The American Journal of Clinical Nutrition 123 (2026) 101100

The American Journal of CLINICAL NUTRITION

journal homepage: https://ajcn.nutrition.org/

Original Research Article

Comparing a diet-wide panel of biomarkers of food intake in whole blood and 24-hour urine and self-reported with known dietary intakes: randomized feeding trial of three 48-hour interventions

Aysu Shahin [1,2], Alastair B Ross [3], Robin M Turner [4], Jim Mann [1,2], Andrew N Reynolds [1,2,*]

[1] *Department of Medicine, University of Otago, Dunedin, New Zealand;* [2] *Riddet Institute Centre of Research Excellence, Massey University, Palmerston North, New Zealand;* [3] *Proteins and Metabolites, Smart Foods and Bioproducts, AgResearch, Lincoln, New Zealand;* [4] *Biostatistics Centre, Department of Preventive and Social Medicine, University of Otago, Dunedin, New Zealand*

A B S T R A C T

Background: Specific metabolites detected in biofluids after food intake have been proposed as objective markers to address limitations with traditional dietary assessment methods. Although hundreds of metabolites have been associated with foods or food groups, they require validation.

Objectives: The aim of this study was to develop a panel from the proposed biomarkers of dietary intake reflecting major food groups and compare them against known and self-reported dietary intakes.

Methods: A randomized crossover trial of 3 interventions, including a day of foods consumed under observation, was performed. Each feeding day was of 3 food groups (e.g., whole grains, dairy, and fish) with optional snacks (e.g., fruit, chicken, and legumes) but absent of 3 other food groups (e.g., meat, vegetables, nuts, and seeds). Fasted whole blood and 24-h urine samples were analyzed by liquid chromatography mass spectrometry to detect previously proposed biomarkers of food intake. Urinary sodium was measured. Pairwise correlation coefficients and generalized linear modeling (GLM) considered relationships between biomarkers and food groups. Comparisons were drawn between self-reported and known dietary intakes.

Results: Twenty-one participants [mean age 24.8, standard deviation (SD) 6.0 y, body mass index: 24.1; SD: 4.0] completed the trial. GLM coefficients indicated fish intake was associated with 3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid [62.15 (95% confidence interval: 35.00, 89.29)], wholegrain intake with 3,5-dihydroxybenzoic acid [87.32 (24.28, 150.35)] and fruit intake with fructose [5.39 (2.53, 8.25)], and *s*-methylcysteine [5.91 (1.24, 10.58)]. GLM rate ratios indicated chicken intake was associated with 3-methylhistidine [0.19 (0.07, 0.31)], anserine [0.21 (0.05, 0.37)], and carnosine [0.11 (0.03, 0.19)], legume intake with glycine betaine [0.21 (0.02, 0.40)] and vegetable intake with sulforaphane [0.30 (0.20, 0.47)], S-methylcysteine [0.23 (0.14, 0.45)], methoxytyramine [0.21 (0.08, 0.35)], and β-carotene [0.05 (0.02, 0.08)]. There was no association between 24-h urinary sodium and known sodium intake [0.11 (−0.06, 0.28)]. Self-reported dietary intake was associated above acceptable level ($r > 0.40$) with known intake.

Concluzions[We identified some previously reported associations between foods and proposed biomarkers, but not all, outlining the need for assessing dietary biomarkers across a range of study designs including food intakes within realistic ranges.

This trial was registered at ACTRN as 12622001036707 (https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=384292&isReview=true).

Keywords: dietary assessment, objective markers, feeding study, biomarkers of food intake, LCMS

Introduction

Biomarkers of food intake (BFIs) have been proposed as alternatives or complements to traditional dietary assessment methods [1–3]. Advances in metabolomics, i.e., the science of studying the metabolome, have resulted in the detection of a greater number of metabolites in biofluids at relatively low cost and are becoming important techniques for the identification of objective measures of dietary intake [4–8]. Most BFIs are single metabolites or a combination of metabolites found in the body after the intake of the related food [8,9]. There are several classes of BFIs, namely recovery, concentration, replacement, and predictive biomarkers, although some biomarkers can fall into more than one of these categories [10]. Ideal BFIs should be based on exogenous metabolites (i.e., a component of the food) and

Abbreviations: BFI, biomarkers of food intake; CI, confidence interval; C18, octadecyl; GLM, generalized linear modeling; HILIC, hydrophilic interaction liquid chromatography; MeOH, methanol; SPE, solid-phase extraction.

* Corresponding author.
E-mail address: andrew.reynolds@otago.ac.nz (A.N. Reynolds).

https://doi.org/10.1016/j.ajcnut.2025.10.017
Received 11 March 2025; Received in revised form 27 August 2025; Accepted 22 October 2025; Available online 9 December 2025

specific to individual foods or food groups. There are many factors that must be considered before confirming candidate metabolites as objective biomarkers of dietary intake. The range and diversity of metabolites found in food, as well as how they may be processed in the human body, are not yet well understood.

Over the past 2 decades, research on metabolite markers has increasingly focused on the following 3 areas: *1*) the discovery of BFIs, *2*) the identification of metabolic pathways altered because of dietary interventions, and *3*) the investigation of the association between BFIs, nutrition, and health status [6,11,12]. Substantial progress has been made in discovering potential BFIs with databases now cataloging hundreds of identified candidate BFIs [13–15]. However, evaluating their association with diet has been relatively limited, as few have been considered within controlled feeding studies. Many dietary biomarker "discovery" studies have used very high intakes to find metabolites that clearly respond to a particular food [16], and it can be unclear how well these markers reflect intake within the range of usual intake. As part of the FoodBall initiative, Dragsted et al. [16] proposed 8 criteria required for the validation of candidate BFIs; however, few studies that meet these validation criteria have been conducted to date. To understand whether proposed objective markers of dietary intake are useful, further work is required to test methods in the context of usual dietary intake and different types of dietary interventions. To further gain knowledge in this field toward BFI validation, we conducted a randomized crossover feeding trial providing foods of known amounts from all major food groups [17]. We carried out a targeted assessment of proposed metabolites in blood and urine to develop a panel reflecting overall diet. We also compared objective (biomarkers), self-reported (24-h recall), and known (what food was consumed under observation) intakes to comment on these assessment methods in this study setting.

Methods

This trial was prospectively registered (ACTRN 12622001036707) and conducted in Aotearoa New Zealand from September 2022 to September 2023 with all feeding interventions delivered over the first 5 wk. Ethical consent was obtained from the University of Otago Human Ethics Committee (UOHEC H22/075) after scientific peer review, and all participants provided informed consent before participation.

Study design

This trial used a randomized crossover design with 3 48-h feeding interventions, each separated by ≥4-d washout. Day 1 of each intervention was an unmonitored day of ad libitum dietary intake, physical activity, and sleep behaviors with 24-h urine collection. Participants were asked on these days to standardize these behaviors, including any medication or supplement use throughout the study period, so that each day of the study was similar. Day 2 of each intervention was a feeding day with all meals provided in known amounts and consumed under observation with a second 24-h urine collection. A fasted blood sample was drawn, and 24-h recall was taken in the morning after each unmonitored standardized day and each feeding day of foods consumed under observation.

Eligibility criteria

Participants aged 18–40 y were sought through online and paper-based recruitment materials. These young adults were chosen to minimizethe occurrence of unknown diseases or metabolic conditions thatmight impede metabolite processing (e.g., renal disease). Participants were included if they were willing to comply with the study requirementsof each intervention and the necessary urine collection andblood draw. Individuals were excluded if they had food allergies, foodsensitivities, or food intolerance, a history of eating disorders and pre-existinghealth conditions that impact digestion and metabolism. Enrollment and consenting were done via Research Electronic Data Capture(REDCap).

Intervention

Each of the 3 feeding days consisted of 3 main meals of known amountsconsumed under observation, with preweighed snacks provided tobe consumed ad libitum outside of mealtimes. All meals were preparedand provided in the same facility, where participants were required toattend breakfast, lunch, and dinner. Menus were designed to provide a varietyof foods at a normal intake level from a selection of 9 major food groups according to the New Zealand Food Guidelines [17]. Food groupswere categorized into daily menus according to 3 key criteria: *1*) ensuringa balanced daily intake with ≥1 animal protein source/d; *2*) designing a palatable, satiating, full-day menu; and *3*) minimizing overlapamong food groups with shared BFIs (e.g., flavonoids in fruits andvegetables). Each feeding day was of meals comprising select 3 food groups(i.e., red meat, vegetables, nuts, and seeds), with optional snacks of3 other food groups (i.e., poultry, legumes, and fruit), and a complete absenceof 3 other food groups (i.e., fish, dairy, and whole grains). Full interventionmenus are shown in Table 1.

Participantswere instructed to eat no other foods on feeding days otherthan the food we provided. Each meal portion was weighed out andpredefined; however, participants could choose to eat any number of portions of any meal of their choice, provided they constipated the full portion each time. The same brand and batch of foods purchased from the 1 supplier were used over the 5 wk where feeding interventions were delivered, to standardize as much as possible potential variability in the biomarker or sodium contents of foods. For seasoning, participants were provided with preweighed packets of salt that, if used, had to be consumed in full to obtain an accurate reflection of dietary sodium intake. Monitoring of all mealtimes, salt packets, empty plates, and empty snack bags was undertaken to ensure intervention adherence.

Randomization and blinding

Participantswere randomly assigned to intervention order using a balanced computer-generated protocol with no stratification. The randomization list was created and uploaded into REDCap by a researchernot involved with recruitment, participant communication, datacollection, or data analysis. The randomization list could not be accessedby study staff enrolling participants. Because of the nature of dietary feeding studies, blinding of participants was not possible; however, participants were not informed of their allocated intervention orderat any time point.

Measurements

Baseline data collection included demographic data, as well as weight, height, and blood pressure. For self-reported dietary data, participantsprovided a 24-h recall the morning after each standard and feedingday using a multipass method [18] for a total of 6 24-h recalls during the study. Recalls were entered into FoodWorks 9 (Xyris SoftwareAustralia Pty Ltd) for nutrient and food group analysis using local food composition tables [17]. All 24-h urine samples were

TABLE 1
Full intervention menu.

Wholegrain, seafood, and dairy		Red meat, vegetables, nuts, and seeds		Chicken, egg, fruit, and legumes	
Food item	Amount	Food item	Amount	Food item	Amount
Breakfast		Breakfast		Breakfast	
Rolled oats	30 g	Canola oil	25 g	Apple	30 g
Milk, full fat	60 g	Onion, white	40 g	Pear	30 g
Honey	5 g	Potato	70 g	Kiwifruit	30 g
Greek yogurt, full fat	20 g	Garlic	5 g	Green grape	30 g
Cinnamon, ground	0.5 g	Tomato	60 g	Lemon juice	6 g
		Red capsicum	40 g	Egg, hard boiled	102 g
Lunch		Baby spinach	20 g	Chicken, smoked	50 g
Tuna, canned in oil	92 g	Tomato paste, NSA	25 g		
Wholegrain bread	88 g	Sausage, pork	110 g	Lunch	
Cheddar cheese	60 g	Salt	2.5 g	Red lentils, dry	110 g
Parsley, dried	0.25 g	Black pepper	0.15 g	Chicken breast	83 g
Butter, NSA	10 g	Turmeric	0.2 g	Soybean oil	25 g
Salt	0.2 g	Pumpkin seeds	6 g	Salt	0.3 g
Black pepper	0.2 g			Red kidney beans, canned	50 g
		Lunch		Mayonnaise	20 g
Dinner		Lettuce, iceberg	50 g	Avocado	45 g
Butter, NSA	15 g	Cauliflower	40 g	Black pepper	0.15 g
Sole fillet	100 g	Broccoli	40 g	Paprika	0.15 g
Salt	0.2 g	Kumara	50 g	Turmeric	0.15 g
Black pepper	0.1 g	Salt	4 g	Orange juice	200 mL
Brown rice, steamed	100 g	Beef	70 g		
Feta, full fat	25 g	Black pepper	0.15 g	Dinner	
Canola oil	5 g	Canola oil	19 g	Chicken bone broth	140 g
		Pumpkin seeds	6 g	Lentil pasta, boiled	100 g
Dessert		Lemon juice	30 g	Brown lentils, canned	60 g
Cream, full fat	40 g	Peanut butter	20 g	Black pepper	0.2 g
Sugar	18 g			Smoked paprika	0.1 g
Cream cheese, full fat	44 g	Dinner		Soybean oil	25 g
Vanilla extract	1 g	Lamb	200 g	Salt	0.15 g
Butter, NSA	20 g	Canola oil	16 g	Chicken breast	320 g
Rolled oats	25 g	Salt	2 g		
		Garlic powder	2.5 g	Dessert	
		Black pepper	0.1 g	Mixed berries, frozen	80 g
		Potato	80 g	Sugar, white	15 g
		Parsnip	115 g	Lemon juice	5 g
		Eggplant	50 g		
		Red capsicum	30 g		
		Onion, white	55 g		
		Mushroom	50 g		
		Garlic	5 g		
		Turmeric	0.1 g		
		Pumpkin seeds	6 g		
		Orange juice (optional)	200 mL		
		Dessert			
		Chocolate, dark	17 g		

Abbreviation: NSA, no salt added.
All meat and vegetable weights are given as raw weight unless specified.

collected following WHO protocols for population-level sodium determination [19]. All fasted venous blood samples were collected by a trained phlebotomist the morning of the feeding day and the morning after feeding day, and then frozen as whole blood at –80°C until analysis.

Sample preparation

We used the existing literature to select already identified metabolites reflecting dietary intake for screening [11–15,20,21] and included other compounds that are not specific dietary biomarkers but are related to food intake (e.g., amino acids and vitamins) as potential BFI. Because the selected potential dietary biomarkers varied widely in their polarity, 2 types of chromatography were used for our targeted analysis: hydrophilic interaction liquid chromatography (HILIC) for separating polar compounds and octadecyl (C18) for separating semipolar compounds. Analyses were designed to capture the full spectra of metabolites in each sample (untargeted), whereas the output from the instruments was processed to detect only the selected BFIs (targeted) for analysis, in line with our approach to consider known and proposed markers rather than identify new markers.

Urine was extracted following the protocol of Gonzalez-Dominguez et al. [7]. In brief, defrosted 1000 µL aliquots of urine samples were acidified using 20 µL of 4% phosphoric acid and 50 µL of internal standard mix [fumaric acid-d4, U13C-glucose, L-Dopa-d3, benzoic acid-d5, cholic acid-d5, choline-d9 (all from Cambridge Isotopes Ltd), and taxifolin (Sigma Aldrich) each at 1 µg/mL in 10% methanol (MeOH)]. The urine was mixed using a vortex mixer (10 s), then cleaned using solid-phase extraction (SPE) (Oasis HLB, Waters). The SPE plate

was conditioned with MeOH and water (with 0.1% formic acid) and 10 mM ammonium formate. Two fractions were collected for further analysis; polar compounds were eluted with 1.5 mL of water with 0.1% formic acid and 10 mM ammonium formate, and semipolar compounds were eluted with 1.5 mL MeOH with 0.1% formic acid and 10 mM ammonium formate. All samples were dried in a centrifuge evaporator (Christ RVC 2-18 CD plus, Martin Christ GmbH) and redissolved in 50% MeOH before analysis. Sample blanks were prepared in parallel to samples, with water added instead of urine, and quality control samples were prepared by pooling 10 µL aliquots from all samples.

We adjusted all urine samples for urine density (metabolite value divided by urine density for each sample). Urine density was measured using a hand-held refractometer (Milwaukee MA886 Digital Sodium Chloride Refractometer) calibrated with deionized water and dried with paper tissue according to the manufacturer's instructions. Once calibrated, 100 µL of sample was transferred onto the refractometer dish, and the specific gravity value was recorded in duplicate, with temperature also being noted. The mean of the 2 measures was used as the final density.

Whole blood samples were prepared for running with both C18 and HILIC chromatography to increase metabolic coverage [7]. Whole blood samples were prepared for running with both C18 chromatography and HILIC. Several solvents were tested to determine the most appropriate solvent for analysis: MeOH with 1% formic acid, acetonitrile with 1% formic acid, MeOH with 10 mM ammonium formate, acetone with 10 mM ammonium formate, and acetone with 1% formic acid. MeOH led to clotting with no supernatant. The remaining 4 extraction solvents were analyzed as per the methodology described in Supplemental Material 1, and acetone containing 1% formic acid was selected as the most appropriate solvent for whole blood based on the number of BFIs detected. All blood samples were thawed and initially extracted with acetone containing 1% formic acid (100 µL whole blood to 900 µL of extraction solvent) and dried using vacuum centrifugation. For analysis of polar compounds, acetone extract was resuspended in 50% acetonitrile with 0.1% formic acid. For analysis of semipolar compounds, the method of Dominguez-Fernandez et al. [22] was followed. Acetone extract was resuspended in 350 µL of water and cleaned using SPE, using the same method as for urine. The second extraction method (MeOH elution) was used. Whole blood samples were analyzed in electrospray ionization positive mode only. Expanded methodology is described in Supplemental Material 1.

Metabolite analysis

Samples were analyzed using liquid-chromatograph–high-resolution mass spectrometry (LC-MS, Shimadzu LCMS 9030 quadrupole-time-of-flight mass spectrometer). Samples were analyzed in randomized order based on individual, with samples from each individual analysis in a single block randomized based on sample collection time point. The targeted dietary biomarkers were analyzed before analysis to confirm that they were detectable, with mass as the primary adduct, and to determine retention time. Metabolites were subsequently identified based on retention time, high-resolution mass (<5 ppm), and isotope pattern matching score (>30%).

Because several metabolites were detected with both HILIC and C18 chromatography, the column that gave the best peak shape and intensity was selected for further analysis. Insight Explore (Shimadzu) was used to extract data. Compounds detected in <90% of samples were removed from the dataset. Overall variation from sample to sample was determined based on taxifolin and L-Dopa internal standards for urine and whole blood, respectively, meaning that the values obtained were semiquantified values, as full quantification is not possible due to the lack of accepted standards for the proposed metabolites we targeted. Expanded methodology is described in Supplemental Material 1.

Urine sodium assessment

Urine sodium was measured for each 24-h urine sample using the Ion-Selective Electrode module of the Roche/Hitachi Cobas C systems for the quantitative determination of sodium, potassium, and chloride in urine according to the manufacturer's protocols. Standardized quality control samples were used and spread out in 20-sample intervals between the samples. Any sample with an out-of-range value was repeated in triplicate. Urinary sodium measures from the 24-h urine collections were converted to dietary sodium equivalent using methods previously described [23]. Urinary sodium values were compared with the recorded sodium intake from provided foods using the same generalized linear modeling (GLM) steps as other BFIs described below.

Statistical analysis

As an exploratory study, insufficient data were available from previous studies to conduct a robust power calculation. We instead based our sample size on those used in previous studies [24–26] to set sample size at 21 completing participants, ceasing study recruitment when this number was reached. All available data of completed interventions were used, even if the participant had not completed all 3 interventions. Missing metabolite data were treated as random.

Data were tested for assumptions of normality and spread using normal distribution curves. All dietary data were presented per grams for each food group per feeding day and for the 3 standard days combined. Measures for energy, water, and sodium intake were also included. To consider metabolite concentrations in relation to food provided, exploratory multivariate analyses were carried out using principal component analyses and partial least square regression to create heat maps of Pearson correlation coefficients between dietary intake and potential BFIs in each biological sample on each intervention day.

The analysis to compare the relationship between BFIs and known dietary intake was completed in several steps using GLM, with the identity link function used for continuous outcomes and a log link for count outcomes (i.e., Poisson regression). Metabolite and self-reported dietary data from standardized days were included as covariates in the model to provide a baseline level for both diet and BFI amounts before intervention start. Findings were reported as GLM regression coefficients using data from main meals and exclusion days only (i.e., no snack days). A random effect was included to account for the clustering from repeated measures for participants. To investigate the linearity assumption multivariable fractional polynomial modeling for each BFI was used to determine the functional form of each BFI in the model (e.g., continuous, log, square, and cubic). BFI data (in ppm) were scaled to allow for comparison with dietary data [in grams (g)], using z scores for each data point, using the following formula: [variable − mean (variable)]/SD (variable)]. The BFI content in each biological sample was compared with known dietary intake without any covariates in the model, and a second analysis was conducted including previous-day BFI content as covariates for baseline BFI concentration adjustment. Findings were summarized as regression coefficients (identity link) or exponentiated rate ratios (Poisson model) with 95% confidence interval (CI) and P values. Sensitivity analyses were conducted to compare the effects of complete with incomplete 24-h urine samples on overall primary analysis results, with each previous step in primary analyses being repeated.

Analyses exploring the association between self-reported and known (true) dietary intake by feeding day were conducted by separating main meals intake (i.e., foods of known amount provided by the researchers and consumed under observation) and snack intake. Pairwise correlation coefficients were calculated using main meals only to capture known intake compared with self-reported intake for each food group variable depending on the distributions (Pearson's for normally distributed, Spearman's for nonnormally distributed). A correlation coefficient >0.40 was considered as acceptable based on previous literature in dietary validation studies [23,27]. Similar to the analysis of BFIs as predictors of dietary intake, GLM was conducted with self-reported intake as predictors of known intake and reported as regression coefficients. Further analysis comparing self-reported to known intake without clustering by feeding day was conducted. We did not adjust for multiple comparisons given that outcomes are correlated and doing so would increase the type II error rate [28] Data are presented as mean (SD) unless otherwise stated. Principal component analyses and partial least square were carried out using SIMCA 17 (Satorius), and all other analyses were carried out using StataSE version 17.0 [29].

Results

A total of 120 individuals expressed interest in the study, with 24 (74% female) providing consent and starting the study (Figure 1). Mean age was 24.8 (SD: 6.0) y and BMI was 24.1 (SD 4.0) kg/m^2. Participants self-identified as being New Zealand European (70%), Māori (22%), Chinese (4%), or Dutch (4%). Three (13%) participants were current smokers, 8 (33%) reported taking a dietary supplement in the last 3 mo and 5 (21%) were currently taking medications, although none with obvious metabolic effects. Full intervention menus are shown in Table 1. No participant reported issues in collecting their urine and blood samples, with 90% of 24-h urine collections being reported as all urine produced that day. Mean urine volume of the 131 urine samples was 1.5 (0.8) L. A total of 125 and 123 suggested BFIs and diet-related compounds were detected in whole blood and urine, respectively. Of these, 36 metabolites were detected in over 90% of samples and with a mean peak area >1000 and were included for further analysis.

Details of the dietary intake from intervention and standardized days are shown in Table 2. The GLM regression model coefficients between known dietary intake and BFIs in blood and urine for continuous variables and count variables, clustered by participant ID (Table 3) showed evidence of an association between ≥1 metabolite and its corresponding food group in 7 of 9 food groups (not for red meat and nuts and seeds), although the associations between alkylresorcinols in blood and whole grains, between isovalerylglycine in blood and dairy, and between 3,3′-diindolylmethane and capsaicin in urine and nonstarchy vegetables were negative. Adjusting for baseline levels of metabolites impacted the relationship between isovalerylglycine in urine and dairy and between methoxytyramine in

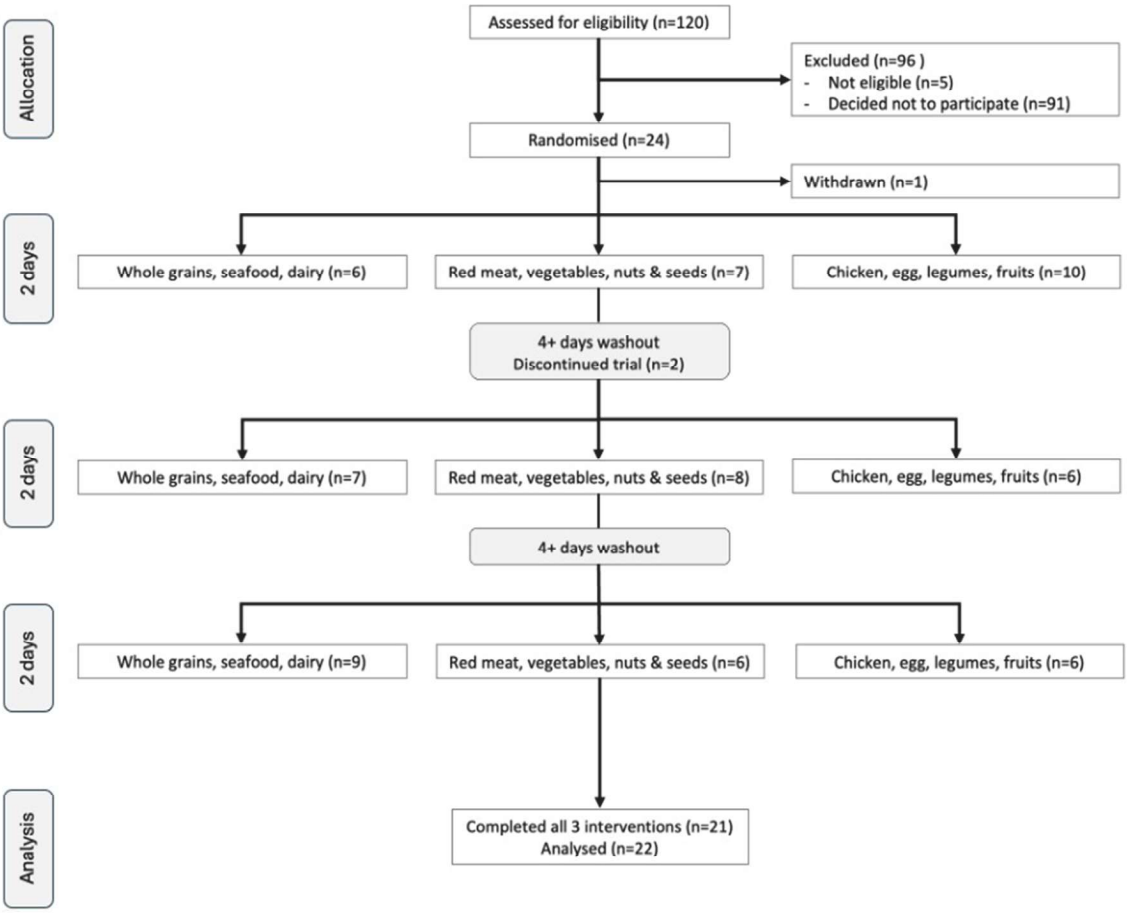

FIGURE 1. Flowchart of participants throughout the feeding trial (n = number of participants). 1 participant consumed study foods but did not provide data needed for analysis and discontinued after 1 intervention period; 1 participant completed 1 full intervention period with full data and was included in analysis and discontinued after 1 intervention period.

TABLE 2
Dietary intake of food groups known intake from intervention days and self-reported 3-d average intake from standard days [mean (SD)].

Nutrients	Intervention day: meat, vegetables, nuts, and seeds ($n = 21$)	Intervention day: fish, seafood, dairy, and whole grains ($n = 22$)	Intervention day: chicken, eggs, legumes, and fruit ($n = 21$)	Self-reported 3-standard days average ($n = 64$)
Energy (kJ)	10,722.1 (2653.0)	12,276.0 (3206.6)	12,202.6 (2876.2)	9280.1 (3599.0)
Water (g)	2793.5 (1050.5)	2250.55 (943.6)	2407.6 (847.4)	2463.8 (1015.4)
Sodium (mg)	4791.9 (1399.3)	3505.7 (1259.5)	3504.7 (1098.5)	3744.0 (2628.0)
Food group serves				
Grains (g)	0.0 (0.0)	392.3 (115.4)	130.7 (101.7)	286.7 (154.7)
Refined grains	0.0 (0.0)	0.0 (0.0)	74.67 (66.4)	197.2 (116.9)
Whole grains	0.0 (0.0)	392.3 (115.4)	56.1 (52.4)	89.5 (115.1)
Fruits (g)	59.8 (36.4)	0.00 (0.00)	91.5 (29.8)	24.1 (33.5)
Fruit juice (mL)	111.2 (98.7)	0.00 (0.00)	168.1 (91.0)	5.0 (19.9)
Vegetables (g)	343.7 (173.7)	58.6 (101.0)	0.0 (0.0)	173.9 (171.3)
Nonstarchy vegetables	446.5 (140.1)	0.3 (0.1)	0.0 (0.0)	181.5 (251.0)
Starchy vegetables	146.9 (79.2)	58.4 (101.0)	0.0 (0.0)	75.7 (121.1)
Red meat (g)	219.6 (114.7)	0.0 (0.0)	0.0 (0.0)	61.6 (95.0)
Processed meat (g)	47.0 (26.9)	0.0 (0.0)	0.0 (0.0)	9.0 (20.7)
Poultry (g)	0.0 (0.0)	0.0 (0.0)	447.8 (130.2)	46.7 (90.0)
Eggs (g)	60.8 (65.4)	0.0 (0.0)	73.0 (31.1)	32.8 (45.0)
Fish (g)	0.0 (0.0)	279.1 (99.0)	0.0 (0.0)	5.6 (19.9)
Nuts and seeds (g)	114.9 (57.4)	114.9 (97.1)	0.0 (0.0)	9.9 (22.4)
Legumes (g)	0.0 (0.0)	0.0 (0.0)	519.2 (131.56)	24.3 (77.4)
Dairy (g)	5.6 (3.2)	223.8 (75.5)	35.3 (33.0)	222.4 (193.7)
Milk	5.6 (3.2)	85.0 (36.7)	0.0 (0.0)	166.7 (154.1)
Cheese	0.0 (0.0)	102.1 (42.7)	35.3 (33.0)	19.3 (26.5)
Yogurt	0.0 (0.0)	36.7 (14.5)	0.0 (0.0)	36.5 (89.6)
Alcoholic beverage (std)	0.0 (0.0)	0.0 (0.0)	0.1 [0.89 (4.18)]	0.3 (0.8)

Abbreviations: std, standard drink; *n*, number of participants.

urineand fruit, resulting in a a negative association, and choline in blood andegg, resulting in a positive association. The existing associations betweenβ-carotene in blood and 3,3'-diindolylmethane and capsaicin in urine with nonstarchy vegetable intake were no longer observed after baseline adjustment. From the Pearson correlation coefficients whencomparing all food groups and BFIs (Supplemental Material 2), therewas also evidence of an association between whole grains, fish, legume,vegetables, and fruit and their corresponding BFIs.

Sensitivity analyses when considering only complete 24-h urine samples (Supplemental Material 2) showed a larger and positive relationship between 3,5-dihydroxybenzoic acid and whole grain intake, 3-carboxy-4-methyl-5-propyl-2-furanpropanoate and trimethylamine N-oxide and fish intake, and sulforaphane, S-methylcysteineand methoxytyramine, and vegetable intake. There was also an association between lactose and dairy intake and a larger but negative association between urinary methoxytyramine adjusted for baseline levels and fruit intake, with no other relationship between remaining food groups and their associated BFIs found when consideringonly complete 24-h urine samples.

Comparisons between known and self-reported dietary intakes (Table 4) showed a strong positive relationship with correlation coefficients $r > 0.55$ between known dietary intake and self-reported dietaryintake of whole grains, fish, fruit, fruit juice, and vegetables. GLM regression coefficients showed a relationship between selfreportedand known intake of all food groups, ranging from a coeffi-cientof 0.0005–0.6074, except for dairy and nuts and seeds. There was no association found between the current reference standard urine sodium excretion and known sodium intake. When converted to dietarysalt intake, urine sodium showed a GLM regression coefficient (95%CI) of 0.11 (−0.06, 0.28), which increased to 0.16 (−0.03, 0.35) whenadjusted for baseline (i.e., standard day) urine sodium concentrations. Sensitivity analyses using complete urine samples only did not improve these associations (Supplemental Material 2). Comparisons between known and self-reported dietary intakes without clustering by feeding day showed a positive relationship for all food groups, with a Pearson's correlation coefficient r of 0.40 for sodium and r ranging from 0.55 to 0.94 for all other food groups (Supplemental Material 2).

Dose-response considerations between food group and metabolite are shown in the Figure 2. There were no noticeable dose-response relationships between dairy and fruits and their associated BFIs. There was a stepwise increase in urinary anserine and 3-methylhistine after 400 g of chicken intake and proline betaine after 200 g of fruit juice intake. Urinary excretion of 3-carboxy-4-methyl-5-propyl-2-furanpropanoate and trimethylamine N-oxide showed a nearly linear relationship with fish intake, dropping after 4 serves (400 g) of intake. Urinary sulforaphane and S-methylcysteine increased per 300 g increase in dietary intake of nonstarchy vegetables, seeming to plateau at 600 g of intake. There was also an increase urinary trigonelline and pipecolic acid between 450 g to 6500 g of dietary intake of legumes. Finally, urinary carnitine and 1-methylhistidine in blood and urine increased after 200 g of red meat intake.

Considering BFIs that were found in both urine and whole blood, pairwise correlation (Supplemental Material 2) illustrates which BFI had a stronger relationship with a food group depending on the biofluid. Hesperetin and 3-methylhistidine showed a stronger correlation within blood, although these did not meet the threshold for significance. All others had a stronger and positive response in urine, with S-methylcysteine and trigonelline meeting the threshold for significance.

Examples of the exploration of the relationships between BFIs in urine and the average self-reported dietary intake of standardization days were done through scatter plots (Supplemental Material 2). Unlike the primary analysis, the 3-d average shows that fruit intake had a relationship with proline betaine in urine with a Pearson's correlation

TABLE 3
Biomarkers of food intake associations with known dietary intakes.

Food/food group	Metabolite name	Known metabolite mean (SD) (ppm)	Tissue	Functional form[1]	GLM coefficient (95% CI) metabolites scaled (*z*-scored)	*P*	GLM coefficient (95% CI) metabolites scaled (*z*-scored) adjusted for baseline	*P*
GLM for a continuous outcome								
Whole grains	Total alkylresorcinols[2] ($n = 39$)	1.03 (4.41)	Blood	Log	−8.27 (−14.61, −1.94)	0.01	−8.49 (−14.38, −2.61)	<0.01
	Alkylresorcinol C17:0 ($n = 39$)	0.05 (0.17)	Blood	Log	12.61 (−20.74, 45.95)	0.46	10.75 (−23.22, 44.72)	0.54
	Alkylresorcinol C19:0 ($n = 39$)	0.03 (0.09)	Blood	Log	4.53 (−88.09, 97.16)	0.92	5.79 (−87.55, 99.12)	0.90
	Alkylresorcinol C20:0 ($n =39$)	0.38 (1.33)	Blood	Log	−57.30 (−89.70, −24.90)	<0.01	−57.28 (−90.35, −24.21)	<0.01
	Alkylresorcinol C21:0 ($n = 39$)	0.60 (3.00)	Blood	Log	−26.66 (−33.25, −20.07)	<0.01	−27.49 (−33.60, −21.38)	<0.01
	Alkylresorcinol C23:0 ($n =39$)	0.02 (0.07)	Blood	Log	−19.41 (−51.87, 13.06)	0.24	−19.09 (−50.86, 12.68)	0.24
	3,5-Dihydroxybenzoic acid ($n =40$)	3.96 (5.42)	Urine	Linear	87.32 (24.28, 150.35)	<0.01	102.08 (26.03, 178.13)	<0.01
	HPAA ($n =43$)	6.39 (10.69)	Urine	Linear	57.76 (−70.85, 186.36)	0.38	107.49 (−104.49, 319.47)	0.32
Wheat bread	Total alkylresorcinols[1] ($n = 39$)	1.03 (4.41)	Blood	Log	−46.28 (−79.70, −12.86)	< 0.01	−42.41 (−77.14, −7.68)	0.02
	Alkylresorcinol C17:0 ($n = 39$)	0.05 (0.17)	Blood	Log	−13.02 (−19.24, −6.80)	< 0.01	−13.18 (−19.87, −6.49)	<0.01
	Alkylresorcinol C19:0 ($n = 39$)	0.03 (0.09)	Blood	Log	−21.03 (−43.10, 1.05)	0.06	−23.73 (−46.57, −0.89)	0.04
	Alkylresorcinol C20:0 ($n =39$)	0.38 (1.33)	Blood	Log	142.46 (−255.56, 540.49)	0.48	138.15 (−275.68, 551.99)	0.51
	Alkylresorcinol C21:0 ($n = 39$)	0.60 (3.00)	Blood	Log	−3.60 (−222.36, 215.15)	0.97	75.63 (−130.36, 281.63)	0.47
	Alkylresorcinol C23:0 ($n =39$)	0.02 (0.07)	Blood	Log	−37.85 (−66.90, −8.81)	0.01	−44.67 (−76.66, −12.67)	0.01
Fish	CMPF ($n = 43$)	4.15 (3.41)	Urine	Linear	62.15 (35.00, 89.29)	<0.01	48.94 (9.61, 88.28)	0.02
	TMAO ($n = 40$)	104.05 (159.67)	Urine	Log	60.10 (39.05, 81.15)	<0.01	67.13 (41.98, 92.29)	<0.01
	TMAO ($n = 39$)	1.39 (3.32)	Blood	Log	13.78 (−35.69, 63.25)	0.59	21.48 (−21.64, 64.61)	0.33
Dairy	Isovalerylglycine ($n = 43$)	173.61 (173.74)	Urine	Linear	−31.26 (−80.55, 18.03)	0.21	−16.52 (−30.97, −2.07)	0.03
	Isovalerylglycine ($n = 39$)	11.59 (33.04)	Blood	Log	−14.76 (−23.88, −5.64)	<0.01	−14.57 (−24.55, −4.59)	<0.01
	Lactose ($n = 42$)	13.20 (20.91)	Urine	Log	15.38 (−18.31, 49.07)	0.37	14.28 (−7.38, 35.94)	0.20
Cheese	Isovalerylglycine ($n = 43$)	173.61 (173.74)	Urine	Linear	0.00 (−0.00, 0.00)	0.53	0.00 (−0.00, 0.00)	0.53
	Isovalerylglycine ($n = 39$)	11.59 (33.04)	Blood	Log	−0.00 (−0.00, 0.00)	0.20	−0.00 (−0.00, 0.00)	0.48
Fruit	Fructose ($n = 41$)	63.63 (129.47)	Urine	Log	2.76 (−5.46, 10.98)	0.51	−16.99 (−43.24, 9.26)	0.21
	Fructose ($n = 41$)	0.34 (1.34)	Blood	Log	5.39 (2.53, 8.25)	<0.01	5.37 (2.51, 8.24)	<0.01
	Hesperetin ($n = 38$)	0.61 (0.78)	Urine	Log	5.12 (−12.26, 22.50)	0.56	12.42 (−9.82, 34.65)	0.61
	Hesperetin ($n = 41$)	0.10 (0.34)	Blood	Log	3.92 (−1.22, 9.07)	0.14	4.10 (−0.60, 8.79)	0.09
	Naringenin ($n = 42$)	0.92 (1.13)	Urine	Log	3.15 (−10.05, 16.34)	0.64	10.37 (−7.41, 28.16)	0.25
	Proline betaine ($n = 42$)	560.59 (638.22)	Urine	Log	−2.27 (−13.57, 9.04)	0.69	−2.22 (−13.65, 9.22)	0.70
	Phloretin ($n = 38$)	1.56 (3.05)	Urine	Log	−5.17 (−16.47, 6.12)	0.37	N/A in model due to collinearity	
	S-methylcysteine ($n = 41$)	0.13 (0.39)	Blood	Non-logged	5.91 (1.24, 10.58)	0.01	5.74 (1.16, 10.32)	0.01
	Epicatechin ($n = 42$)	0.58 (1.28)	Urine	Log	-0.47 (-5.44, 4.51)	0.85	−0.49 (−5.50, 4.53)	0.85
	Methoxytyramine ($n = 42$)	0.44 (0.93)	Urine	Log	−6.45 (−14.05, 1.15)	0.10	−8.11 (−16.21, −0.01)	0.05
Nuts and seeds	γ-tocopherol ($n = 39$)	0.43 (1.71)	Blood	Log	6.25 (−8.13, 20.63)	0.39	3.72 (−10.58, 18.01)	0.61
	α-tocopherol ($n = 41$)	16.77 (46.68)	Blood	Log	−9.00 (−28.65, 10.66)	0.37	−8.62 (−30.12, 12.89)	0.43
Capsicum	Capsaicin ($n = 31$)	0.22 (0.57)	Urine	Linear	0.00 (−0.01, 0.02)	0.83	0.00 (−0.00, 0.00)	0.83
Food/food group	Metabolite name	Known metabolite mean (SD) (ppm)	Tissue	Functional form[1]	GLM rate ratio (95% CI) metabolites scaled (*z*-scored)	*P*	GLM rate ratio (95% CI) metabolites scaled (*z*-scored) adjusted for baseline	*P*
GLM for Poisson outcome								
Chicken	3-Methylhistidine ($n = 41$)	359.64 (450.03)	Urine	Linear	0.1897 (0.0690, 0.3104)	<0.01	0.2617 (0.1130, 0.4104)	<0.01
	3-Methylhistidine ($n = 41$)	8.72 (25.24)	Blood	Log	−0.2912 (−0.6727, 0.0902)	0.14	−0.3090 (−0.6810, 0.0630)	0.10
	Anserine ($n = 40$)	261.78 (416.98)	Urine	Linear	0.2073 (0.0494, 0.3652)	0.01	0.1887 (0.0376, 0.3398)	0.01
	Anserine ($n = 36$)	0.05 (0.09)	Blood	Log	−0.0940 (−0.2934, 0.1053)	0.36	−0.1836 (−0.5420, 0.1768)	0.32
	Carnosine ($n = 41$)	0.10 (0.31)	Blood	Linear	0.1082 (0.0259, 0.1904)	0.01	0.0796 (0.0058, 0.1533)	0.04
Egg	3-Methylhistidine ($n = 41$)	359.64 (450.03)	Urine	Linear	−0.0002 (−0.1512, 0.1508)	0.99	−0.0444 (−0.1892, 0.1004)	0.55
	3-Methylhistidine ($n = 41$)	8.72 (25.24)	Blood	Log	−0.1013 (−4023, 1.996)	0.51	−0.1045 (−0.4070, 0.1980)	0.50
	Choline ($n = 35$)	0.14 (0.25)	Blood	Log	0.0858 (−0.0059, 0.1775)	0.07	0.1377 (0.0645, 0.2109)	< 0.01

(*continued on next page*)

TABLE 3 (*continued*)

Food/food group	Metabolite name	Known metabolite mean (SD) (ppm)	Tissue	Functional form[1]	GLM coefficient (95% CI) metabolites scaled (*z*-scored)	*P*	GLM coefficient (95% CI) metabolites scaled (*z*-scored) adjusted for baseline	*P*
Red meat	1-Methylhistidine ($n = 31$)	209.36 (209.97)	Urine	Linear	0.2653 (−0.0844, 0.6149)	0.14	0.2262 (−0.1862, 0.6387)	0.28
	1-Methylhistidine ($n = 40$)	6.71 (19.91)	Blood	Log	−0.0981 (−0.9976, 0.8015)	0.83	−0.0908 (−1.1630, 0.9814)	0.87
	Carnitine ($n = 41$)	260.98 (284.24)	Urine	Linear	0.2031 (−0.1207, 0.5268)	0.22	0.1768 (−0.1693, 0.5229)	0.32
	Carnitine ($n = 40$)	92.71 (236.78)	Blood	Linear	−0.0751 (−0.3016, 0.1514)	0.52	−0.1087 (−0.3447, 0.1274)	0.37
	Acetyl carnitine ($n = 40$)	33.56 (96.88)	Blood	Linear	−0.1170 (-0.3188, 0.0847)	0.36	−0.1919 (-0.4653, 0.0815)	0.17
	Carnosine ($n = 40$)	0.05 (0.11)	Blood	Log	−0.2082 (−0.6815, 0.2651)	0.39	−0.1871 (−0.6704, 0.2962)	0.45
Legumes	Trigonelline ($n = 42$)	263.86 (279.46)	Urine	Linear	0.0323 (−0.2490, 0.3137)	0.82	−0.0552 (−0.3090, 0.1986)	0.67
	Trigonelline ($n = 41$)	3.43 (9.73)	Blood	Linear	−0.1197 (−0.4570, 0.2176)	0.49	0.3306 (−0.1535, 0.8148)	0.18
	Glycine betaine ($n = 41$)	47.71 (144.42)	Blood	Linear	0.2128 (0.0230, 0.3956)	0.02	0.2151 (0.0508, 0.3794)	0.01
	Pipecolic acid ($n = 40$)	1.52 (2.18)	Urine	Linear	0.1172 (−0.1519, 0.3864)	0.39	0.1300 (−0.1497, 0.4100)	0.36
	Pipecolic acid ($n = 41$)	0.69 (1.98)	Blood	Linear	0.0877 (−0.0150, 0.1904)	0.09	0.0840 (−0135, 0.1814)	0.09
	Quinic acid ($n = 42$)	0.56 (0.87)	Urine	Linear	0.0815 (−0.0398, 0.2027)	0.19	0.0766 (−0.0431, 0.1963)	0.21
Vegetables (nonstarchy)	Sulforaphane ($n = 31$)	22.25 (32.94)	Urine	Linear	0.2954 (0.1996, 0.4712)	<0.01	0.6760 (0.3123, 1.0400)	<0.01
	S-methylcysteine ($n = 29$)	0.24 (0.26)	Urine	Linear	0.2300 (0.1449, 0.4542)	<0.01	0.3002 (0.1649, 0.4354)	<0.01
	S-methylcysteine ($n = 40$)	0.13 (0.39)	Blood	Linear	0.0803 (0.0295, 0.1311)	<0.01	0.0824 (0.0318, 0.1330)	<0.01
	β-carotene ($n = 40$)	2.08 (12.22)	Blood	Linear	0.0494 (0.0238, 0.0750)	<0.01	0.0281 (−0.0194, 0.7559)	0.25
	Gallic acid ($n = 39$)	10.77 (27.73)	Blood	Linear	−0.2311 (−0.6395, 0.1773)	0.27	−0.2544 (−0.7565, 0.2476)	0.32
	3,3'-diindolylmethane ($n = 32$)	0.21 (0.25)	Urine	Linear	−0.7724 (−1.5106, −0.0341)	0.04	−0.8710 (−1.7703, 0.0282)	0.06
	Capsaicin ($n = 41$)	0.22 (0.57)	Urine	Linear	−1.3284 (−2.6552, −0.0017)	0.05	−1.4428 (−2.8982, 0.0126)	0.05
	Methoxytyramine ($n = 41$)	0.48 (0.93)	Urine	Linear	0.2137 (0.0822, 0.3453)	<0.01	0.2290 (0.0861, 0.3720)	<0.01
Broccoli	Sulforaphane ($n = 31$)	22.25 (32.94)	Urine	Linear	0.0000 (−0.0000, 0.0000)	0.56	0.0000 (−0.0000, 0.0000)	0.99
Apple and pear	Phloretin ($n = 38$)	1.56 (3.05)	Urine	Log	0.0945 (0.0529, 0.1361)	<0.001	0.1178 (−0.0128, 0.2485)	0.08
Fruit juice	Proline betaine ($n = 42$)	402.30 (595.84)	Urine	Linear	0.1339 (−0.1753, 0.4432)	0.40	0. 1224 (−0.1959, 0.4407)	0.45
	Hesperetin ($n = 38$)	0.61 (0.78)	Urine	Log	0.0000 (−0.0001, 0.0002)	0.38	0.0000 (−0.0001, 0.0002)	0.38
	Hesperetin ($n = 41$)	0.10 (0.34)	Blood	Log	0.0000 (−0.0000, 0.0000)	0.33	0.0000 (−0.0000, 0.0000)	0.33

Abbreviations: CI, confidence interval; CMPF, 3-carboxy-4-methyl-5-propyl-2-furanpropanoate; GLM, general linear model; HPAA, N-(2-hydroxyphenyl) acetamide; IRR, incidence rate ratio; TMAO, trimethylamine N-oxide.

n: number of data points when accounting for the intervention day with and fully excluding that food group, resulting in 2 data points per participant.

[1] Functional form: linear = where there is a straight-line association between predictor and outcome; log = where there is a log function fit between predictor and outcome.

[2] Includes only alkylresorcinol homologs C17:0, C19:0 and C21:0.

TABLE 4
Self-reported versus known dietary intake on each corresponding feeding day.

Food/ Food groups	Known intake mean (SD) (g)		Self-reported intake mean (SD) (g)		Self-reported vs. known intake pairwise correlation (95% CI)		Self-reported vs. known intake GLM coefficient (95% CI)	*P*
	Main meals	Snacks	Main meals	Snacks	Main meals	Snacks	Main meals	
Whole grains ($n = 21$)	392.3 (115.4)	56.1 (52.4)	251.1 (157.9)	0.0 (0.0)	0.79 (0.48, 1.00)	0.06 (−0.37, 0.50)	0.5610 (0.3505, 0.7717)	<0.01
Chicken ($n = 22$)	447.8 (130.2)	0.0 (0.0)	242.8 (126.1)	0.0 (0.0)	0.21 (−0.23, 0.64)	—	0.0005 (0.0003, 0.0006)	<0.01
Red meat ($n = 21$)	219.6 (114.7)	0.0 (0.0)	191.7 (142.9)	0.0 (0.0)	0.19 (−0.27, 0.64)	—	0.0007 (0.0004, 0.0009)	<0.01
Fish ($n = 21$)	279.1 (99.0)	0.0 (0.0)	233.6 (117.1)	0.0 (0.0)	0.55 (0.41, 1.00)	—	0.6074 (0.3427, 0.8722)	<0.01
Dairy ($n = 21$)	223.8 (75.5)	35.3 (33.0)	189.0 (136.3)	58.8 (85.8)	0.34 (−0.08, 0.76)	0.65 (0.31, 0.98)	0.1880 (−0.0463, 0.423)	0.12
Legume ($n = 22$)	519.2 (131.6)	0.0 (0.0)	176.4 (117.4)	0.0 (0.0)	0.15 (0.09, 0.78)	—	0.0003 (0.0001, 0.0005)	<0.01
Fruit ($n = 22$)	91.5 (29.8)	59.8 (36.4)	103.6 (36.6)	68.2 (48.2)	0.68 (0.36, 0.98)	−0.07 (−0.53, 0.40)	0.5719 (0.3177, 0.82614)	<0.01
Fruit juice ($n = 22$)	168.1 (91.0)	0.0 (0.0)	241.0 (119.9)	0.0 (0.0)	0.58 (0.23, 0.94)	—	0.0026 (0.0023, 0.0029)	<0.01
Vegetables (nonstarchy) ($n = 21$)	446.5 (140.1)	0.3 (0.1)	192.3 (169.3)	3.2 (5.6)	0.55 (0.16, 0.93)	−0.33 (−0.81, 0.14)	0.0009 (0.0008, 0.0010)	<0.01
Nuts and seeds ($n = 21$)	114.9 (57.4)	114.9 (97.1)	49.6 (54.3)	39.9 (38.8)	−0.08 (−0.62, 0.36)	0.76 (0.46, 1.00)	−0.1266 (−0.6122, 0.3591)	0.61
Sodium (mg)[1]	3916.1 (1394.5)		2411.4 (1648.6)		0.40 (0.17, 0.65)		0.35 (0.20, 0.50)	<0.01

Abbreviations: *n*, number of participants; CI, confidence interval; GLM, general linear model; IRR, incidence rate ratio.
[1] Sodium measures not clustered by feeding day type, accounting for all intervention days' sodium intake (n = 64).

coefficient of 0.32 when unscaled and unadjusted for baseline. Furthermore, the effect size from a Gaussian model was larger than when BFIs were compared with known intake (although the relationship with proline betaine was not statistically significant), with regression coefficients (95% CI) of 45.88 (−9.40, 101.16) and 18.04 (6.91, 29.16) for proline betaine and fructose, respectively. Other BFI such as carnitine and 1-methylhistidine showed a similar trend where neither of the BFIs had an association with red meat intake, despite a positive linear relationship.

Discussion

We have conducted a crossover feeding study to evaluate identified BFI against known intakes and self-reported intakes. To do this, we built a panel of metabolite BFI reflecting 9 different food groups and used this panel to assess biomarker concentrations in both fasted blood and 24-h urine samples reflecting each feeding day. Our primary focus was to evaluate the relationships between food-specific biomarkers and their corresponding food group to increase the plausibility of the assessment [16]; however, more general food-related biomarkers were also considered where food-specific markers were not available. We found significant diet–metabolite associations for 7 of the 9 food groups our participants consumed under observation with previously proposed dietary biomarkers. Adjusting for baseline metabolite levels did little to alter the direction of the relationships observed. Perhaps most surprisingly, 24-h urinary sodium was not associated with known dietary sodium intake in this study where foods were consumed under observation. We observed associations above acceptable level ($r >$ 0.40) [23,27] when comparing self-reported intakes with known intakes provided to participants during this study. The accuracy of self-reported intakes in this feeding study is likely a product of the study design, where participants knew they would be asked to recall what they consumed the preceding day, the food consumed was novel and therefore memorable, and often prepared for consumption in front of them.

Several previously noted relationships between biomarkers and food intakes were not identified in the current study. Several studies have shown proline betaine to be a marker of the habitual consumption of citrus fruits [30–32] when measured in isolation. The excretion of proline betaine is known to peak within a few hours of intake and be completely excreted shortly after that [31,33], indicating a 24-h urine sample should capture complete or most proline betaine excretion. This was not observed in the current study. The amount of citrus fruit provided (36 g fruit per breakfast portion and 200 mL of orange juice) may have been insufficient to initiate a detectable response in proline betaine excretion. Contrary to previous works [34–39], the current study did not find associations between red meat intake and 1-methylhistidine, carnosine, or carnitine [34,40]. Although being abundant in beef and pork, carnitine is almost entirely absent from poultry, fish, and dairy indicating potential as a specific BFI of red meat [38,39]. Several other intervention studies have found carnosine concentrations to increase in urine after the ingestion of beef [36,39,41,42], however not in plasma [41,42]. However, carnosine and carnitine are also endogenous metabolites, released in blood and urine as part of normal muscle turnover, which can confound associations with meat intakes. Finally, trigonelline has been proposed as a general candidate BFI for legumes, despite being linked to the consumption of other foods [43]. Again, this was not associated with legume intake in our findings despite an average serving of 519.2 g (SD 131.6) on the relevant feeding day.

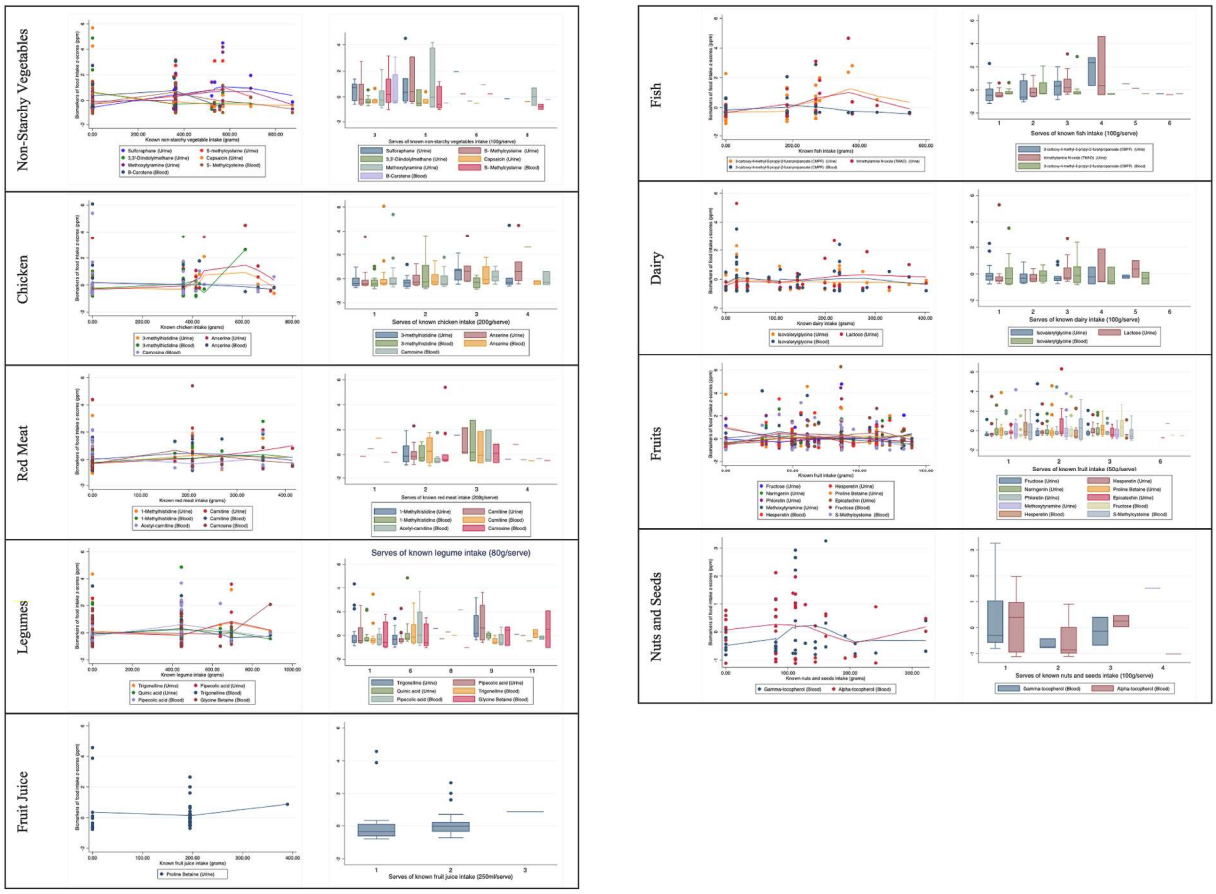

FIGURE 2. Scattered plots and box plots showing relationship between food groups and their associated Biomarkers of Food Intakes by grams or serves.

There were several notable strengths in this trial. The crossover study design was most appropriate, given that it allows each participant to act as their own control, thereby lowering variability associated with physiology, metabolism, and physical activity [44]. Our development of a panel of biomarkers enabled the parallel consideration of a range of markers of the whole diet, which is a novel strength of this work. A more complete capture of the dietary profile increases relevance of BFI in nutrition research and clinical settings. To use this panel, we created a standard sample preparation and detection method using a combination of published protocols along with high-resolution mass and retention time to detect proposed BFIs. Other methodologies covering many BFIs have used multiple reaction monitoring with tandem MS, which has the advantage of wider dynamic range than a time-of-flight detector, although high-resolution mass provides a high degree of specificity, and scanning allows retrospective interrogation of the data as new BFIs are proposed. Food intake levels in our study were designed to reflect a full days' worth of meals and imitate real-life amounts at normal intake levels, increasing generalizability of our results for usual intakes.

Although we were able to identify several potential candidate BFIs within our analysis, there were some potential limitations. Our intervention feeding days were of 1 d duration; however, metabolites with slower half-lives, such as alkylresorcinols, more reflective of habitual intakes may require repeat exposures to foods over several days to show a response relating to dietary intakes. There are also few studies that have used whole blood as a biological specimen to measure metabolites, with plasma or serum being more routinely assessed [6,7,45]. Our use of whole blood instead of plasma or serum may have impacted the sensitivity of detection, particularly due to the high concentration of phospholipids in whole blood. The foods provided in the interventions did not have their own compositional analyses undertaken to measure their metabolite concentration. Due to the focus on a wide range of reported BFIs and a lack of suitable internal standards, only relative quantitation of detected BFIs has been reported. This tends to lead to higher analytical variance, as would the large intra- and intervariability of the bioavailability of some of the proposed metabolites. We did not consider the pharmacokinetics of the proposed biomarkers, known on unknown, other than to not measure phase 2 metabolism conjugates of BFIs (e.g., glucuronide or sulfate conjugates) [7] as we lacked standards for these. The high degree of compound class variation between different BFIs is a complicating factor for a universal method, and we chose to use 2 sets of chromatography to cover a wide range of BFIs, especially as many elute around the solvent front with C18 chromatography [7]. Lastly, participants were free-living and could not be controlled for all activities that may have influenced BFI-dietary intake associations.

There are several limitations of current BFI research that need acknowledging, namely: moving beyond detecting to quantifying food-specific metabolites, nonstandardization of sample handling, lack of dose-response testing, and differing analysis methods. These factors make it currently challenging to understand how candidate BFIs might perform in applied settings. Although there have been several recent reviews summarizing the best quality studies [8,46–49], more work and greater collaborative approaches to research are needed. Notable examples of such collaborations include the FoodBall project [14,15], which provide a start for bridging the gap between researching

biochemical properties of metabolites and their applications in clinical research settings. Future research needs to better understand internal processing of candidate BFIs, such as the impacts of renal function and microbiome on biomarker presence, half-life, and validity as a marker of dietary intakes. Future dietary interventions need to use candidate BFIs as objective measures of compliance and increase the amount of data available in this topic and our understanding of its use [50]. A range of research on multiple tissue types needs to be undertaken to consider cell lines that capture actual and habitual intakes. This current work supports the need to move the field of BFI from biomarker discovery to validation using a variety of study designs and methodologies.

In conclusion, the present feeding study has drawn on previous work detecting potential BFI to develop and apply a panel of metabolites reflecting all main food groups. Our results were similar to previous research supporting the associations between some candidate BFIs and the corresponding food group, but not all. Our data suggest that current proposed biomarkers for red meat and nut and seed intakes were not associated with intake of these foods, and that new biomarkers should be evaluated. Further steps needed in this emerging field include standardized methodologies for targeted metabolomics and dose-response testing to understand biomarker concentrations on consumption of realistic amounts. Although the lure of objective measures of dietary intakes is appealing, with implications for research and clinical practice, more development work and greater collaboration in the field are needed to assess their potential.

Acknowledgments

We thank the study phlebotomist, Ivy Salih, for collecting samples even during a blizzard, and the study participants for their time.

Author contributions

The authors' responsibilities were as follows – AS: involved in study design and conceptualization, methodology, project administration, training of research assistants, collaboration with third parties, laboratory analyses, statistical analyses, data curation and writing (original draft and editing); ABR: involved in methodology, laboratory analyses, supervision and writing (review); RMT: involved in statistical analyses and writing (review); JM: involved in writing (review); ANR: involved in funding the study and study design, methodology, supervision and writing (review); and all authors: read and approved the final manuscript.

Conflict of interest

ANR reports financial support was provided by Health Research Council of New Zealand. ANR reports financial support was provided by Riddet Institute Centre of Research Excellence funding. All other authors report no conflicts of interest.

Funding

This work was supported by a Health Research Council of New Zealand Emerging Researcher First Grant (PI ANR 20/591) and Riddet Institute Centre of Research Excellence funding (PI ANR). AS is funded by a Riddet Institute PhD scholarship, and ANR is funded as the Heart foundation-Hynds Senior Research Fellowship. The study funders had no role in the design, execution, or interpretation of the research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ajcnut.2025.10.017.

References

[1] N. Potischman, J.L. Freudenheim, Biomarkers of nutritional exposure and nutritional status: an overview, J. Nutr. 133 (3) (2003) 873S–874S.
[2] D. Corella, J.M. Ordovás, Biomarkers: background, classification and guidelines for applications in nutritional epidemiology, Nutr. Hosp. 31 (Suppl 3) (2015) 177–188.
[3] E.M. Brouwer-Brolsma, L. Brennan, C.A. Drevon, H. Van Kranen, C. Manach, L.O. Dragsted, et al., Combining traditional dietary assessment methods with novel metabolomics techniques: present efforts by the Food Biomarker Alliance, Proc. Nutr. Soc. 76 (4) (2017) 619–627.
[4] M. Levatte, A.H. Keshteli, P. Zarei, D.S. Wishart, Applications of metabolomics to precision nutrition, Lifestyle Genom 15 (1) (2022) 1–9.
[5] M. Guasch-Ferré, S.N. Bhupathiraju, F.B. Hu, Use of metabolomics in improving assessment of dietary intake, Clin. Chem. 64 (1) (2018) 82–98.
[6] M.C. Playdon, L.F. Tinker, R.L. Prentice, E. Loftfield, K.M. Hayden, L. Van Horn, et al., Measuring diet by metabolomics: a 14-d controlled feeding study of weighed food intake, Am. J. Clin. Nutr. 119 (2) (2024) 511–526.
[7] R. González-Domínguez, M. Urpi-Sarda, O. Jáuregui, P.W. Needs, P. A. Kroon, C. Andrés-Lacueva, Quantitative Dietary Fingerprinting (QDF)—a novel tool for comprehensive dietary assessment based on urinary nutrimetabolomics, J. Agric. Food Chem. 68 (7) (2020) 1851–1861.
[8] E.D. Clarke, J.J. Ferguson, J. Stanford, C.E. Collins, Dietary assessment and metabolomic methodologies in human feeding studies: a scoping review, Adv. Nutr. 14 (6) (2023) 1453–1465.
[9] A.E. McNamara, J. Walton, A. Flynn, A.P. Nugent, B.A. McNulty, L. Brennan, The potential of multi-biomarker panels in nutrition research: total fruit intake as an example, Front. Nutr. 7 (2020) 577720.
[10] M. Jenab, N. Slimani, M. Bictash, P. Ferrari, S.A. Bingham, Biomarkers in nutritional epidemiology: applications, needs and new horizons, Hum. Genet. 125 (5–6) (2009) 507–525.
[11] R. González-Domínguez, O. Jáuregui, P. Mena, K. Hanhineva, F.J. Tinahones, D. Angelino, et al., Quantifying the human diet in the crosstalk between nutrition and health by multi-targeted metabolomics of food and microbiota-derived metabolites, Int. J. Obes. (Lond). 44 (12) (2020) 2372–2381.
[12] D. González-Peña, L. Brennan, Recent advances in the application of metabolomics for nutrition and health, Annu. Rev. Food Sci. Technol. 10 (2019) 479–519.
[13] D.S. Wishart, Y.D. Feunang, A. Marcu, A.C. Guo, K. Liang, R. Vázquez-Fresno, et al., HMDB 4.0: the human metabolome database for 2018, Nucleic Acids Res 46 (D1) (2018) D608–D617.
[14] M.K. Singh, S. Han, S. Ju, J.S. Ranbhise, S. Akter, S.S. Kim, I. Kang, Fruit carbohydrates and their impact on the glycemic index: A study of key determinants., FooDB, Internet]. [cited 25 October 2024]. Available from: https://foodb.ca/.
[15] Food metabolome web portal, FOODBALL: the Food Biomarker Alliance, 2017 [cited 2024 Oct 25]. Available from: https://foodmetabolome.org/.
[16] L.O. Dragsted, Q. Gao, A. Scalbert, G. Vergères, M. Kolehmainen, C. Manach, et al., Validation of biomarkers of food intake-critical assessment of candidate biomarkers, Genes Nutr 13 (2018) 14.
[17] Food Composition Database NZ, New Zealand Food Composition Database, The New Zealand Institute for Plant and Food Research Limited, 14th Edition, Ministry of Health, 2021.
[18] U. Harttig, J. Haubrock, S. Knüppel, H. Boeing, The MSM program: web-based statistics package for estimating usual dietary intake using the Multiple Source Method, Eur. J. Clin. Nutr. 65 (Suppl 1) (2011) S87–S91.
[19] WHO (World Health Organization), Guidelines for the collection of clinical specimens during field investigation of outbreaks WHO/CDS/CSR/EDC/2000, 4, World Health Organization, 2005, pp. 1–51.
[20] E.D. Clarke, M.E. Rollo, C.E. Collins, R.L. Haslam, K. Pezdirc, Urinary biomarkers of dietary intake: a review, Nutr. Rev. 78 (5) (2020) 364–381.
[21] A. Shahin, A. Ross, S. Green, J. Mann, A.N. Reynolds, Are metabolite markers of food intake ready for real-world application? Systematic review with dose response meta-analyses, Trends Food Sci. Technol. 163 (2025) 105190. Available from: https://www.sciencedirect.com/science/article/pii/S0924224425003267.
[22] M. Domínguez-Fernández, Y. Xu, P. Young Tie Yang, W. Alotaibi, R. Gibson, W.L. Hall, et al., Quantitative assessment of dietary (Poly)phenol intake: a high-throughput targeted metabolomics method for blood and urine samples, J. Agric. Food Chem. 69 (1) (2021) 537–554.

[23] R.M. McLean, V.L. Farmer, A. Nettleton, C.M. Cameron, N.R. Cook, N.R. C. Campbell, Assessment of dietary sodium intake using a food frequency questionnaire and 24-hour urinary sodium excretion: a systematic literature review, J. Clin. Hypertens. 19 (12) (2017) 1214–1230.
[24] K.L. MacKenzie-Shalders, N.A. King, N.M. Byrne, G.J. Slater, Increasing protein distribution has no effect on changes in lean mass during a rugby preseason, Int. J. Sport Nutr. Exerc. Metab. 26 (1) (2016) 1–7.
[25] A.E. Mcnamara, X. Yin, C. Collins, L. Brennan, Metabolomic based approach to identify biomarkers of broccoli intake, Food Funct. 14 (18) (2023) 8586–8596.
[26] F. Paradis, B. Lamarche, J. Robitaille, C. Couillard, J Lafrenière, A. J. Tremblay, et al., Validation of an automated self-administered 24-hour dietary recall web application against urinary recovery biomarkers in a sample of French-speaking adults of the province of Québec, Canada, Appl. Physiol. Nutr. Metab. 47 (2) (2022) 173–182.
[27] L.S. Freedman, J.M. Commins, J.E. Moler, W. Willett, L.F. Tinker, A. F. Subar, et al., Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for potassium and sodium intake, Am. J. Epidemiol. 181 (7) (2015) 473–487.
[28] R.J. Feise, Do multiple outcome measures require p-value adjustment? BMC Med. Res. Methodol. 2 (2002) 8.
[29] StataCorp. StataCorp, 2021.
[30] G. Favé, M. Beckmann, A.J. Lloyd, S. Zhou, G. Harold, W. Lin, et al., Development and validation of a standardized protocol to monitor human dietary exposure by metabolite fingerprinting of urine samples, Metabolomics 7 (4) (2011) 469–484.
[31] S.S. Heinzmann, I.J. Brown, Q. Chan, M. Bictash, M.E.E. Dumas, S. Kochhar, et al., Metabolic profiling strategy for discovery of nutritional biomarkers: proline betaine as a marker of citrus consumption, Am. J. Clin. Nutr. 92 (2) (2010) 436–443.
[32] A.J. Lloyd, M. Beckmann, G. Favé, J.C. Mathers, J. Draper, Proline betaine and its biotransformation products in fasting urine samples are potential biomarkers of habitual citrus fruit consumption, Br. J. Nutr. 106 (6) (2011) 812–824.
[33] M.B.S. Andersen, M. Kristensen, C. Manach, E. Pujos-Guillot, S.K. Poulsen, T.M. Larsen, et al., Discovery and validation of urinary exposure markers for different plant foods by untargeted metabolomics, Anal. Bioanal. Chem. 406 (7) (2014) 1829–1844.
[34] L. Bernard, J. Chen, H. Kim, K.E. Wong, L.M. Steffen, B. Yu, et al., Serum metabolomic markers of protein-rich foods and incident CKD: results from the atherosclerosis risk in communities study, Kidney Med 6 (4) (2024) 100793.
[35] F. Wang, P.D. Chandler, O.A. Zeleznik, K. Wu, Y. Wu, K. Yin, et al., Plasma metabolite profiles of red meat, poultry, and fish consumption, and their associations with colorectal cancer risk, Nutrients 14 (5) (2022) 978.
[36] H. Abe, E. Okuma, H. Sekine, A. Maeda, S. Yoshiue, Human urinary excretion of L-histidine-related compounds after ingestion of several meats and fish muscle, Int. J. Biochem. 25 (9) (1993) 1245–1249.
[37] M. Gil-Agustí, J. Esteve-Romero, S. Carda-Broch, Anserine and carnosine determination in meat samples by pure micellar liquid chromatography, J. Chromatogr. A. 1189 (1–2) (2008) 444–450.
[38] H. Abe, Distribution of free L-histidine and related dipeptides in the muscle of fresh-water fishes, Comp. Biochem. Physiol. B. 76 (1) (1983) 35–39.
[39] W.D. Block, R.W. Hubbard, B.F. Steele, Excretion of histidine and histidine derivatives by human subjects ingesting protein from different sources, J. Nutr. 85 (1965) 419–425.
[40] C. Cuparencu, G. Praticó, L.Y. Hemeryck, P.S.C. Sri Harsha, S. Noerman, C. Rombouts, et al., Biomarkers of meat and seafood intake: an extensive literature review, Genes Nutr 14 (1) (2019) 35.
[41] K.J. Yeum, M. Orioli, L. Regazzoni, M. Carini, H. Rasmussen, R.M. Russell, et al., Profiling histidine dipeptides in plasma and urine after ingesting beef, chicken, or chicken broth in humans, Amino Acids 38 (3) (2010) 847–858.
[42] W. Cheung, P. Keski-Rahkonen, N. Assi, P. Ferrari, H. Freisling, S. Rinaldi, et al., A metabolomic study of biomarkers of meat and fish intake, Am. J. Clin. Nutr. 105 (3) (2017) 600–608.
[43] N.D. Willis, A.J. Lloyd, L. Xie, M. Stiegler, K. Tailliart, I. Garcia-Perez, et al., Design and characterisation of a randomized food intervention that mimics exposure to a typical UK diet to provide urine samples for identification and validation of metabolite biomarkers of food intake, Front Nutr 7 (2020) 561010.
[44] K.A. Guertin, E. Loftfield, S.M. Boca, J.N. Sampson, S.C. Moore, Q. Xiao, et al., Serum biomarkers of habitual coffee consumption may provide insight into the mechanism underlying the association between coffee consumption and colorectal cancer, Am. J. Clin. Nutr. 101 (5) (2015) 1000–1011.
[45] M.M. Ulaszewska, C.H. Weinert, A. Trimigno, R. Portmann, C. Andres Lacueva, R. Badertscher, et al., Nutrimetabolomics: an integrative action for metabolomic analyses in human nutritional studies, Mol. Nutr. Food Res. 63 (1) (2019) e1800384.
[46] T. Rafiq, S.M. Azab, K.K. Teo, L. Thabane, S.S. Anand, K.M. Morrison, et al., Nutritional metabolomics and the classification of dietary biomarker candidates: a critical review, Adv. Nutr. 12 (6) (2021) 2333–2357.
[47] S. Noerman, R. Landberg, Blood metabolite profiles linking dietary patterns with health—toward precision nutrition, J. Intern. Med. 293 (4) (2023) 408–432.
[48] S. Andraos, K.L. Beck, M.B. Jones, T.L. Han, C.A. Conlon, J.V. De Seymour, Characterizing patterns of dietary exposure using metabolomic profiles of human biospecimens: a systematic review, Nutr. Rev. 80 (4) (2022) 699–708.
[49] R.L. Prentice, Intake biomarkers for nutrition and health: review and discussion of methodology issues, Metabolites 14 (5) (2024) 276.
[50] A.N. Reynolds, F. Hood, R. Wilson, A. Ross, S. Neumann, R. Turner, et al., Healthy grocery delivery in the usual care for adults recovering from an acute coronary event: protocol for a three-arm randomised controlled trial, BMJ Open 13 (11) (2023) 1–9.