Metabolome of healthy and unhealthy dietary patterns

Writing Assignment
Utrecht University
Biology of Disease

Jonne Damen, BSc
5823374

Supervisor: Marjolein Harbers, MSc
Examiner: Prof. Yvonne van der Schouw, PhD
Second reviewer: Prof. Monique Verschuren, PhD

Layman's summary

Een gezond dieet heeft een positief effect op de gezondheid. Verschillende voedingspatronen, waaronder het Mediterraans dieet, het *DASH* dieet, het *New Nordic Diet*, en het vegetarische dieet behoren tot de gezonde diëten. Deze diëten hebben allemaal hetzelfde uitgangspunt, gezien er in deze diëten voornamelijk volkoren granen, groenten en fruit, en onverzadigde vetten wordt gegeten, terwijl verzadigde vetten, zout, en bewerkte voeding achterwege wordt gelaten. Onderzoeken laten zien dat gezond eten een positieve invloed heeft op verschillende ziekten, zoals hart- en vaatziekten, suikerziekte, en bepaalde vormen van kanker. Waarom gezonde voeding een positief effect heeft op de gezondheid wordt nog niet volledig begrepen. Om dit te onderzoeken kunnen metabolieten worden gemeten. Metabolieten zijn de tussen- of eindproducten van voedingsstoffen die door de stofwisseling in het lichaam ontstaan. Deze kunnen onder andere in het bloed en de urine worden gemeten. Door metabolieten te meten bij mensen met een gezond voedingspatroon en te vergelijken met mensen met een relatief ongezond voedingspatroon, kan worden onderzocht of er verschillen zijn in de samenstelling van de metabolieten tussen gezonde en ongezonde diëten. In dit artikel wordt een overzicht gegeven van de resultaten van 20 klinische studies die de verschillen tussen een gezond en ongezond dieet hebben onderzocht door metabolieten te meten in het bloed of de urine. Er worden veel verschillen gezien in de samenstelling van de klassen waartoe de metabolieten horen tussen een gezond en een ongezond dieet. Zo werden er bijvoorbeeld meer metabolieten gevonden die geassocieerd worden met groente- en fruitinname bij de gezonde diëten. Een aantal metabolieten werd genoemd in meerdere artikelen, wat een associatie met ofwel een gezond ofwel een ongezond dieet waarschijnlijker maakt. Zo werd een metaboliet die geassocieerd wordt met de inname van rood vlees in de ongezonde diëten gevonden. Ook is gekeken of de metabolieten die in de analyse zijn gevonden bij bepaalde *pathways* (groepen metabolieten die in een bepaalde volgorde een bepaalde biologische functie vervullen) horen en of er verschillen kunnen worden gevonden tussen de *pathways* die geassocieerd zijn met gezonde en ongezonde voedingspatronen. Er is een aantal *pathways* gevonden dat geassocieerd is met ofwel een gezond ofwel een ongezond voedingspatroon. Deze resultaten geven aan dat er indicaties zijn dat er verschillen zijn tussen de metaboliet samenstelling van gezonde en ongezonde diëten. Er is meer onderzoek nodig om te begrijpen wat deze verschillen in metaboliet samenstelling betekenen voor de gezondheid.

Abstract

Background and aim: Healthy diet has been shown to have a positive impact on health status. However, underlying mechanisms are not completely understood. Metabolomics, in which metabolites in biofluids are analyzed and quantified, can be utilized to uncover metabolomic signatures associated with healthy and unhealthy dietary patterns. Utilizing metabolomics can improve understanding of the effects of healthy and unhealthy dietary patterns on metabolic processes and eventually provide a greater understanding of the underlying mechanisms in which healthy diet contributes to health status. This review aimed to conduct a systematic search to provide an overview of relevant literature and to gather and analyze available data on the metabolomic signatures associated with healthy and unhealthy dietary patterns.

Methods: A systematic search was conducted to identify randomized controlled trials investigating endogenous metabolome in plasma, serum, or urine in subjects consuming a healthy dietary pattern compared to subjects consuming a relatively unhealthier dietary pattern. Metabolites reported as significant were extracted and used to compare metabolite class distributions between dietary patterns. Metabolites reported in multiple studies were highlighted. Pathway over-representation analyses of metabolites associated with a healthy or unhealthy dietary pattern were performed.

Results: 20 relevant randomized controlled trials were identified and article characteristics were summarized. Differences in metabolite class distributions were found between dietary patterns and metabolite class compositions were dependent on sample specimen. Several metabolites were associated with either a healthy or unhealthy dietary pattern in multiple articles. Pathway over-representation analyses show several metabolic pathways significantly enriched in healthy or unhealthy dietary patterns.

Conclusions: Differing metabolomic signatures were found to be associated with healthy and unhealthy dietary patterns. Identified differences do not provide a unanimous and clear consensus regarding connections to health status. Further research is needed to understand correlations between metabolomic profiles and health status. More comparable study parameters will enable future meta-analyses to more accurately uncover metabolomic profiles associated with dietary patterns.

Introduction

The metabolization of foods introduces metabolite compounds in the body and is thereby able to alter endogenous metabolic processes[1]. Therefore, healthy dietary pattern may influence biological processes and contribute to health status. Healthy diet has been shown to positively impact and prevent chronic diseases, such as type 2 diabetes mellitus[2] and cardiovascular diseases[3–5], and high-quality dietary patterns may be associated with a lower risk of certain cancer types[6–8]. Healthy dietary patterns include, amongst others, the Mediterranean diet, Dietary Approaches to Stop Hypertension (DASH), New Nordic Diet, and vegetarian diets. While these diets can differ in recommended servings of food and beverages, they share similar characteristics, generally consuming more whole grains, fruits and vegetables, and unsaturated fats, while avoiding saturated fats, salt, and processed foods[9]. However, the underlying mechanisms in which dietary pattern contributes to health status are not completely understood[10,11].

To provide more insights into the underlying mechanisms in which dietary pattern contributes to health status, metabolomics can be utilized. Metabolomics provides an explorative and untargeted approach to nutrition research, as it doesn't require a hypothesis based on a priori knowledge[1]. In metabolomics, metabolites in biofluids, tissues, or cells are analyzed and quantified[10,12]. Metabolomics represents the most downstream stage of the post-genomic technologies and therefore reflects the end products of genetic, epigenetic, and environmental stimuli[1,12]. Therefore, metabolomic profiles can be used as a biomarker to provide an unbiased measure for diet adherence[11,13–16]. However, metabolites are not merely biomarkers of gene and protein activity, but they have far-reaching regulatory activity, as the metabolome can interact with and actively modulate all other -omics levels, making metabolites direct modulators of biological processes and phenotypes[12]. Therefore, integrating metabolomics with nutritional science provides a deeper insight into the relationship between metabolites and health status[10]. Besides reflecting dietary intake, metabolomics can also reflect other sources of variability in metabolism, such as diet-gut microbiome interplay and effects of genetic variation on metabolism, enabling objective assessment of metabolic responses to dietary patterns[11]. Therefore, the use of metabolomics in nutritional research may improve the mechanistic understanding of the impact of diet on human health[10]. Metabolomics can be utilized to compare metabolomic signatures of healthy dietary patterns to unhealthy dietary patterns. This will advance understanding of the effects of either healthy or unhealthy diet on metabolic processes and will provide a greater understanding of the underlying mechanisms in which healthy diet contributes to health status.

This review aimed to gather and analyze available data on the metabolomic signatures associated with healthy dietary patterns compared to unhealthy dietary patterns. Therefore, the aim of this review was to systematically search for and identify relevant published articles and to map out article characteristics to provide an overview of dietary patterns and corresponding metabolomic profiles. Thereby, this review aimed to provide a starting point for further research into the effects of healthy dietary patterns on metabolic and biological processes and health status.

Methods

Search strategy

The search was performed using PubMed on January 11th, 2022. The search string was constructed using two groups of search terms, joined by an AND operator. The first group of search terms includes variations on dietary patterns: *"dietary pattern*" OR "diet pattern*" OR "mediterranean diet" OR DASH OR "dietary approaches to stop hypertension" OR "healthy eating index" OR HEI OR "new nordic diet" OR "healthy nordic diet" OR "vegetarian diet" OR "vegan diet".* The second group of search terms aims to include articles using a metabolomics approach: *metabolome OR metabonome*. No limits or filters were applied in the search.

Eligibility criteria

Studies were eligible for inclusion if they were randomized controlled trials that described a type of healthy dietary pattern consumed in Western countries in combination with measurements of endogenous metabolome in plasma, serum, or urine. Studies that were reviews, study protocols, commentaries, other types of clinical trials besides randomized controlled trials, described animal or cell research, were not written in English, or that had no full-text available, were excluded. Furthermore, studies having a non-metabolomics approach, solely describing the gut or microbiome metabolome, or making use of a feces specimen for metabolomics, associating metabolites to one food or food group instead of a healthy dietary pattern, lacking data relevant to data extraction, or researching other variables alongside dietary pattern (e.g., physical activity), were also excluded.

Record screening

During the first stage of screening, one reviewer performed assessed relevance of the titles and abstracts using Rayyan[17] based on the inclusion- and exclusion criteria mentioned above. Articles that were deemed eligible based on their title and abstract were included in the second stage of screening. During this stage, one reviewer assessed full-text articles and excluded the reports if they did not meet the inclusion- and exclusion criteria. Of all remaining articles, potential bias was determined using the Revised Cochrane risk-of-bias tool for randomized trials[18]. A visual summary of the risk of bias of all included articles is shown in **supplementary figure 1**.

Data extraction and analysis

One reviewer collected data from the included studies. For every report, metabolites statistically significantly associated with a dietary pattern were listed. As the significance thresholds were defined by article authors, the level of significance of listed metabolites was reported as well. Other variables for which data were sought included the type of dietary pattern studied, study design, study population size and characteristics (health status, age, and sex), metabolomics platform used, and specimen type. Data was presented in a table summary. Reported significant metabolites were manually assigned to metabolite classes based on matching records in the Human Metabolome Database. The total number of extracted significant metabolites, as well as unidentified metabolites, metabolites that could be matched to a Human Metabolome Database entry, and the number of metabolites associated with healthy and unhealthy dietary patterns was determined. Furthermore, relative contributions of metabolite classes to the total number of metabolites associating with healthy and unhealthy dietary patterns were determined and visualized using bidirectional bar graphs in R (version 4.0.2)[19] using ggplot2[20]. These results were also split for

metabolites identified in urine samples and blood samples. Moreover, metabolites reported to be significantly associated with dietary pattern in multiple studies were highlighted in a table summary. Lastly, metabolite pathway over-representation analyses were performed and visualized using MetaboAnalyst 5.0[21]. Here, metabolites associating with either healthy or unhealthy dietary patterns were entered into the database to test if a group of metabolites belonging to a curated pathway was represented within the entered metabolite list more than expected by chance, thereby making a pathway enriched. The pathway library that was tested against was the Homo sapiens KEGG library. The enrichment method that was chosen was the hypergeometric test. Enrichment values below a False Discovery Rate adjusted p-value of <0.1 were considered statistically significant. Furthermore, for pathway topology analysis to help identify importance of positions of entered metabolites within a given pathway, relative-betweenness centrality was chosen. Results were visualized using the -log10 transformed p-values and the pathway impact was calculated from pathway topology analysis.

Results

The search yielded 305 potentially relevant articles in PubMed. All 305 articles were included in the first round of relevance screening of the titles and abstract in Rayyan. All clinical trial articles that mentioned measuring metabolites to examine effect of a dietary pattern on metabolic composition or to determine diet adherence were included for the next round of full-text screening. After the first round of relevance screening, 225 articles were excluded, and 80 articles were included in full-text screening. After full text screening, 20 randomized controlled trials were selected for result extraction. The flow of the study selection process is shown in **figure 1**.

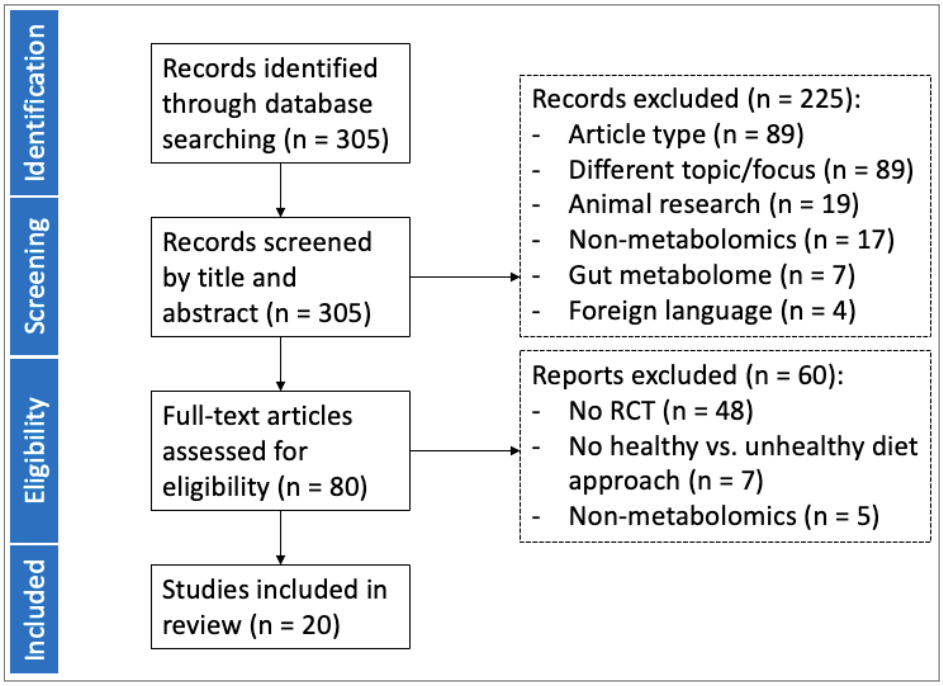


Figure 1 – Flow diagram of study selection process.

Selected randomized controlled trials show differences in design parameters

Article characteristics, such as dietary patterns that were compared, study design, and number of significant extracted metabolites, of all 20 articles were summarized in **table 1**. Different types of healthy dietary patterns were studied, as 6 Mediterranean-style[22–27],

Table 1 – **Summary of included articles.** # reflects the number of significant metabolites extracted and * reflects the type of significance reported.

Year	Author	Dietary pattern	Design	N	Study population	Method	Specimen	# and *	Ref
2020	Meslier *et al.*	MED vs. CONT	Parallel 8 week Free living, food diary	- 43 MED - 39 CONT	- Healthy sedentary overweight/obese - Age: 43±12 - 43 female, 39 male	UHPLC/MS Untargeted	Urine and serum	14 Loading >0.04 on component of PLS-DA	22
2020	Djekic *et al.*	Vegetarian vs. Meat	Crossover 4 week Some food provided, food diary	31	- IHD PCI treated patients - Age: 61-70 - 94% men	HPLC-QTOF-MS Untargeted	Plasma	29 p-adj < 0.05	28
2021	Pouraf-shar *et al.*	DASH vs. CONT	Post-hoc analysis 2 week Some on-site feeding visits, food diary	- 10 DASH - 10 CONT	- Unmedicated HT - Age: DASH 46.1±9, CONT 42.4±6.3 - 13 female, 7 male	GC/MS Untargeted	Plasma and urine	12 p-adj < 0.20	29
2021	Gürdeniz *et al.*	HND vs. CONT	Parallel 18/24 week Some food provided, food diary	- 98 HND - 71 CONT	- Metabolic syndrome - Age: HND 54.5±8.09, CONT 54.7±8.48 - ~65% female	UPLC-QTOF-MS Untargeted	Plasma and urine	39 p-value < 0.05	30
2019	Draper *et al.*	Vegan vs. Meat	Follow-up analysis, crossover 3 day Food provided	21	- Healthy - Age: 25-45 - 11 female, 10 male	UPLC/MS Targeted	Plasma	15 p-adj < 0.10	31
2021	Kim *et al.*	DASH high sodium vs. CONT high sodium	Multicenter parallel 30 day Food provided	- 199 DASH - 193 CONT	- PreHT/stage 1 HT - Age > 22 - ~55% female	UHPLC/MS Untargeted	Urine	143 p-adj < 0.05	32
2018	Rebholz *et al.*	DASH vs. CONT (vs. fruits and vegetables)	Multicenter 8 week Food provided	- 110 DASH - 108 CONT - (111 fruits and vegetables)	- Healthy - Age > 22 - 154 female, 175 male	LC/MS and GC/MS Untargeted	Serum	94 p-adj < $6.11*10^{-5}$	33
2021	Galié *et al.*	MED vs. non-MED + nuts	Crossover 2 month Food diary	44	- Metabolic syndrome - Age 37-65 - Sex not specified	LC/MS and GC/MS Targeted	Plasma	65 Elastic net regression coefficient	23
2019	Navarro *et al.*	WG vs. RG	Crossover 28 day Food provided	80	- Healthy - Age: 18-45 (mean 29.6) - 40 female 40 male	LC/MS Targeted	Plasma	18 p-adj < 0.05	34
2020	Trimigno *et al.*	NND vs. ADD	Parallel 26 weeks Free living, recipes provided	- 89 NND - 56 ADD 142 analyzed	- Centrally obese - Age 22-60 (mean 42) - 100 female, 45 male	^{1}H-NMR Untargeted	Urine	10 q-value < 0.05 after PLS-DA	35
2016	Khakimov *et al.*	NND vs. ADD	Parallel 26 weeks	- 89 NND - 56 ADD	- Centrally obese - Age 22-60 (mean 42)	GC/MS Untargeted	Plasma	12	36

			Free living, recipes provided		- 100 female, 45 male			p-value < 0.05 after PLS-DA	
2014	Andersen *et al.*	NND vs. ADD	Parallel 26 weeks Free living, recipes provided	- 89 NND - 56 ADD 107 analyzed	- Centrally obese - Age 22-60 (mean 42) - 100 female, 45 male	UPLC-QTOF-MS Untargeted	Urine	30 Selected in PLS-DA	37
2019	Acar *et al.*	NND vs. ADD	Parallel 26 weeks Free living, recipes provided	- 90 NND - 56 ADD	- Centrally obese - Age 22-60 (mean 42) - 100 female, 46 male	UPLC-QTOF-MS Untargeted	Plasma	36 Selected in PLS-DA	38
2015	Vázques-Fresno *et al.*	MED vs. CONT low fat	Multicenter parallel 1- and 3-year follow up Free living	- 41 MED + EVOO - 27 MED + nuts - 30 CONT low fat	- High CVD risk - Age 66.6±5.7 - 28.7% male	^{1}H-NMR Untargeted	Urine	24 VIP > 1.5 in PLS-DA	24
2021	McNairn *et al.*	Vegan vs. Meat	Crossover 24 hour Food provided	8	- Healthy undergraduate students - Age: 20-24 - 6 female, 2 male	UPLC/MS Untargeted	Urine and dried blood spot	26 p-value < 0.05 in urine and dried blood spot	39
2017	Garcia-Perez *et al.*	WHO healthy eating guidelines accordance	Crossover 72 hour Inpatient periods	19	- Healthy - Age: 21-65 - 9 female, 10 male	^{1}H-NMR Untargeted	Urine	28 q-value < 0.05	16
2021	Barber *et al.*	MED vs. Western	Crossover 2 week Food provided, food diary	20	- Healthy - Age: 18-38 - Only male	UPLC-QTOF-MS Untargeted	Urine	11 >1.5-fold increase compared to other diet	25
2015	Bondia-Pons *et al.*	RESMENA vs. CONT	Parallel 2 month Menu plan provided, food diary	- 48 RESMENA - 45 CONT	- Metabolic syndrome - Age: 49±1 - 41 female, 52 male	LC-QTOF-MS Untargeted	Plasma	34 p-value < 0.05	26
2019	Michielsen *et al.*	MED vs. SFA (vs. MUFA)	Parallel 8 week Most food provided, food diary	- 14 MED - 16 SFA - (17 MUFA)	- At risk of developing metabolic syndrome - Age: 45-60 - 27 female, 20 male	^{1}H-NMR Targeted	Serum	52 p-adj < 0.05	27
2019 **	Wellington *et al.*	Prudent (Western at baseline) vs. Western (Prudent at baseline)	Parallel 2 week Food provided, food diary	- 24 Prudent - 18 Western 84 originally randomized	- Healthy - Age: 50±18 Prudent, 43±20 Western - 27 female, 15 male	MI-CE-MS Untargeted	Plasma and urine	18 p-value < 0.05 in 2/3 statistical models and in correlation	40

Abbreviations: MED: Mediterranean diet, CONT: control diet, HP: high performance, UP: ultra-performance, UHP: ultra-high performance, LC: liquid chromatography, GC: gas chromatography, MS: mass spectrometry, QTOF: Quadrupole Time of Flight, PLS-DA: Partial Least Squares Discriminant Analysis, IHD: ischemic heart disease, PCI: percutaneous coronary intervention, GLMM: generalized linear mixed models, HT: hypertension, p-adj: adjusted p-value, HND: Healthy Nordic diet, WG: low-glycemic whole-grain diet, RG: refined grains and added sugars diet, NND: new Nordic diet, ADD: average Danish diet, ^{1}H-NMR: proton nuclear magnetic resonance, EVOO: extra virgin olive oil, CVD: cardiovascular disease, VIP: variable importance in projection, WHO: World Health Organization, RESMENA: metabolic syndrome reduction in Navarra (based on Mediterranean dietary pattern), SFA: saturated fatty acids, MUFA: monounsaturated fatty acids, MI-CE: multisegment injection-capillary electrophoresis. **This report was found to have a higher risk of bias.

5 Nordic[30,35–38], 3 DASH[29,32,33], 3 vegetarian or vegan[28,31,39], and 3 other healthy dietary patterns[16,34,40] were researched in comparison to relatively more unhealthy dietary patterns. Differences in study design, such as intervention duration, can be observed, ranging from interventions of 24 hours to 3 years. Furthermore, 7 studies had a crossover design[16,23,25,28,31,34,39], while 13 studies had a parallel design[22,24,37,38,40,26,27,29,30,32,33,35,36]. Different intervention strategies were used, as several studies chose to provide foods, of which some let study participants eat meals in the study facility and some used food diaries to determine diet adherence. Other studies provided recipes or menu plans to study participants in a free-living environment with regular diet checks based on food diaries. Differences in study size can also be observed, ranging from a study size of 8 to 392 study subjects. The study populations generally consisted of either healthy participants or individuals with cardiovascular risk profiles. Furthermore, except for 2 studies[25,28], female participant contribution was generally equal to- or higher than male participant contribution. The types of metabolomics measuring methods used include liquid chromatography mass spectrometry or gas chromatography mass spectrometry, sometimes used in combination with quadrupole time of flight mass spectrometry, and proton nuclear magnetic resonance mass spectrometry. Furthermore, one study used multisegment injection-capillary electrophoresis mass spectrometry[40]. Four articles used targeted metabolomics[23,27,31,34], while most used an untargeted metabolomics approach. Specimen samples used for metabolomics analysis include plasma, serum, and urine, and one study used dry blood spot specimens for metabolomics analysis[39]. While 4 articles researched the same study population[35–38], the discriminating factors between these studies are the type of sample specimen and the type of metabolomics platform used. The number of extracted significant metabolites ranges from 10 to 143. Furthermore, differences in data analysis methods result in different significance outcomes. Some studies used clustering methods, such as PLS-DA, to discriminate between dietary patterns and then selected metabolites most significantly contributing to dietary pattern discrimination. Other studies determined significant metabolites based on one-by-one differences in measured metabolite concentrations. Different significance thresholds are used in the studies, as some studies reported metabolites that are significant below a nominal p-value, while others reported significance after multiple testing correction. The study performed by Wellington *et al.*[40] selected study subjects after randomization based on their baseline diets, resulting in a higher risk of bias, as visualized in **supplementary figure 1**.

Metabolite class distributions differ between healthy and unhealthy dietary patterns

Of all 20 articles, 710 metabolites were extracted, of which 56 were unidentified. 355 metabolites were reported with a relatively healthy dietary pattern, while 355 were reported with a relatively unhealthy dietary pattern. 18 metabolites were reported by Wellington *et al.*[40] and were not included in further analyses due to the higher risk of bias. Of all remaining 692 extracted metabolites, 433 (340 unique) metabolites could be matched to an entry in the Human Metabolome Database. 227 of these categorized metabolites were found to be associated with healthy dietary patterns, while 206 were found to be associated with unhealthy dietary patterns. **Figure 2** shows the distribution of all identified metabolite classes associated with healthy and unhealthy dietary patterns. Both similarities and differences in metabolite class distributions are apparent between healthy and unhealthy dietary patterns. The two metabolite classes that are the largest contributor to both groups are amino acids and glycerophospholipids. Contributions of amino acids are comparable between healthy (25.2%) and unhealthy (25.7%) dietary patterns, whereas the glycerophospholipids class has

a higher contribution to the unhealthy dietary patterns (22.3%) compared to the healthy dietary patterns (13.3%). Other classes, such as carbohydrates, benzoic acids, and cinnamic acids show a higher contribution to healthy dietary patterns (6.2%, 4.9%, and 4.0%) compared to unhealthy dietary patterns (1.5%, 1,9%, and 0.5%, respectively). Organonitrogen compounds, cofactors and vitamins, carboxylic acids, hydroxy acids, bile acids, ketone bodies, and glycerolipids also show a higher contribution to healthy dietary patterns compared to unhealthy dietary patterns. Conversely, fatty acid esters and imidazopyrimidines have a higher contribution to unhealthy dietary patterns (9.7% and 7.8%) compared to healthy dietary patterns (0.4% and 0.9%, respectively). Furthermore, several metabolite classes are found exclusively in either healthy or unhealthy dietary patterns. For example, sulfuric acids are only found in healthy dietary patterns, making up 5.3% of all metabolites found in healthy dietary patterns, while nucleosides are only found in unhealthy dietary patterns (3.4% of all metabolites).

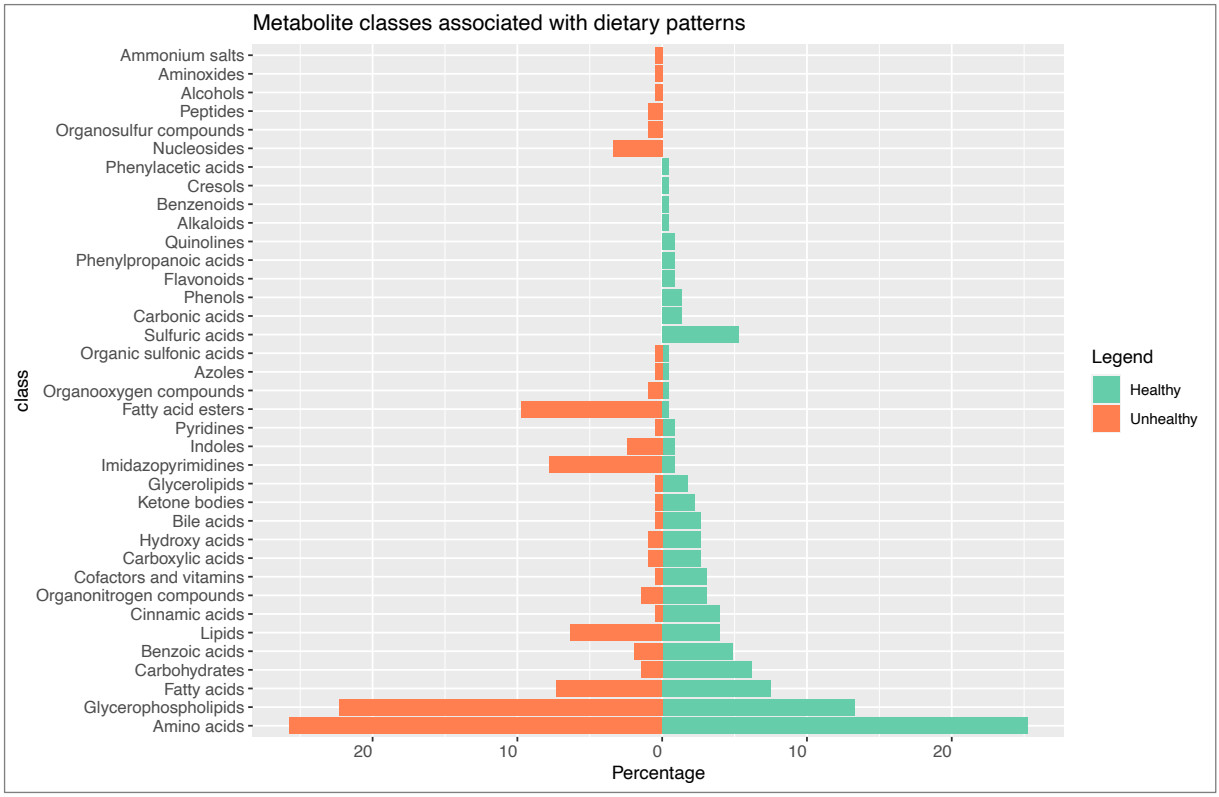


Figure 2 – **Distribution of all identified metabolite classes associated with dietary patterns.** Relative contributions of metabolite classes to the total number of metabolites associated with healthy (green) or unhealthy (orange) dietary patterns are shown. The y-axis shows all identified metabolite classes. The x-axis shows the relative contribution of the metabolite classes to the total number of metabolites, where 0% is found in the middle of the x-axis and relative contributions of metabolite classes to healthy dietary patterns are on the right side and those of unhealthy dietary patterns are on the left side.

Similar differences in metabolite compositions between dietary patterns can also be observed when results are split into metabolites identified in plasma or serum samples and metabolites identified in urine samples (**figure 3A and 3B**). Interestingly, metabolite class differences between dietary patterns are more pronounced in urine samples. Furthermore, glycerophospholipids and glycerolipids were only identified in plasma or serum samples, while many metabolite classes, amongst which phenols, nucleosides, and alcohols, could only be identified in urine samples.

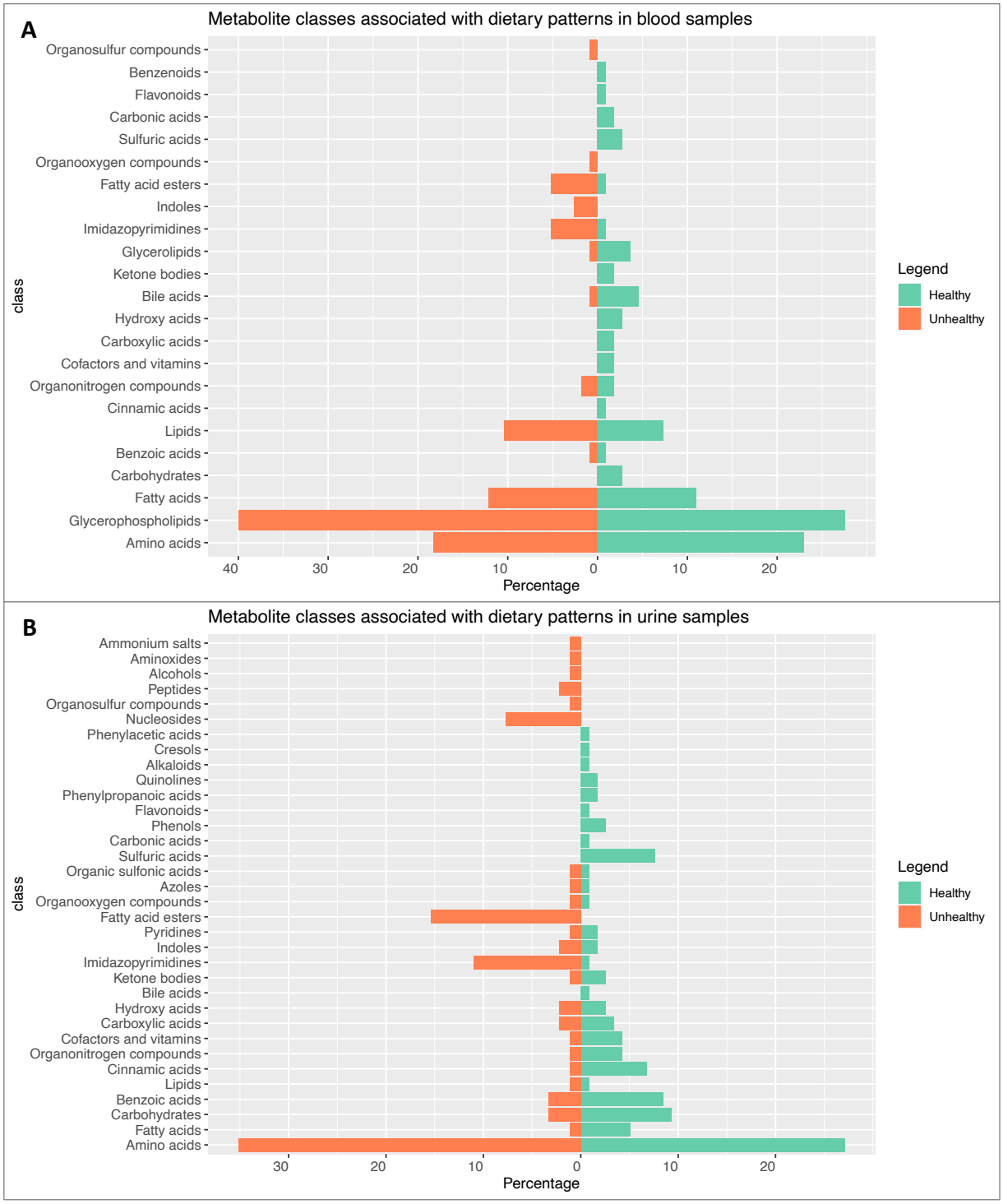


Figure 3 - **Distribution of all identified metabolite classes associated with dietary patterns in (A) blood samples and (B) urine samples.** Relative contributions of metabolite classes to the total number of metabolites associated with healthy (green) or unhealthy (orange) dietary patterns are shown. The y-axis shows all identified metabolite classes. The x-axis shows the relative contribution of the metabolite classes to the total number of metabolites, where 0% is found in the middle of the x-axis and relative contributions of metabolite classes to healthy dietary patterns are on the right side and those of unhealthy dietary patterns are on the left side.

Several metabolites are significant in multiple reports

Several metabolites were significantly associated with either a healthy or unhealthy dietary pattern in multiple reports, as shown in **table 2**. To account for potential bias, metabolites reported by Trimigno *et al.*[35], Khakimov *et al.*[36], Andersen *et al.*[37], and Acar *et al.*[38], that study the same population, were counted once when found in multiple of these reports. Interestingly, while trimethylamine N-oxide (TMAO) was found to be associated with healthy dietary patterns in 5 studies, 3 of which belonged to the studies mentioned before, TMAO was also reported to be associated with unhealthy dietary pattern in one study. Similarly, carnitine was associated with unhealthy dietary pattern in 5 studies and to healthy dietary pattern in one study.

Table 2 – Top metabolites associated with dietary patterns. Freq indicates number of reports in which the metabolite was reported to be significantly associated with the dietary pattern.

Dietary pattern	Metabolite	Freq
Healthy	3-Hydroxybutyric acid	4
Healthy	Phenylalanine	4
Healthy	Lenticin	4
Healthy	Lysophosphatidylcholine (22:6)	3
Healthy	Phosphatidylcholine (18:0/22:6)	3
Healthy	Trimethylamine N-oxide (TMAO)	3
Unhealthy	Carnitine	5
Unhealthy	4-Hydroxyproline	4
Unhealthy	3,7-Dimethyluric acid	3
Unhealthy	7-Methyluric acid	3
Unhealthy	7-Methylxanthine	3
Unhealthy	Theobromine	3

Different metabolomic pathways relating to dietary pattern

To uncover potentially differing biological pathways between healthy and unhealthy dietary patterns, pathway over-representation analyses of metabolites associating to these dietary

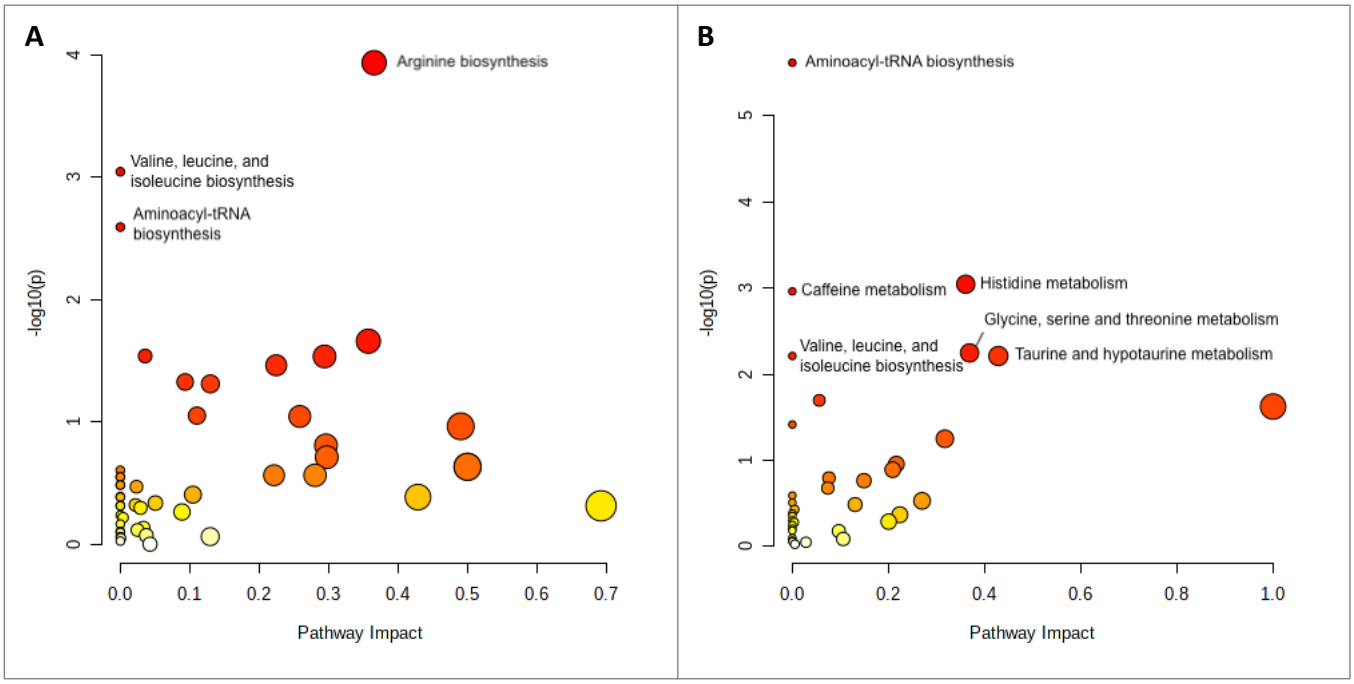


Figure 4 – Visual summary of metabolite pathway over-representation analysis of (A) metabolites associated with healthy dietary patterns and (B) metabolites associated with unhealthy dietary patterns. -log10 p-values are shown on the y-axis and are represented by node color, where a darker red color indicates a higher significance. Pathway impact, as calculated from pathway topology analysis, is shown on the x-axis, and is represented by node radius.

patterns were performed. Human Metabolome Database identifiers of all unique metabolite compounds corresponding to a dietary pattern were entered into the MetaboAnalyst 5.0 database. For healthy dietary patterns, 189 unique metabolites were entered, of which 144 were annotated well enough in the MetaboAnalyst database to be included in the analysis. 3 pathways were significantly enriched below an FDR adjusted p-value of 0.1: arginine biosynthesis, valine, leucine, and isoleucine biosynthesis, and aminoacyl-tRNA biosynthesis, as shown in **figure 4A** (and **supplementary table 1**). Here, pathway enrichment significance (-log10 scale) is plotted against pathway impact, as calculated from pathway topology analysis. A higher pathway impact indicates a higher importance of the metabolites entered in the analysis in that pathway. For unhealthy dietary patterns, 171 unique metabolites were entered into the MetaboAnalyst database, of which 148 were annotated well enough to be included in the analysis. **Figure 4B** (and **supplementary table 1**) shows 6 pathways that were significant below an FDR adjusted p-value of 0.1: aminoacyl-tRNA biosynthesis, histidine metabolism, caffeine metabolism, glycine, serine, and threonine metabolism, valine, leucine and isoleucine biosynthesis, and taurine and hypotaurine metabolism. Interestingly, threobromine, 3,7-dimethyluric acid, 7-methylxanthine, and 7-methyluric acid, which are all reported to be associated with unhealthy dietary patterns in 3 separate articles, as can be seen in **table 2**, are all part of the caffeine metabolism pathway. Furthermore, L-phenylalanine, which was associated with healthy dietary patterns in 4 separate articles, is part of the aminoacyl-tRNA biosynthesis pathway.

Discussion

This review has attempted to provide an overview of literature describing metabolite patterns associated with healthy compared to relatively unhealthy dietary patterns. 20 articles were identified in a systematic search for articles that used metabolomics to uncover metabolites associated with healthy and unhealthy dietary patterns. Differences in metabolite class compositions were found, as glycerophospholipids, imidazopyrimidines, and fatty acid esters show a higher contribution to unhealthy dietary patterns, while carbohydrates, benzoic acids, and cinnamic acids show a higher contribution to healthy dietary patterns (**figure 2**). Metabolite compositions were also found to be dependent on sample specimen, as metabolite class distributions differ between blood and urine samples (**figure 3**). Additionally, several metabolites were reported in multiple articles and pathway over-representation analyses of metabolites associated with either healthy or unhealthy dietary patterns returned several significantly enriched pathways (**figure 4**).

Current results are influenced by several factors. Firstly, only metabolites that were reported as significant were included in the analyses, excluding all metabolites that were not significant below these arbitrary thresholds. Of all significant metabolites, the strength of the association to dietary patterns was not considered in the current analysis. Therefore, to uncover stronger associations, metabolites that were mentioned in multiple articles were highlighted. Secondly, not all metabolites could be matched to a Human Metabolome Database entry, resulting in loss of data. Lastly, not all metabolites were characterized well enough in MetaboAnalyst to be included in pathway analysis. Therefore, more efforts should be made to identify and characterize metabolites to expand current metabolite databases to advance functional interpretation of metabolomics data.

Due to heterogeneity of the articles, such as different dietary patterns, sample specimens, metabolomics methods, and outcome values, interpreting the combined results of these articles should be done with caution. However, some interesting findings are

discussed here. Firstly, the benzoic and cinnamic acid classes were found to have a higher contribution to healthy dietary patterns. These classes belong to the phytochemical compounds[41], reflecting fruit and vegetable intake in the healthy dietary pattern. Cinnamic acids have been shown to decrease inflammation in a periodontitis animal model[42] and have been associated with a beneficial influence on type 2 diabetes[41]. Moreover, lenticin was found to be associated with healthy dietary patterns in four separate articles and has been proposed as a biomarker of legume consumption[33]. Lenticin has been shown to inhibit adipocyte differentiation and improve insulin sensitivity *in vitro*, thereby having the potential to reduce the risk of diabetes mellitus in obese individuals and treat insulin resistance[43]. Therefore, the higher contribution of cinnamic acids and lenticin to healthy dietary patterns may result in beneficial effects on health status. Secondly, TMAO was mostly found to be associated with healthy dietary patterns and is likely to be a result of fish intake in these dietary patterns[16,38]. Association of this metabolite with healthy dietary patterns was not expected, as high TMAO level is shown to be associated with major adverse cardiovascular events and all-cause mortality[44]. TMAO is shown to be synthesized by the gut microbiota from ingested phosphatidylcholines[45,46], 2 of which were also found to be associated with healthy dietary patterns in 3 separate studies. Phosphatidylcholines are part of the glycerophospholipids group, which was found in greater abundance in unhealthy dietary patterns (**figure 2**). Therefore, unhealthy dietary patterns may provide more phosphatidylcholine substrate for gut microbiome conversion into TMAO compared to healthy dietary patterns. Thirdly, carnitine, which is found in high concentrations in red meat[47], was found to be associated with unhealthy dietary patterns in 5 separate studies. Carnitine plays a key role in lipid metabolism, as it facilitates transport of long-chain fatty acids into the mitochondrial matrix[47,48]. Findings regarding health outcomes are contrasting. Carnitine has been described to exert cardioprotective effects, reducing oxidative stress, inflammation, and necrosis of cardiomyocytes[48]. However, it was also shown to be converted to TMAO by the gut microbiota[49], thereby increasing cardiovascular risk. To summarize, while findings on cinnamic acids and lenticin reinforce associations to healthy dietary patterns and potential beneficial effects of healthy diet on health status, findings on TMAO, phosphatidylcholines, and carnitine regarding association to dietary pattern and health outcome are contrasting and require further research.

4 out of 6 top metabolites associated with unhealthy dietary patterns are involved in the caffeine metabolism pathway. Furthermore, the caffeine metabolism pathway was found to be significantly enriched in unhealthy dietary patterns in the pathway over-representation analysis. Caffeine has many effects in different tissues, including central nervous system, skeletal muscle, cardiovascular, renal, and pulmonary tissue[50]. Interestingly, while caffeine has mostly been described to have antioxidant capacities, high doses of caffeine have been shown to increase proinflammatory cytokine levels[50]. Coffee consumption has been associated with a lower risk of cardiovascular disease incidence and mortality[51,52] and coffee and tea consumption has been associated with a lower risk of coronary heart disease mortality[53]. However, one of these studies found similar associations to cardiovascular disease risk between caffeinated and decaffeinated coffee, making caffeine an unlikely cause of the observed associations[51]. Another study showed protective effects of coffee on cardiovascular disease, while caffeine intake was associated with increased risk of cardiovascular disease outcomes[54]. Therefore, further research into the effects of caffeine and related metabolites on health outcomes is needed.

In this review, differences in metabolite compositions were found between plasma and urine samples (**figure 3**). Interestingly, lipid-soluble metabolites are only present in plasma and cannot be measured in urine[55]. This underscores the importance of separating sample specimen results in future investigations. Furthermore, future studies should ideally make use of comparable metabolite measuring methods and study designs. This will enable functional meta-analyses of mass spectrometry peaks, in which metabolomics results of different studies can be integrated[21], taking into account the strength of metabolite associations to dietary patterns, thus providing a more accurate metabolite profile of dietary patterns.

Conclusion

This review aimed to conduct a systematic search to provide an overview of relevant literature and to gather and analyze available data on the metabolomic signatures associated with healthy and unhealthy dietary patterns. 20 relevant randomized controlled trials were identified. Combined results show differences in metabolomic profiles of healthy and unhealthy dietary patterns. However, identified differences do not provide a unanimous and clear consensus regarding connections to health status. More research on metabolites is needed to expand metabolite databases to facilitate metabolite identification and functional interpretation of metabolomics data. Additionally, more comparable study parameters will enable meta-analyses of mass spectrometry peaks to more accurately identify metabolite profiles associated with healthy and unhealthy dietary patterns.

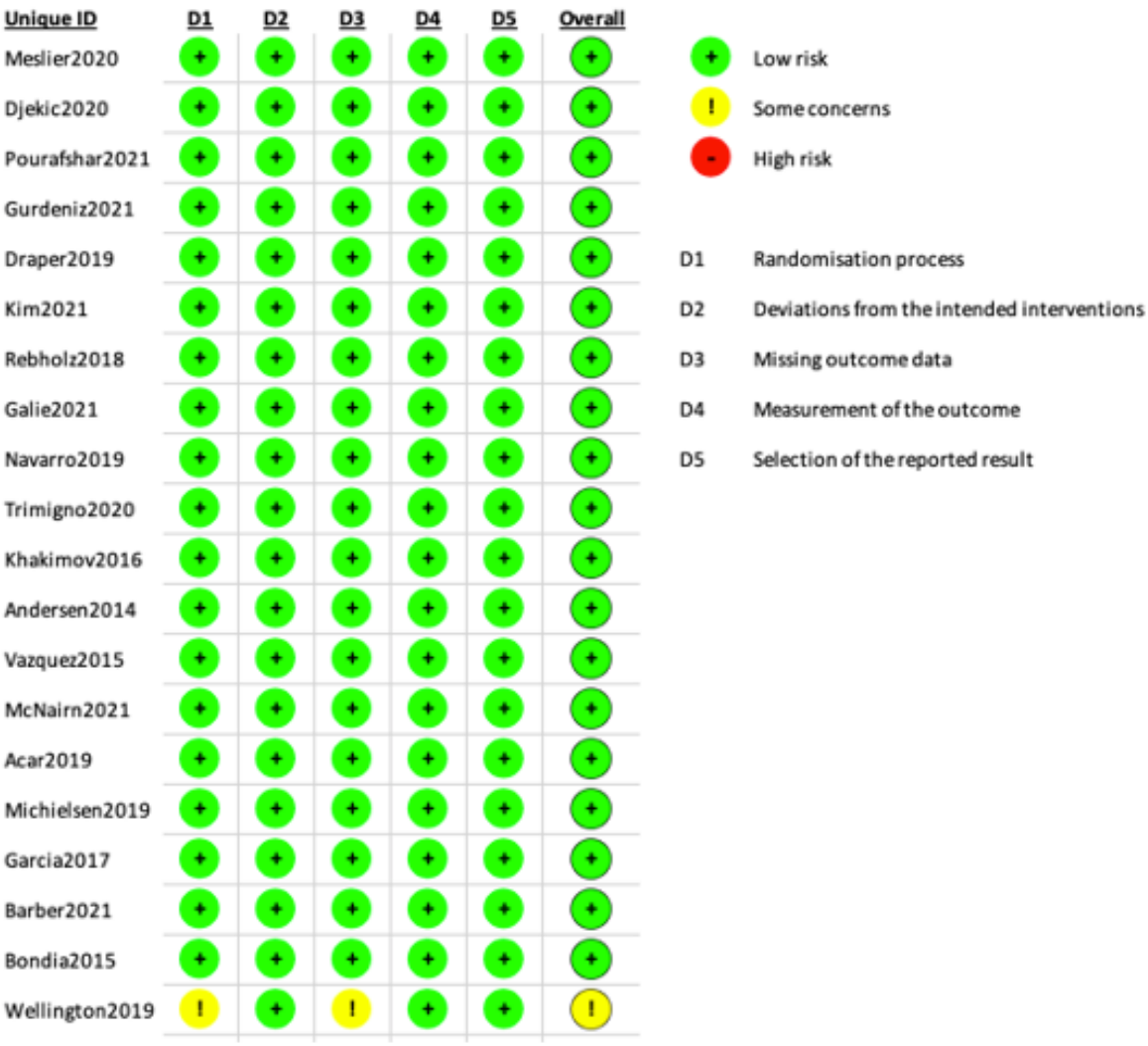
References

1. Bertram, H. C. & Jakobsen, L. M. A. Nutrimetabolomics: integrating metabolomics in nutrition to disentangle intake of animal-based foods. *Metabolomics* **14**, 1–10 (2018).
2. Jannasch, F., Kröger, J. & Schulze, M. B. Dietary patterns and Type 2 diabetes: A systematic literature review and meta- analysis of prospective studies. *J. Nutr.* **147**, 1174–1182 (2017).
3. Tindall, A. M. *et al.* Replacing Saturated Fat With Walnuts or Vegetable Oils Improves Central Blood Pressure and Serum Lipids in Adults at Risk for Cardiovascular Disease: A Randomized Controlled-Feeding Trial. *J. Am. Heart Assoc.* **8**, 9–12 (2019).
4. Estruch, R. *et al.* Primary Prevention of Cardiovascular Disease with a Mediterranean Diet Supplemented with Extra-Virgin Olive Oil or Nuts. *N. Engl. J. Med.* **378**, e34 (2018).
5. Dehghan, M. *et al.* Relationship between healthy diet and risk of cardiovascular disease among patients on drug therapies for secondary prevention a prospective cohort study of 31 546 high-risk individuals from 40 countries. *Circulation* **126**, 2705–2712 (2012).
6. Dandamudi, A., Tommie, J., Nommsen-Rivers, L. & Couch, S. Dietary patterns and breast cancer risk: A systematic review. *Anticancer Res.* **38**, 3209–3222 (2018).
7. Zheng, J. *et al.* Dietary patterns and risk of pancreatic cancer: A systematic review. *Nutr. Rev.* **75**, 883–908 (2017).
8. Grosso, G. *et al.* Possible role of diet in cancer: Systematic review and multiple meta-analyses of dietary patterns, lifestyle factors, and cancer risk. *Nutr. Rev.* **75**, 405–419 (2017).
9. Bach, K. E. *et al.* Healthy dietary patterns and incidence of CKD: A meta-analysis of cohort studies. *Clin. J. Am. Soc. Nephrol.* **14**, 1441–1449 (2019).
10. Ulaszewska, M. M. *et al.* Nutrimetabolomics: An Integrative Action for Metabolomic Analyses in Human Nutritional Studies. *Mol. Nutr. Food Res.* **63**, 1–38 (2019).
11. Li, J. *et al.* The Mediterranean diet, plasma metabolome, and cardiovascular disease risk. *Eur. Heart J.* **41**, 2645–2656 (2020).
12. Rinschen, M. M., Ivanisevic, J., Giera, M. & Siuzdak, G. Identification of bioactive metabolites using activity metabolomics. *Nat. Rev. Mol. Cell Biol.* **20**, 353–367 (2019).
13. Esko, T. *et al.* Metabolomic profiles as reliable biomarkers of dietary composition. *Am. J. Clin. Nutr.* **105**, 547–554 (2017).
14. Playdon, M. C. *et al.* Identifying biomarkers of dietary patterns by using metabolomics. *Am. J. Clin. Nutr.* **105**, 450–465 (2017).
15. O'Sullivan, A., Gibney, M. J. & Brennan, L. Dietary intake patterns are reflected in metabolomic profiles: Potential role in dietary assessment studies. *Am. J. Clin. Nutr.* **93**, 314–321 (2011).
16. Garcia-Perez, I. *et al.* Objective assessment of dietary patterns by use of metabolic phenotyping: a randomised, controlled, crossover trial. *Lancet Diabetes Endocrinol.* **5**, 184–195 (2017).
17. Ouzzani, M., Hammady, H., Fedorowicz, Z. & Elmagarmid, A. Rayyan-a web and mobile app for systematic reviews. *Syst. Rev.* **5**, 1–10 (2016).
18. Sterne, J. A. C. *et al.* RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ* **366**, 1–8 (2019).
19. R Core Team. R: A Language and Environment for Statistical Computing. (2020).
20. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag New York, 2016).
21. Pang, Z. *et al.* MetaboAnalyst 5.0: Narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res.* **49**, W388–W396 (2021).
22. Meslier, V. *et al.* Mediterranean diet intervention in overweight and obese subjects lowers plasma cholesterol and causes changes in the gut microbiome and metabolome independently of energy intake. *Gut* **69**, 1258–1268 (2020).
23. Galié, S. *et al.* Effects of Mediterranean Diet on plasma metabolites and their relationship with insulin resistance and gut microbiota composition in a crossover randomized clinical trial. *Clin. Nutr.* **40**, 3798–3806 (2021).
24. Vázquez-Fresno, R. *et al.* Metabolomic pattern analysis after mediterranean diet intervention in a nondiabetic population: a 1- and 3-year follow-up in the PREDIMED study. *J. Proteome Res.* **14**, 531–540 (2015).

25. Barber, C. *et al.* Differential Effects of Western and Mediterranean-Type Diets on Gut Microbiota: A Metagenomics and Metabolomics Approach. *Nutrients* **13**, (2021).
26. Bondia-Pons, I. *et al.* Effects of short- and long-term Mediterranean-based dietary treatment on plasma LC-QTOF/MS metabolic profiling of subjects with metabolic syndrome features: The Metabolic Syndrome Reduction in Navarra (RESMENA) randomized controlled trial. *Mol. Nutr. Food Res.* **59**, 711–728 (2015).
27. Michielsen, C. C. J. R., Hangelbroek, R. W. J., Feskens, E. J. M. & Afman, L. A. Disentangling the Effects of Monounsaturated Fatty Acids from Other Components of a Mediterranean Diet on Serum Metabolite Profiles: A Randomized Fully Controlled Dietary Intervention in Healthy Subjects at Risk of the Metabolic Syndrome. *Mol. Nutr. Food Res.* **63**, e1801095 (2019).
28. Djekic, D. *et al.* Effects of a Vegetarian Diet on Cardiometabolic Risk Factors, Gut Microbiota, and Plasma Metabolome in Subjects With Ischemic Heart Disease: A Randomized, Crossover Study. *J. Am. Heart Assoc.* **9**, e016518 (2020).
29. Pourafshar, S. *et al.* Urine and Plasma Metabolome of Healthy Adults Consuming the DASH (Dietary Approaches to Stop Hypertension) Diet: A Randomized Pilot Feeding Study. *Nutrients* **13**, (2021).
30. Gürdeniz, G. *et al.* Analysis of the SYSDIET Healthy Nordic Diet randomized trial based on metabolic profiling reveal beneficial effects on glucose metabolism and blood lipids. *Clin. Nutr.* **41**, 441–451 (2021).
31. Draper, C. F. *et al.* Vegan and Animal Meal Composition and Timing Influence Glucose and Lipid Related Postprandial Metabolic Profiles. *Mol. Nutr. Food Res.* **63**, e1800568 (2019).
32. Kim, H. *et al.* Urine Metabolites Associated with the Dietary Approaches to Stop Hypertension (DASH) Diet: Results from the DASH-Sodium Trial. *Mol. Nutr. Food Res.* **65**, e2000695 (2021).
33. Rebholz, C. M., Lichtenstein, A. H., Zheng, Z., Appel, L. J. & Coresh, J. Serum untargeted metabolomic profile of the Dietary Approaches to Stop Hypertension (DASH) dietary pattern. *Am. J. Clin. Nutr.* **108**, 243–255 (2018).
34. Navarro, S. L. *et al.* Plasma metabolomics profiles suggest beneficial effects of a low-glycemic load dietary pattern on inflammation and energy metabolism. *Am. J. Clin. Nutr.* **110**, 984–992 (2019).
35. Trimigno, A. *et al.* Human urine 1H NMR metabolomics reveals alterations of protein and carbohydrate metabolism when comparing habitual Average Danish diet vs. healthy New Nordic diet. *Nutrition* **79**, 110867 (2020).
36. Khakimov, B. *et al.* New Nordic Diet versus Average Danish Diet: A Randomized Controlled Trial Revealed Healthy Long-Term Effects of the New Nordic Diet by GC-MS Blood Plasma Metabolomics. *J. Proteome Res.* **15**, 1939–1954 (2016).
37. Andersen, M.-B. S. *et al.* Untargeted metabolomics as a screening tool for estimating compliance to a dietary pattern. *J. Proteome Res.* **13**, 1405–1418 (2014).
38. Acar, E. *et al.* Biomarkers of Individual Foods, and Separation of Diets Using Untargeted LC-MS-based Plasma Metabolomics in a Randomized Controlled Trial. *Mol. Nutr. Food Res.* **63**, e1800215 (2019).
39. McNairn, M. *et al.* Postprandial Dried Blood Spot-Based Nutritional Metabolomic Analysis Discriminates a High-Fat, High-Protein Meat-Based Diet from a High Carbohydrate Vegan Diet: A Randomized Controlled Crossover Trial. *J. Acad. Nutr. Diet.* **121**, 931-941.e2 (2021).
40. Wellington, N. *et al.* Metabolic Trajectories Following Contrasting Prudent and Western Diets from Food Provisions: Identifying Robust Biomarkers of Short-Term Changes in Habitual Diet. *Nutrients* **11**, (2019).
41. Adisakwattana, S. Cinnamic acid and its derivatives: Mechanisms for prevention and management of diabetes and its complications. *Nutrients* **9**, (2017).
42. Karatas, O. *et al.* Cinnamic acid decreases periodontal inflammation and alveolar bone loss in experimental periodontitis. *J. Periodontal Res.* **55**, 676–685 (2020).
43. Luan, G. *et al.* Hypaphorine, an Indole Alkaloid Isolated from Caragana korshinskii Kom., Inhibites 3T3-L1 Adipocyte Differentiation and Improves Insulin Sensitivity in Vitro. *Chem.*

Biodivers. **14**, (2017).
44. Guasti, L. *et al.* TMAO as a biomarker of cardiovascular events: a systematic review and meta-analysis. *Intern. Emerg. Med.* **16**, 201–207 (2021).
45. Wang, Z. *et al.* Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **472**, 57–65 (2011).
46. Tang, W. H. W. *et al.* Intestinal Microbial Metabolism of Phosphatidylcholine and Cardiovascular Risk. *N. Engl. J. Med.* **368**, 1575–1584 (2013).
47. Dragsted, L. O. Biomarkers of meat intake and the application of nutrigenomics. *Meat Sci.* **84**, 301–307 (2010).
48. Wang, Z. Y., Liu, Y. Y., Liu, G. H., Lu, H. Bin & Mao, C. Y. L-Carnitine and heart disease. *Life Sci.* **194**, 88–97 (2018).
49. Koeth, R. A. *et al.* L-Carnitine in omnivorous diets induces an atherogenic gut microbial pathway in humans. *J. Clin. Invest.* **129**, 373–387 (2019).
50. Barcelos, R. P., Lima, F. D., Carvalho, N. R., Bresciani, G. & Royes, L. F. Caffeine effects on systemic metabolism, oxidative-inflammatory pathways, and exercise performance. *Nutr. Res.* **80**, 1–17 (2020).
51. Van Dongen, L. H., Mölenberg, F. J. M., Soedamah-Muthu, S. S., Kromhout, D. & Geleijnse, J. M. Coffee consumption after myocardial infarction and risk of cardiovascular mortality: A prospective analysis in the Alpha Omega Cohort. *Am. J. Clin. Nutr.* **106**, 1113–1120 (2017).
52. Ding, M., Bhupathiraju, S. N., Satija, A., Van Dam, R. M. & Hu, F. B. Long-term coffee consumption and risk of cardiovascular disease: A systematic review and a dose-response meta-analysis of prospective cohort studies. *Circulation* **129**, 643–659 (2014).
53. De Koning Gans, J. M. *et al.* Tea and coffee consumption and cardiovascular morbidity and mortality. *Arterioscler. Thromb. Vasc. Biol.* **30**, 1665–1671 (2010).
54. Gaeini, Z., Bahadoran, Z., Mirmiran, P. & Azizi, F. Tea, coffee, caffeine intake and the risk of cardio-metabolic outcomes: Findings from a population with low coffee and high tea consumption. *Nutr. Metab.* **16**, 1–10 (2019).
55. Gibney, M. J. *et al.* Metabolomics in human nutrition: Opportunities and challenges. *Am. J. Clin. Nutr.* **82**, 497–503 (2005).

Supplementary



Supplementary figure 1 – **Risk of bias determined based on 5 domains (D1-D5).** Determined using the revised Risk of Bias tool by Cochrane. Unique ID corresponds to first author and publication year.

Supplementary table 1 – Pathway over-representation results of metabolites associated with (A) healthy and (B) unhealthy dietary patterns. Pathways significant below an FDR adjusted p-value < 0.01 are shown. Total is the total number of metabolites in the pathway. Expected is the number of metabolites that can be expected to be found by chance in the entered metabolite list. Hits represents the number of metabolites present in the entered metabolite list. Raw p is the original p-value calculated from the over-representation analysis. Holm adjust is the p-value adjusted by the Holm-Bonferroni method and FDR is the p-value adjusted using False Discovery Rate. The pathway impact value is calculated by pathway topology analysis.

A	Total	Expected	Hits	Raw p	-log10(p)	Holm adjust	FDR	Impact
Arginine biosynthesis	14	0.89419	6	0.00011555	3.9372	0.0097059	0.0097059	0.36548
Valine, leucine, and isoleucine biosynthesis	8	0.51097	4	0.00089952	3.046	0.07466	0.03778	0
Aminoacyl-tRNA biosynthesis	48	3.0658	9	0.0025478	2.5938	0.20892	0.071339	0

B	Total	Expected	Hits	Raw p	-log10(p)	Holm adjust	FDR	Impact
Aminoacyl-tRNA biosynthesis	48	2.4774	12	2,413E-03	5.6175	0.00020269	0.00020269	0
Histidine metabolism	16	0.82581	5	0.00090151	3.045	0.074826	0.030536	0.36065
Caffeine metabolism	10	0.51613	4	0.0010906	2.9624	0.089426	0.030536	0
Glycine, serine, and threonine metabolism	33	1.7032	6	0.0056567	2.2474	0.45819	0.086043	0.36871
Valine, leucine, and isoleucine biosynthesis	8	0.4129	3	0.0061459	2.2114	0.49167	0.086043	0
Taurine and hypotaurine metabolism	8	0.4129	3	0.0061459	2.2114	0.49167	0.086043	0.42857