015).

15 Mondul, A. M. *et al.* 1-stearoylglycerol is associated with risk of prostate cancer: results from serum metabolomic profiling. *Metabolomics* **10**, 1036-1041 (2014).

16 Schmidt, J. A. *et al.* Patterns in metabolite profile are associated with risk of more aggressive prostate cancer: A prospective study of 3,057 matched case-control sets from EPIC. *Int J Cancer* **146**, 720-730 (2020).

17 Schmidt, J. A. *et al.* Pre-diagnostic metabolite concentrations and prostate cancer risk in 1077 cases and 1077 matched controls in the European Prospective Investigation into Cancer and Nutrition. *BMC Med* **15**, 122 (2017).

18 Lin, X. *et al.* Plasma Metabolomics for Discovery of Early Metabolic Markers of Prostate Cancer Based on Ultra-High-Performance Liquid Chromatography-High Resolution Mass Spectrometry. *Cancers (Basel)* **13** (2021).

19 Lécuyer, L. *et al.* NMR metabolomic profiles associated with long-term risk of prostate cancer. *Metabolomics* **17**, 32 (2021).

20 Feng, X. *et al.* Association of Prediagnostic Blood Metabolomics with Prostate Cancer Defined by ERG or PTEN Molecular Subtypes. *Cancer Epidemiol Biomarkers Prev* **30**, 1000-1008 (2021).

21 Röhnisch, H. E. *et al.* Identification of metabolites associated with prostate cancer risk: a nested case-control study with long follow-up in the Northern Sweden Health and Disease Study. *BMC Medicine* **18**, 187 (2020).

22 Wang, Y., Jacobs, E. J., Carter, B. D., Gapstur, S. M. & Stevens, V. L. Plasma Metabolomic Profiles and Risk of Advanced and Fatal Prostate Cancer. *Eur Urol Oncol* **4**, 56-65 (2021).

23 National Comprehensive Cancer Network (NCCN). *Practice Guidelines in Oncology: Prostate Cancer. Version 2*, https://www.nccn.org/professionals/physician_gls/pdf/prostate.pdf Accessed 12/23, (2023).

24 Hayden, J. A., Côté, P. & Bombardier, C. Evaluation of the Quality of Prognosis Studies in Systematic Reviews. *Annals of Internal Medicine* **144**, 427-437 (2006).

25 Wishart, D. S. *et al.* HMDB 5.0: the Human Metabolome Database for 2022. *Nucleic Acids Res* **50**, D622-d631 (2022).

26 Demicco, M., Liu, X.-Z., Leithner, K. & Fendt, S.-M. Metabolic heterogeneity in cancer. *Nature Metabolism* **6**, 18-38 (2024).

27 Dudka, I., Lundquist, K., Wikström, P., Bergh, A. & Gröbner, G. Metabolomic profiles of intact tissues reflect clinically relevant prostate cancer subtypes. *J Transl Med* **21**, 860 (2023).

28 Yu, J., Wang, Y. & Peng, S. The high expression of lncRNA TUG1 correlates with progressive tumor condition and less satisfying survival profiles in prostate cancer patients. *Transl Cancer Res* **8**, 1817-1825 (2019).

29 Lin, Y. *et al.* Effects of glutamate and aspartate on prostate cancer and breast cancer: a Mendelian randomization study. *BMC Genomics* **23**, 213 (2022).

30 Koochekpour, S. *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* **18**, 5888-5901 (2012).

31 Lima, A. R., Pinto, J., Bastos, M. d. L., Carvalho, M. & Guedes de Pinho, P. NMR-based metabolomics studies of human prostate cancer tissue. *Metabolomics* **14**, 88 (2018).

32 Huang, H. *et al.* The Ku70-SIX1-GPT2 axis regulates alpha-ketoglutarate metabolism to drive progression of prostate cancer. *Oncogene* 10.1038/s41388-024-03209-8 (2024).

33 Lieu, E. L., Nguyen, T., Rhyne, S. & Kim, J. Amino acids in cancer. *Experimental & Molecular Medicine* **52**, 15-30 (2020).

34 Huang, J. *et al.* Serum Metabolomic Profiling of All-Cause Mortality: A Prospective Analysis in the Alpha-Tocopherol, Beta-Carotene Cancer Prevention (ATBC) Study Cohort. *Am J Epidemiol* **187**, 1721-1732 (2018).

35 Cross, A. J. *et al.* A prospective study of serum metabolites and colorectal cancer risk. *Cancer* **120**, 3049-3057 (2014).

36 Shao, X. *et al.* Screening and verifying endometrial carcinoma diagnostic biomarkers based on a urine metabolomic profiling study using UPLC-Q-TOF/MS. *Clin Chim Acta* **463**, 200-206 (2016).

37 Cramer, S. L. *et al.* Systemic depletion of L-cyst(e)ine with cyst(e)inase increases reactive oxygen species and suppresses tumor growth. *Nature Medicine* **23**, 120-127 (2017).

38 Saha, A. *et al.* Cysteine depletion sensitizes prostate cancer cells to agents that enhance DNA damage and to immune checkpoint inhibition. *J Exp Clin Cancer Res* **42**, 119 (2023).

39 Amić, A. *et al.* Free radical scavenging and COX-2 inhibition by simple colon metabolites of polyphenols: A theoretical approach. *Computational Biology and Chemistry* **65**, 45-53 (2016).

40 Hanahan, D. Hallmarks of Cancer: New Dimensions. *Cancer Discov* **12**, 31-46 (2022).

41 van der Veen, J. N. *et al.* The critical role of phosphatidylcholine and phosphatidylethanolamine metabolism in health and disease. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1859**, 1558-1572 (2017).

42 Saito, R. F., Andrade, L. N. S., Bustos, S. O. & Chammas, R. Phosphatidylcholine-Derived Lipid Mediators: The Crosstalk Between Cancer Cells and Immune Cells. *Front Immunol* **13**, 768606 (2022).

43 Guida, F. *et al.* The blood metabolome of incident kidney cancer: A case-control study nested within the MetKid consortium. *PLoS Med* **18**, e1003786 (2021).

44 Dhavale, M. *et al.* Androgen Receptor Signaling and the Emergence of Lethal Neuroendocrine Prostate Cancer With the Treatment-Induced Suppression of the Androgen Receptor: A Literature Review. *Cureus* **13**, e13402 (2021).

45 Auchus, R. J. & Sharifi, N. Sex Hormones and Prostate Cancer. *Annual Review of Medicine* **71**, 33-45 (2020).

46 Lightning, T. A., Gesteira, T. F. & Mueller, J. W. Steroid disulfates - Sulfation double trouble. *Molecular and Cellular Endocrinology* **524**, 111161 (2021).

47 Aurilio, G. *et al.* Androgen Receptor Signaling Pathway in Prostate Cancer: From Genetics to Clinical Applications. *Cells* **9** (2020).

48 Rodrigues, L. M. *et al.* The action of β-hydroxybutyrate on the growth, metabolism and global histone H3 acetylation of spontaneous mouse mammary tumours: evidence of a β-hydroxybutyrate paradox. *Cancer Metab* **5**, 4 (2017).

49 Labanca, E. *et al.* Prostate cancer castrate resistant progression usage of non-canonical androgen receptor signaling and ketone body fuel. *Oncogene* **40**, 6284-6298 (2021).

50 Goveia, J. *et al.* Meta-analysis of clinical metabolic profiling studies in cancer: challenges and opportunities. *EMBO Mol Med* **8**, 1134-1142 (2016).

51 Zhang, X. *et al.* Identification of characteristic metabolic panels for different stages of prostate cancer by 1H NMR-based metabolomics analysis. *Journal of Translational Medicine* **20**, 275 (2022).

52 Mao, T. *et al.* Elevated serum β-hydroxybutyrate, a circulating ketone metabolite, accelerates colorectal cancer proliferation and metastasis via ACAT1. *Oncogene* **42**, 1889-1899 (2023).

53 Saraon, P. *et al.* Evaluation and prognostic significance of ACAT1 as a marker of prostate cancer progression. *Prostate* **74**, 372-380 (2014).

54 Goudarzi, A. The recent insights into the function of ACAT1: A possible anti-cancer therapeutic target. *Life Sci* **232**, 116592 (2019).

55 DeGuzman, A., Lorenson, M. Y. & Walker, A. M. Bittersweet: relevant amounts of the common sweet food additive, glycerol, accelerate the growth of PC3 human prostate cancer xenografts. *BMC Res Notes* **15**, 101 (2022).

56 Lounis, M. A. *et al.* Elevated Expression of Glycerol-3-Phosphate Phosphatase as a Biomarker of Poor Prognosis and Aggressive Prostate Cancer. *Cancers (Basel)* **13** (2021).

57 Houshmand, M. *et al.* Dihydroorotate dehydrogenase inhibition reveals metabolic vulnerability in chronic myeloid leukemia. *Cell Death & Disease* **13**, 576 (2022).

58 Huang, J. *et al.* Serum metabolomic profiling of prostate cancer risk in the prostate, lung, colorectal, and ovarian cancer screening trial. *Br J Cancer* **115**, 1087-1095 (2016).

59 Huang, J., Zhao, B., Weinstein, S. J., Albanes, D. & Mondul, A. M. Metabolomic profile of prostate cancer-specific survival among 1812 Finnish men. *BMC Medicine* **20**, 362 (2022).

60 Hanigan, M. H. Gamma-glutamyl transpeptidase: redox regulation and drug resistance. *Adv Cancer Res* **122**, 103-141 (2014).

61 Kawakami, K. *et al.* Gamma-glutamyltransferase activity in exosomes as a potential marker for prostate cancer. *BMC Cancer* **17**, 316 (2017).

62 Hanigan, M. H., Gallagher, B. C., Townsend, D. M. & Gabarra, V. Gamma-glutamyl transpeptidase accelerates tumor growth and increases the resistance of tumors to cisplatin in vivo. *Carcinogenesis* **20**, 553-559 (1999).

63 Zavala, V. A. *et al.* Cancer health disparities in racial/ethnic minorities in the United States. *Br J Cancer* **124**, 315-332 (2021).

64 Giaquinto, A. N. *et al.* Cancer statistics for African American/Black People 2022. *CA: A Cancer Journal for Clinicians* **72**, 202-229 (2022).

65 Joshi, A. D. *et al.* An epidemiological introduction to human metabolomic investigations. *Trends in Endocrinology & Metabolism* **34**, 505-525 (2023).

66 Rathbone, J., Hoffmann, T. & Glasziou, P. Faster title and abstract screening? Evaluating Abstrackr, a semi-automated online screening program for systematic reviewers. *Systematic Reviews* **4**, 80 (2015).

67 hmdbQuery: utilities for exploration of human metabolome database. v. R package version 1.28.0 (2025).

68 Kim, S. *et al.* PubChem 2025 update. *Nucleic Acids Research* **53**, D1516-D1525 (2024).

69 Langan, D. *et al.* A comparison of heterogeneity variance estimators in simulated random-effects meta-analyses. *Research Synthesis Methods* **10**, 83-98 (2019).

70 Morze, J. *et al.* Metabolomics and Type 2 Diabetes Risk: An Updated Systematic Review and Meta-analysis of Prospective Cohort Studies. *Diabetes Care* **45**, 1013-1024 (2022).

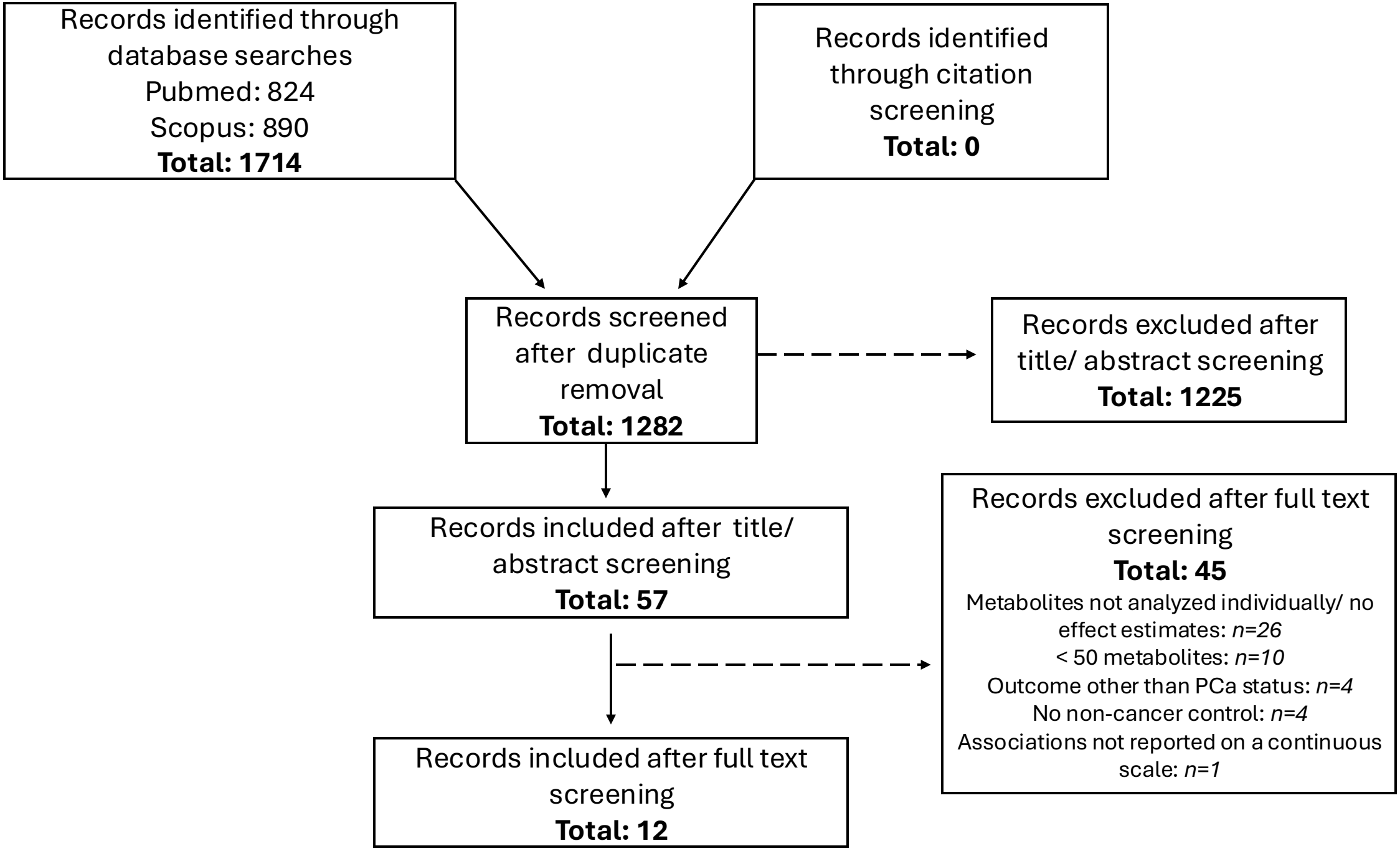
71 de Matos Simoes, R., Tripathi, S. & Emmert-Streib, F. Organizational structure and the periphery of the gene regulatory network in B-cell lymphoma. *BMC Systems Biology* **6**, 38 (2012).

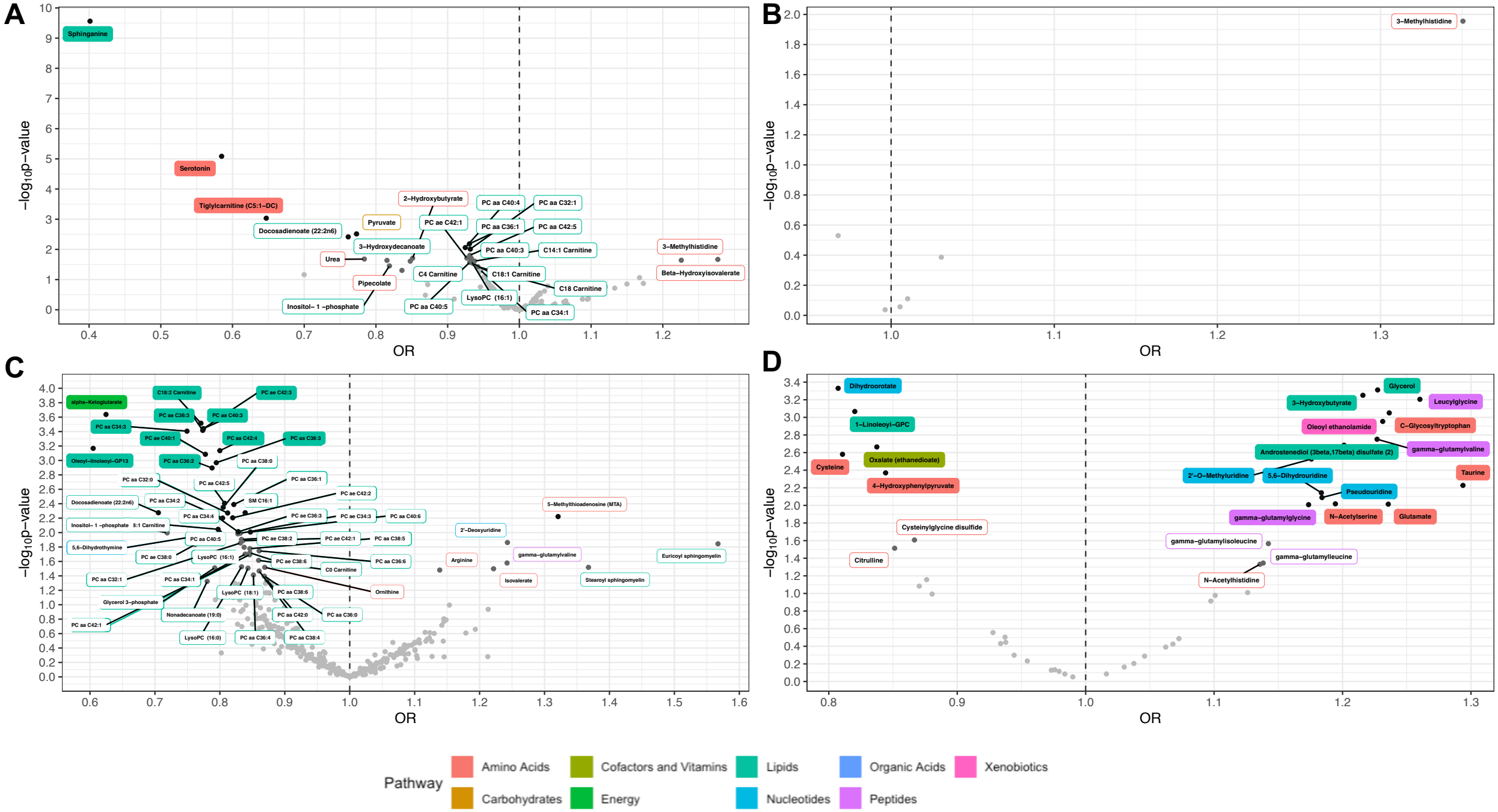
72 Stroup, D. F. *et al.* Meta-analysis of Observational Studies in EpidemiologyA Proposal for Reporting. *JAMA* **283**, 2008-2012 (2000).

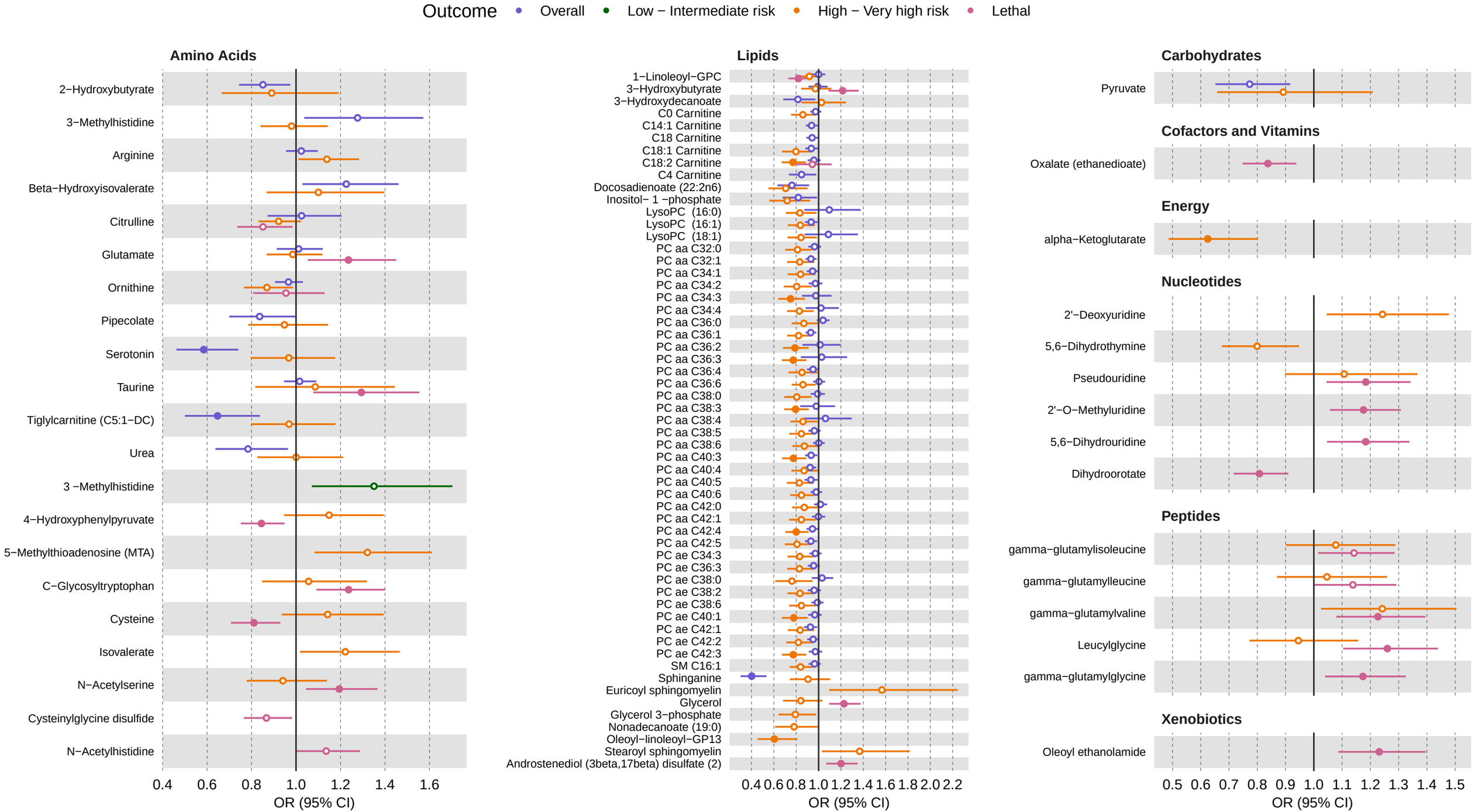
73 Page, M. J. *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **372**, n71 (2021).

74 Wishart, D. S. *et al.* DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res* **46**, D1074-d1082 (2018).

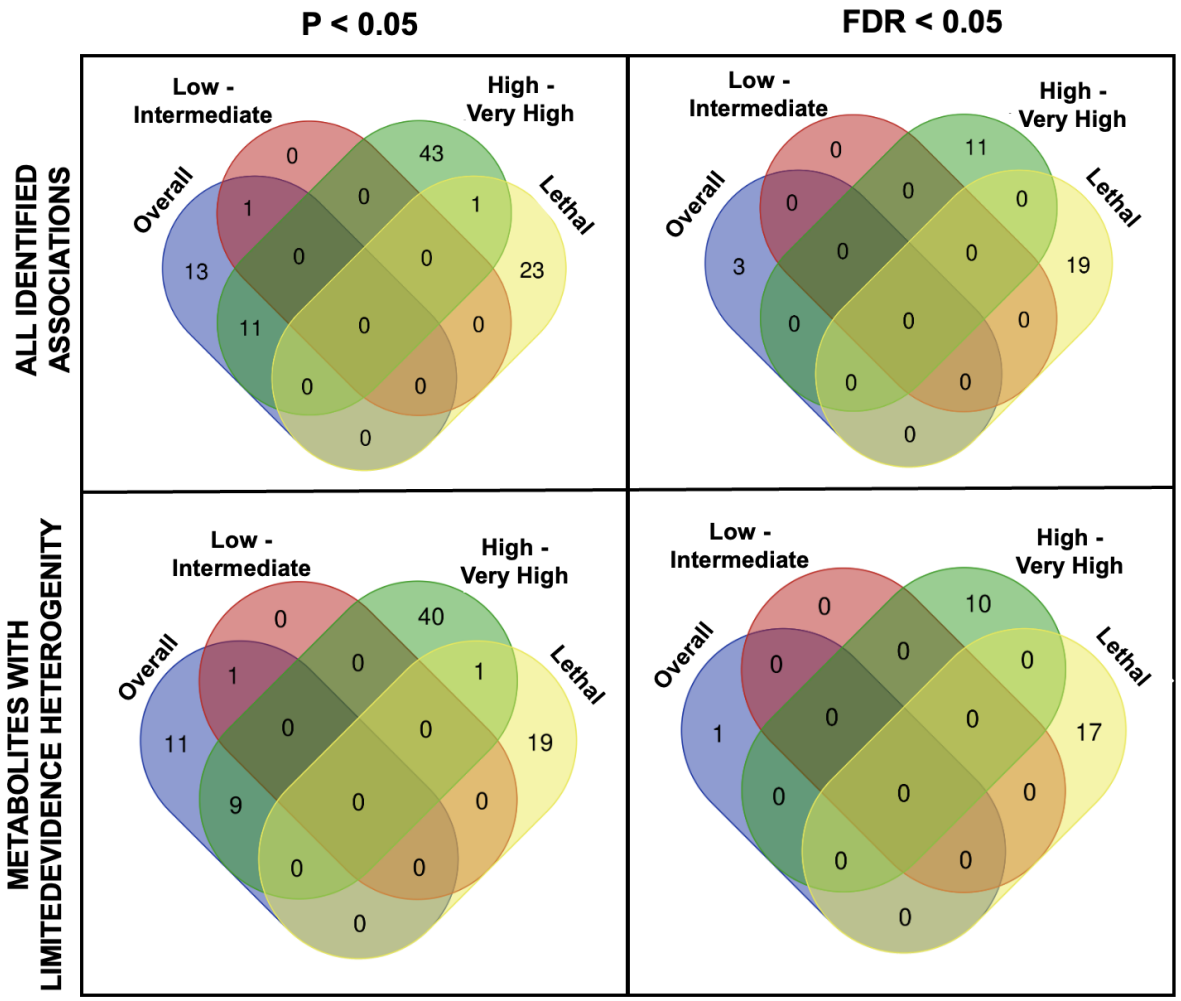
75 Östman, J. R. *et al.* Identification of prediagnostic metabolites associated with prostate cancer risk by untargeted mass spectrometry-based metabolomics: A case-control study nested in the Northern Sweden Health and Disease Study. *Int J Cancer* **151**, 2115-2127 (2022).


Records identified through database searches
Pubmed: 824
Scopus: 890
Total: 1714
Records identified through citation screening
Total: 0
Records screened after duplicate removal
Total: 1282
Records excluded after title/ abstract screening
Total: 1225
Records included after title/ abstract screening
Total: 57
Records excluded after full text screening
Total: 45
Metabolites not analyzed individually/ no effect estimates: n=26
< 50 metabolites: n=10
Outcome other than PCa status: n=4
No non-cancer control: n=4
Associations not reported on a continuous scale: n=1
Records included after full text screening
Total: 12


A
Sphinganine
Serotonin
Tiglylcarnitine (C5:1–DC)
Docosadienoate (22:2n6)
Pyruvate
2–Hydroxybutyrate
3–Hydroxydecanoate
Urea
Pipecolate
Inositol– 1 –phosphate
PC ae C42:1
PC aa C40:4
PC aa C32:1
PC aa C36:1
PC aa C42:5
PC aa C40:3
C14:1 Carnitine
C4 Carnitine
C18:1 Carnitine
C18 Carnitine
LysoPC (16:1)
PC aa C40:5
PC aa C34:1
3–Methylhistidine
Beta–Hydroxyisovalerate
−log10p−value
OR
B
3–Methylhistidine
−log10p−value
OR
C
alpha–Ketoglutarate
C18:2 Carnitine
PC ae C42:3
PC aa C36:3
PC aa C40:3
PC aa C34:3
PC aa C40:1
PC aa C42:4
PC aa C38:3
Oleoyl–linoleoyl–GP13
PC aa C36:2
PC aa C38:0
PC aa C32:0
PC aa C42:5
PC aa C36:1
Docosadienoate (22:2n6)
PC aa C34:2
SM C16:1
PC ae C42:2
PC aa C34:4
PC ae C36:3
PC aa C34:3
PC aa C40:6
Inositol– 1 –phosphate
8:1 Carnitine
PC aa C40:5
PC ae C38:2
PC ae C42:1
PC aa C38:5
5,6–Dihydrothymine
PC ae C38:0
LysoPC (16:1)
PC ae C38:6
PC aa C36:6
C0 Carnitine
PC aa C32:1
PC aa C34:1
LysoPC (18:1)
PC aa C38:6
Glycerol 3–phosphate
Nonadecanoate (19:0)
PC aa C42:0
PC aa C36:0
PC aa C42:1
LysoPC (16:0)
PC aa C36:4
PC aa C38:4
Ornithine
5–Methylthioadenosine (MTA)
2'–Deoxyuridine
Arginine
gamma–glutamylvaline
Isovalerate
Stearoyl sphingomyelin
Euricoyl sphingomyelin
−log10p−value
OR
D
Dihydroorotate
1–Linoleoyl–GPC
Oxalate (ethanedioate)
Cysteine
4–Hydroxyphenylpyruvate
Cysteinylglycine disulfide
Citrulline
Glycerol
3–Hydroxybutyrate
Leucylglycine
Oleoyl ethanolamide
C–Glycosyltryptophan
Androstenediol (3beta,17beta) disulfate (2)
gamma–glutamylvaline
2'–O–Methyluridine
5,6–Dihydrouridine
Taurine
Pseudouridine
gamma–glutamylglycine
N–Acetylserine
Glutamate
gamma–glutamylisoleucine
gamma–glutamylleucine
N–Acetylhistidine
−log10p−value
OR
Pathway
Amino Acids
Cofactors and Vitamins
Lipids
Organic Acids
Xenobiotics
Carbohydrates
Energy
Nucleotides
Peptides

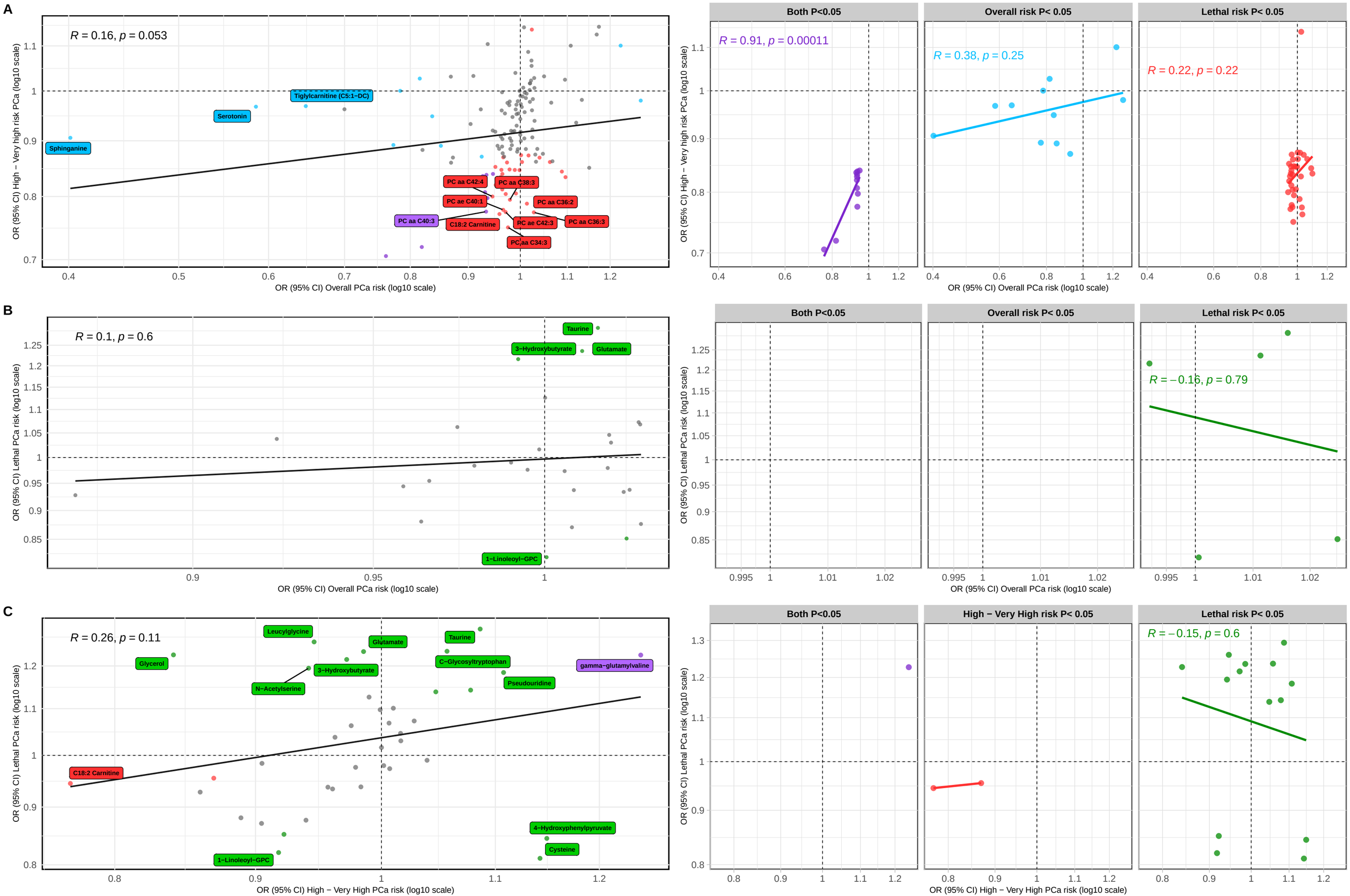

Outcome
Overall
Low − Intermediate risk
High − Very high risk
Lethal
Amino Acids
2−Hydroxybutyrate
3−Methylhistidine
Arginine
Beta−Hydroxyisovalerate
Citrulline
Glutamate
Ornithine
Pipecolate
Serotonin
Taurine
Tiglylcarnitine (C5:1−DC)
Urea
3 −Methylhistidine
4−Hydroxyphenylpyruvate
5−Methylthioadenosine (MTA)
C−Glycosyltryptophan
Cysteine
Isovalerate
N−Acetylserine
Cysteinylglycine disulfide
N−Acetylhistidine
0.4 0.6 0.8 1.0 1.2 1.4 1.6
OR (95% CI)
Lipids
1−Linoleoyl−GPC
3−Hydroxybutyrate
3−Hydroxydecanoate
C0 Carnitine
C14:1 Carnitine
C18 Carnitine
C18:1 Carnitine
C18:2 Carnitine
C4 Carnitine
Docosadienoate (22:2n6)
Inositol− 1 −phosphate
LysoPC (16:0)
LysoPC (16:1)
LysoPC (18:1)
PC aa C32:0
PC aa C32:1
PC aa C34:1
PC aa C34:2
PC aa C34:3
PC aa C34:4
PC aa C36:0
PC aa C36:1
PC aa C36:2
PC aa C36:3
PC aa C36:4
PC aa C36:6
PC aa C38:0
PC aa C38:3
PC aa C38:4
PC aa C38:5
PC aa C38:6
PC aa C40:3
PC aa C40:4
PC aa C40:5
PC aa C40:6
PC aa C42:0
PC aa C42:1
PC aa C42:4
PC aa C42:5
PC ae C34:3
PC ae C36:3
PC ae C38:0
PC ae C38:2
PC ae C38:6
PC ae C40:1
PC ae C42:1
PC ae C42:2
PC ae C42:3
SM C16:1
Sphinganine
Euricoyl sphingomyelin
Glycerol
Glycerol 3−phosphate
Nonadecanoate (19:0)
Oleoyl−linoleoyl−GP13
Stearoyl sphingomyelin
Androstenediol (3beta,17beta) disulfate (2)
0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2
OR (95% CI)
Carbohydrates
Pyruvate
Cofactors and Vitamins
Oxalate (ethanedioate)
Energy
alpha−Ketoglutarate
Nucleotides
2'−Deoxyuridine
5,6−Dihydrothymine
Pseudouridine
2'−O−Methyluridine
5,6−Dihydrouridine
Dihydroorotate
Peptides
gamma−glutamylisoleucine
gamma−glutamylleucine
gamma−glutamylvaline
Leucylglycine
gamma−glutamylglycine
Xenobiotics
Oleoyl ethanolamide
0.5 0.6 0.7 0.8 0.9 1.0 1.1 1.2 1.3 1.4 1.5
OR (95% CI)

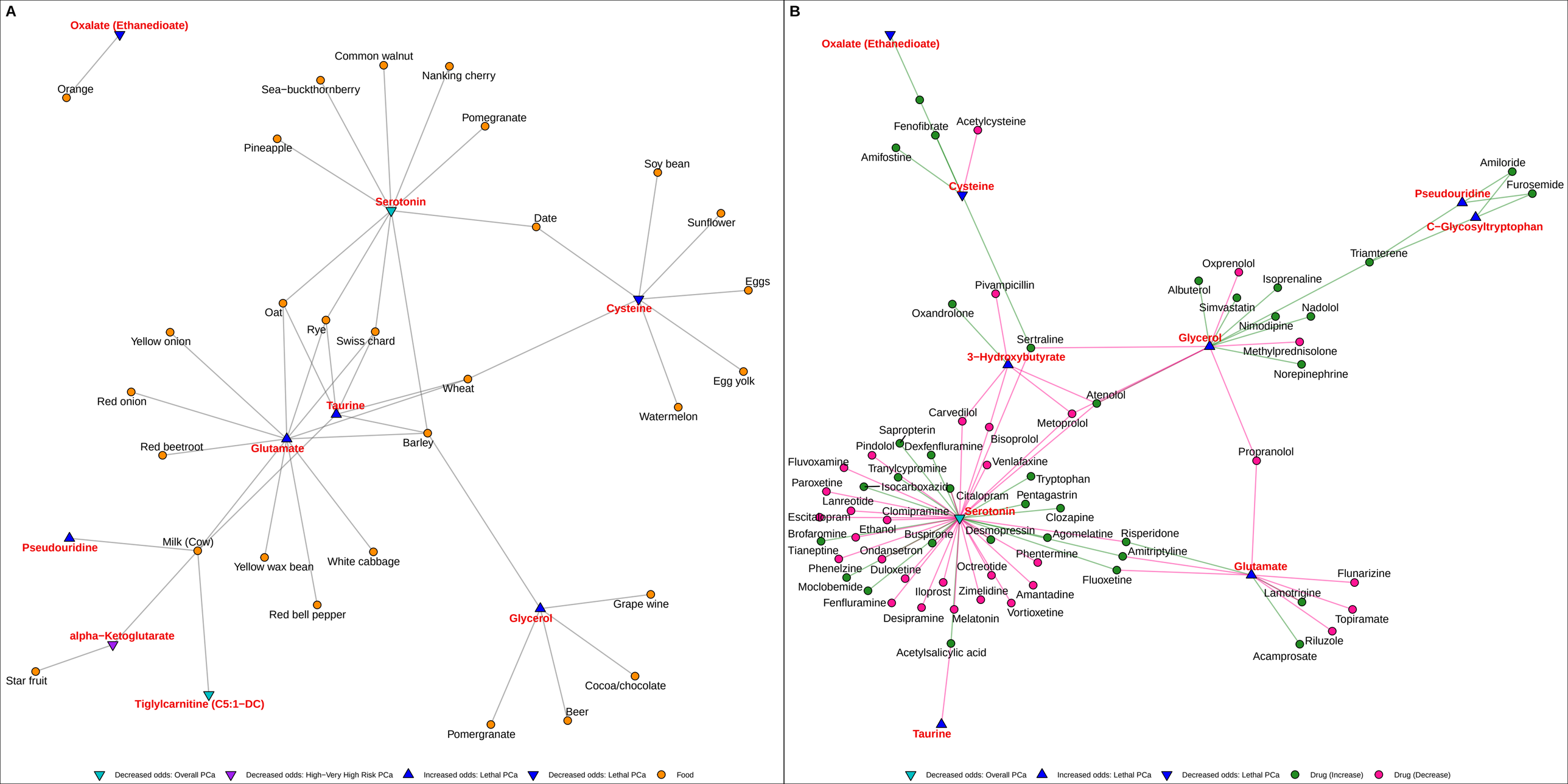
A



B

Outcome	Nmets	Median I^2	Nmets I^2=0 (%)	Nmets Q.value <0.05 (%)
All tested metabolites				
Overall	150	34.01	63 (42)	31 (21)
Low – Intermediate	6	35.04	2 (33.3)	0 (0)
High – Very High	391	6.11	190 (49)	52 (13)
Lethal	48	0	25 (52)	11 (23)
Nominally significant metabolites				
Overall	25	0	16 (64)	4 (16)
Low – Intermediate	1	0	1 (100)	0 (0)
High – Very High	55	12.44	23 (41.82)	5 (9.09)
Lethal	24	0	14 (58.3)	4 (16.67)
FDR significant metabolites				
Overall	3	91.72	0	2 (66%)
Low – Intermediate	0	-	-	-
High – Very High	11	33.39	2 (18.18)	1 (9.09)
Lethal	19	0	13 (68.42)	2 (10.53)


A
R = 0.16, p = 0.053
Tiglylcarnitine (C5:1−DC)
Serotonin
Sphinganine
PC aa C42:4
PC aa C38:3
PC ae C40:1
PC aa C36:2
PC aa C40:3
C18:2 Carnitine
PC ae C42:3
PC aa C36:3
PC aa C34:3
OR (95% CI) High − Very high risk PCa (log10 scale)
OR (95% CI) Overall PCa risk (log10 scale)
Both P<0.05
R = 0.91, p = 0.00011
Overall risk P< 0.05
R = 0.38, p = 0.25
Lethal risk P< 0.05
R = 0.22, p = 0.22
B
R = 0.1, p = 0.6
Taurine
3−Hydroxybutyrate
Glutamate
1−Linoleoyl−GPC
OR (95% CI) Lethal PCa risk (log10 scale)
R = −0.16, p = 0.79
C
R = 0.26, p = 0.11
Leucylglycine
Glutamate
Taurine
Glycerol
3−Hydroxybutyrate
C−Glycosyltryptophan
gamma−glutamylvaline
N−Acetylserine
Pseudouridine
C18:2 Carnitine
4−Hydroxyphenylpyruvate
Cysteine
1−Linoleoyl−GPC
OR (95% CI) High − Very High PCa risk (log10 scale)
High − Very High risk P< 0.05
R = −0.15, p = 0.6


A
Oxalate (Ethanedioate)
Orange
Common walnut
Nanking cherry
Sea−buckthornberry
Pomegranate
Pineapple
Soy bean
Serotonin
Date
Sunflower
Eggs
Cysteine
Oat
Rye
Swiss chard
Yellow onion
Egg yolk
Wheat
Red onion
Taurine
Watermelon
Red beetroot
Glutamate
Barley
Pseudouridine
Milk (Cow)
Yellow wax bean
White cabbage
Red bell pepper
Grape wine
Glycerol
alpha−Ketoglutarate
Star fruit
Cocoa/chocolate
Tiglylcarnitine (C5:1−DC)
Beer
Pomergranate
Decreased odds: Overall PCa
Decreased odds: High−Very High Risk PCa
Increased odds: Lethal PCa
Decreased odds: Lethal PCa
Food
B
Oxalate (Ethanedioate)
Fenofibrate
Acetylcysteine
Amifostine
Cysteine
Amiloride
Pseudouridine
Furosemide
C−Glycosyltryptophan
Triamterene
Oxprenolol
Isoprenaline
Albuterol
Pivampicillin
Simvastatin
Nadolol
Oxandrolone
Nimodipine
Sertraline
Glycerol
Methylprednisolone
3−Hydroxybutyrate
Norepinephrine
Atenolol
Carvedilol
Metoprolol
Sapropterin
Bisoprolol
Pindolol
Dexfenfluramine
Propranolol
Fluvoxamine
Venlafaxine
Tranylcypromine
Tryptophan
Paroxetine
Isocarboxazid
Citalopram
Pentagastrin
Lanreotide
Clomipramine
Serotonin
Escitalopram
Clozapine
Brofaromine
Ethanol
Buspirone
Desmopressin
Agomelatine
Risperidone
Tianeptine
Ondansetron
Phentermine
Amitriptyline
Phenelzine
Duloxetine
Octreotide
Glutamate
Flunarizine
Moclobemide
Iloprost
Zimelidine
Amantadine
Fluoxetine
Lamotrigine
Fenfluramine
Vortioxetine
Desipramine
Melatonin
Topiramate
Riluzole
Acetylsalicylic acid
Acamprosate
Taurine
Decreased odds: Overall PCa
Increased odds: Lethal PCa
Decreased odds: Lethal PCa
Drug (Increase)
Drug (Decrease)

❑ **Methods and Statistical Approaches**

- Detail scaling, standardization and transformation procedures applied.

❑ **Sample Information**

- Provide information on fasting status of participants.
- Report on the time of sample collection in relation to outcome.

❑ **Metabolite Descriptions**

- Report commonly used metabolite identifiers (HMDB, PubChem, KEGG, platform specific IDs).
- Use conventional spelling for metabolite names and conventional formatting for abbreviations, including structural abbreviations within metabolite names.
- State the meaning of used abbreviations.
- State clearly if composite metabolite biomarkers are utilized and which metabolites are considered in these measures.
- Explain how metabolite pathways were assigned.

❑ **Outcome Measurement**

- Provide quantitative measures of time until diagnosis, including a measure of average time until diagnosis (eg., mean, medianand measures of spread (eg., range, interquartile range).

❑ **Reporting of Findings**

- Provide all association results, including non-significant findings.
- Provide association results on a continuous scale that is comparable across populations.
- Report unadjusted and adjusted results, including information on all adjusted variables.
- Clearly state any cohort characteristics that may result in findings being less generalizable across populations.