



Cite this: *Food Funct.*, 2023, **14**, 9635

## Development of a novel (poly)phenol-rich diet score and its association with urinary (poly)phenol metabolites†

Yifan Xu,‡ Yong Li,‡ Jiaying Hu, Rachel Gibson and Ana Rodriguez-Mateos  \*

**Background:** Estimating (poly)phenol intake is challenging due to inadequate dietary assessment tools and limited food content data. Currently, *a priori* diet scores to characterise (poly)phenol-rich diets are lacking. This study aimed to develop a novel (poly)phenol-rich diet score (PPS) and explore its relationship with circulating (poly)phenol metabolites. **Methods:** A total of 543 healthy free-living participants aged 18–80 years completed a food frequency questionnaire (FFQ) (EPIC-Norfolk) and provided 24 h urine samples. The PPS was developed based on the relative intake (quintiles) of 20 selected (poly)phenol-rich food items abundant in the UK diet, including tea, coffee, red wine, whole grains, chocolate and cocoa products, berries, apples and juice, pears, grapes, plums, citrus fruits and juice, potatoes and carrots, onions, peppers, garlic, green vegetables, pulses, soy and soy products, nuts, and olive oil. Foods included in the PPS were chosen based on their (poly)phenol content, main sources of (poly)phenols, and consumption frequencies in the UK population. Associations between the PPS and urinary phenolic metabolites were investigated using linear models adjusting energy intake and multiple testing (FDR adjusted  $p < 0.05$ ). **Result:** The total PPS ranged from 25 to 88, with a mean score of 54. A total of 51 individual urinary metabolites were significantly associated with the PPS, including 39 phenolic acids, 5 flavonoids, 3 lignans, 2 resveratrol and 2 other (poly)phenol metabolites. The total (poly)phenol intake derived from FFQs also showed a positive association with PPS (stdBeta 0.32, 95% CI (0.24, 0.40),  $p < 0.01$ ). Significant positive associations were observed in 24 of 27 classes and subclasses of estimated (poly)phenol intake and PPS, with stdBeta values ranging from 0.12 (0.04, 0.20) for theaflavins/thearubigins to 0.43 (0.34, 0.51) for flavonols ( $p < 0.01$ ). **Conclusion:** High adherence to the PPS diet is associated with (poly)phenol intake and urinary biomarkers, indicating the utility of the PPS to characterise diets rich in (poly)phenols at a population level.

Received 17th May 2023,  
Accepted 19th September 2023

DOI: 10.1039/d3fo01982a  
[rsc.li/food-function](http://rsc.li/food-function)

## Introduction

A suboptimal diet is one of the most important modifiable risk factors for non-communicable diseases (NCDs). Among all the components of the human diet, plant foods are the most essential ones relating to the prevention of NCDs. Four out of five leading dietary risk factors for death and disability-adjusted life-years (DALYs) are related to inadequate intake of plant foods, including wholegrains, fruits, nuts and seeds, and vegetables.<sup>1,2</sup> Evidence from large prospective cohort studies and meta-analysis of cohort studies showed a protective effect of adequate consumption of these plant foods. Consuming

five portions of fruits and vegetables per day was found to be associated with a lowering in total mortality of 13%, cardiovascular disease (CVD) mortality of 12%, cancer mortality of 10% and respiratory disease mortality of 35%, compared to two portions of daily intake.<sup>3</sup> Eating wholegrains instead of refined grains was associated with a lower risk of CVD<sup>4</sup> and type 2 diabetes.<sup>5,6</sup> Higher nut consumption has also been linked to a lower relative risk of CVD mortality, cancer mortality, and all-cause mortality.<sup>7</sup> The health benefits of plant-based foods have been well established, although the mechanisms behind their effect are still unclear. In recent decades, the bioactive effects of many phytochemical compounds present in these foods, such as (poly)phenols, carotenoids, phytosterols, and so on have been recognized and have gained growing attention in nutritional research due to their potential effects to promote overall health and lower NCD risk.<sup>8</sup>

(Poly)phenols are a large family of compounds naturally existing in plants and they are widely distributed in our diet. Accumulating evidence from both clinical trials and epidemiological

Department of Nutritional Sciences, School of Life Course and Population Sciences, Faculty of Life Sciences and Medicine, King's College London, London, UK.  
E-mail: [ana.rodriguez-mateos@kcl.ac.uk](mailto:ana.rodriguez-mateos@kcl.ac.uk)

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d3fo01982a>

‡ Co-first authors contributed equally to this work.



logical studies suggests that dietary (poly)phenols can improve cardiometabolic health.<sup>9–12</sup> However, it is still difficult to give dietary recommendations on (poly)phenol consumption to promote cardiometabolic health due to the lack of consistent results from observational studies with long-term intake and the inaccurate estimation of habitual (poly)phenol intake in the free-living population. Assessing (poly)phenol consumption is extremely challenging due to inadequate dietary assessment tools and limited food content data. Being a quick and efficient data collection method to measure food groups and nutrient intake in epidemiological studies, food frequency questionnaires (FFQs) have been widely used to estimate (poly)phenol intake in multiple large cohorts and surveys such as the Nurses' Health Study and the Health Professionals Follow-Up Study,<sup>13</sup> or the European Prospective Investigation into Cancer and Nutrition (EPIC).<sup>14,15</sup> These studies make up a substantial part of the current evidence on the health benefits of dietary (poly)phenols. However, most of the questionnaires applied in these research studies were not specifically validated for estimating (poly)phenol consumption.<sup>16</sup> Regarding the existing data on the (poly)phenol content of foods, many food items and compounds are still missing in most comprehensive databases available such as Phenol-Explorer<sup>17</sup> and USDA databases.<sup>18–20</sup> This is because the analysis of (poly)phenol content in foods requires accurate chromatography analytical methods and authentic standards, which are not easily accessible. In addition, food items need to be mapped carefully to the (poly)phenol content. Errors could also be introduced systematically during this process, especially when available data are limited.<sup>16</sup>

Another approach widely used to investigate relationships between diet and health is the analysis of dietary patterns instead of single nutrient/bioactive intake. A number of diet indices have been developed in recent years to reflect adherence to certain dietary patterns based on habitual intake and diet quality related to health,<sup>21</sup> such as the Healthy Eating Index (HEI),<sup>22</sup> the Mediterranean Diet Score (MDS),<sup>23</sup> the Dietary Approaches to Stop Hypertension (DASH)<sup>24</sup> and the Plant-based Diet Index (PDI).<sup>25</sup> Evidence suggests that better compliance to these dietary patterns is associated with a lower risk of cardiovascular diseases.<sup>26</sup> Plant foods are key components of these diet scores, and it is possible that the (poly)phenols present in plant foods may mediate the protective effects on cardiometabolic health, as proposed by us recently.<sup>27</sup> However, none of the currently existing dietary indices are specifically focused on (poly)phenol-rich foods and beverages or aim to estimate adherence to (poly)phenol rich diets. Indeed, while fruits and vegetables are included in most healthy dietary scores such as DASH, MDS, HEI and PDI, they are usually grouped together despite their distinct (poly)phenol profiles. In addition, tea and coffee, which are major sources of (poly)phenols, are not included in most dietary scores, except for the PDI score, which includes both as one food group, despite having very distinct profiles, one being rich in flavan-3-ols and the other in phenolic acids. Cocoa products, which are also good sources of dietary (poly)phenols,

have not been included in any of the previously established dietary scores.

(Poly)phenols are a large and diverse class of compounds, with multiple types of (poly)phenol subclasses being found in the same foods and beverages. It is therefore difficult to attribute the health benefits related to the consumption of a (poly)phenol rich food to a single compound. Evidence of health benefits exists for all the flavonoid subclasses and other types of (poly)phenols such as phenolic acids or lignans,<sup>28</sup> although the evidence is stronger for some, such as flavan-3-ols. This can be due to a larger body of evidence existing for certain compounds, rather than certain compounds having higher bioactivity than others, although this is currently unknown. Therefore, creating an overall score to estimate adherence to a (poly)phenol rich diet is a suitable approach to determine all potential bioactive compounds within the diet.

This study aims to develop a (poly)phenol-rich diet score to reflect habitual consumption of a diet rich in (poly)phenols and explore the association between this score and a comprehensive panel of (poly)phenol metabolite levels in 24 h urine.

## Methods

### Study population

The POLYNTAKE cohort is comprised of free-living healthy participants residing in the UK, who participated in nine dietary intervention trials at King's College London. The design and data analysis process of the POLYNTAKE study has been previously reported.<sup>29</sup> Dietary assessment and 24 h urine collection were conducted during the baseline visit of the studies, prior to any intervention. Data were collected and analysed following the same protocols and tools across the nine studies and integrated into this cross-sectional analysis. All studies were registered and approved by the ethics committee of King's College London (ethics number/clinical trial registration number: RESCM-17/18-5283/NCT03434574; HR-15/16-3739/NCT03041961; HR-17/18-5338/NCT03592966; HR-18/19-9091/NCT04084457; HR-17/18-5703/NCT03553225; RESCM-18/19-9036/NCT03995602; HR-17/18-5353/NCT03573414; HR-18/19-8999/NCT04179136; HR-19/20-14771/NCT04276974). The studies were conducted following the Declaration of Helsinki and written informed consent was provided by all participants before their participation. All participants consented to their data being used in follow-up research.

To eliminate the influence of age and outliers in dietary intake, participants were excluded for the following reasons: (i) age < 18 years old; (ii) no available FFQ or FFQ had more than 10 missing ticks; (iii) energy intake <500 kcal day<sup>-1</sup> or >3500 kcal day<sup>-1</sup> for women, <800 kcal day<sup>-1</sup> or >4000 kcal day<sup>-1</sup> for men; and (iv) energy intake to BMR ratio out of mean  $\pm$  2SD (0.025–2.437 for the current dataset) in the study population. Data from 543 participants were included in the analysis of this paper. Among them, urinary (poly)phenol metabolite excretion levels were available for a subgroup of 229 participants.



### Dietary assessment with FFQ

The EPIC-Norfolk FFQ<sup>30</sup> was used to collect the habitual diet of participants during the first study visit before any intervention was given. Participants were invited to provide the dietary intake information of 130 food items commonly consumed in the UK over the past year by choosing the frequency ranging from “never or less than once a month” to “2–3 times per day”. The FFQs were coded using Microsoft Access software and the frequencies in the questions were transformed into nutrients and food items intake using the FFQ EPIC Tool for Analysis (FETA) software.<sup>31</sup> One or more food codes were mapped to each question in the FFQ with the portion weights sourced from UK population data during FETA analysis.<sup>31</sup> The McCance and Widdowson’s “The composition of Foods (5<sup>th</sup> edition)” and its supplements<sup>32</sup> were applied in the calculation of nutrients and energy intake.

### Estimation of dietary (poly)phenol intake

The (poly)phenol intake of the participants was calculated from the food intake (g d<sup>-1</sup>) obtained from the EPIC-Norfolk FFQ coupled with the corresponding content in foods (mg per 100 g) from a (poly)phenol database (PPDB) developed in-house.<sup>29</sup>

The PPDB integrated (poly)phenol content data of 1260 raw and processed food items or dishes, which were obtained from multiple sources such as the Phenol-Explorer database,<sup>17</sup> USDA databases,<sup>18–20</sup> and published analytical data. The (poly)phenol contents for composite dishes were calculated based on recipes from McCance and Widdowson’s (6<sup>th</sup> edition), see the ESI<sup>†</sup><sup>33</sup> and retailer websites. The food codes from FETA and food items in PPDB were matched as precisely as possible to their subtypes according to the food descriptions. If no specific subtype of food was described (e.g. onions, raw), the content of a general food content was matched to it (onions, raw (average)). The food items with little or no (poly)phenol content (e.g. animal products) were removed from the calculation. Total subclasses, classes, and total (poly)phenol intake were calculated by summarising the intake of all the compounds under the group. Details of (poly)phenol analysis process and PPDB have been reported previously.<sup>29</sup>

### Development of the (poly)phenol-rich diet score

To estimate adherence to a (poly)phenol-rich diet, a (poly)phenol-rich diet score (PPS) was developed. The components of the PPS were selected based on the following workflow (ESI Fig. 1†). Firstly, a list of (poly)phenol-rich food items was proposed based on their (poly)phenol content from established literature<sup>34</sup> and our in-house PPDB.<sup>29</sup> The food items with at least 30 mg per 100 g total (poly)phenols (aglycones equivalent, quantified by chromatography method) or 20 mg per 100 g for certain subclasses of (poly)phenols were included. Next, the list was cross-referenced against the food sources of (poly)phenols estimated from 7-day food diaries in the same study cohort<sup>29</sup> and data from the general UK population<sup>35</sup> to ensure inclusion of widely consumed food items. Then, the

food items that were not included in the EPIC-Norfolk FFQ were removed (*i.e.*, blueberries, herbs & spices, and flaxseeds). Then, some food items were grouped together due to the similarity of (poly)phenol profiles, *e.g.*, citrus fruits and juices, berries, whole grains, and green vegetables. Potatoes and carrots were grouped together as they were relatively less rich in (poly)phenol but contributed a substantial amount to total (poly)phenol intake due to the high consumption in the UK.

A total of 20 plant-based foods from the EPIC-Norfolk FFQ that are important sources of (poly)phenol intake in the UK diet were included in the PPS. These food or food groups include tea, coffee, red wine, whole grains, chocolate and cocoa products, berries, apples and apple juice, pears, grapes, plums, citrus fruits and citrus juice, potatoes and carrots, onions, peppers, garlic, green vegetables, pulses, soybeans, and related products, nuts, and olive oil. Table 1 shows the FFQ items included under the 20 food groups and their total (poly)phenol content (mg per 100 g).

Participants were scored by the quintiles of their intake of each food group in the study population. The participants in the highest quintile scored 5 in this food group, and the participants in the lowest quintile scored 1. To calculate the score, we used relative intake rather than absolute intake since there is not yet enough evidence to propose adequate intake levels for all food items that provide health benefits from (poly)phenols. The PPS was calculated as the total score of all the 20 food group scores, which ranged from 20 to 100. Equal weightage was given to all food items, since there is still limited understanding of the differential effects of these foods on health.

### Urine sample collection and phenolic metabolite analysis

The 24h-urine samples collected at the baseline visit of the clinical trials were used to analyse (poly)phenol metabolite levels. The participants were instructed to collect urine starting from the second urine of the day before the study visit and finishing with the first urine on the day of the study visit in 1–2 plastic containers (2 L each) and store them in the cool bag along with icepacks provided by the researchers. After recording the total volume excreted over 24 h, urine samples were centrifuged at 1800g for 15 min at 4 °C and were spiked with 0.2% formic acid (Thermo Fisher, LC-MS grade, Loughborough, UK) before storing at –80 °C in labelled plastic tubes until analysis.

The processing and analysis of the urine samples followed a validated method.<sup>36</sup> Briefly, samples were thawed on ice for 0.5–1 hour and then centrifuged at 15 000g for 15 min at 4 °C using a temperature controlled microtube centrifuge (5417R, Eppendorf, Hamburg, Germany). The urine samples were diluted 5 fold with HPLC water (Sigma Aldrich, Steinheim, Germany) before the diluted samples (350 µL) were acidified with 4% phosphoric acid (85% HPLC grade, Yorlab, Fluka, York, UK) (v:v 1:1). An aliquot of 600 µL of the mixture was loaded on to the Oasis HLB reversed-phase sorbent µ-SPE 96-well plate (Waters, Eschborn, Germany) and washed with HPLC water (200 µL) and 0.2% acetic acid (200 µL) (glacial



**Table 1** Composition of the (poly)phenol-rich diet score and estimated (poly)phenol content of the food component based on the KCL polyphenol database (PPDB)

Foods	Relevant items in the EPIC-Norfolk FFQ	Total (poly)phenol content (mg per 100 g) <sup>17</sup>
Tea	Tea, black, infusion, average	Black tea (94.96), green tea (87.79)
Coffee	Coffee, infusion, average	Coffee, infusion (316.00)
Red wine	Rose wine, medium <sup>a</sup>	Wine, rose (12.83), wine, red (88.32)
Whole grains	Brown rice, boiled; wholemeal bread, average; crispbread, rye; spaghetti, wholemeal, boiled; porridge, made with water; all breakfast cereals <sup>b</sup>	Brown rice (95.86), wholemeal bread (24.80), crispbread (182.77), wholemeal pasta (43.73), breakfast cereals, bran (285.70), breakfast cereals, muesli (13.75)
Chocolate and cocoa products	Chocolate, fancy and filled <sup>c</sup> ; drinking chocolate powder	Milk chocolate (236.10), dark chocolate (1639.51), drinking chocolate powder (875.15), drinking chocolate powder, made up (289.16)
Berries	Raspberries, raw; strawberries, raw	Raspberries (189.88), strawberries (268.13), blueberries (420.99)
Apples and juice	Apples, eating, average, raw, flesh and skin weighted; apple juice, unsweetened; chutney, apple, homemade	Apples (138.25), apple juice (68.49), apple chutney (155.67)
Pears	Pears, average, raw, peeled or not peeled, weighed with core	Pears (35.61)
Grapes	Grapes, average <sup>d</sup>	Grapes, green (91.60), grapes, black (128.17)
Plums	Plums, average, raw	Plums (366.38)
Citrus fruit and juice	Oranges, weighed with peel and pips; orange juice, unsweetened; grapefruit, raw	Oranges, blond (50.47); orange juice (65.34), grapefruit (73.79)
Potatoes and carrots	Salad potato, with mayonnaise or reduced calorie dressing; potatoes, roast, fat removed; chips, straight cut, fat removed; chips, retail, fried in vegetable oil; old potatoes, boiled in salted water; carrots, old or young, boiled in salted water	Potato, boiled (25.04), potato chips (21.28), carrots, boiled (41.83)
Onions	Onions, raw	Red onion, raw (25.48), yellow onion, raw (15.02)
Peppers	Peppers, capsicum, green or red, raw	Red sweet pepper (14.19), green sweet pepper (19.54), chilli pepper, green (21.81), chilli pepper, yellow (37.40)
Garlics	Garlic, raw	Garlic (184.94)
Green vegetables	Spinach, boiled in salted water; broccoli, green, boiled in salted water; brussels sprouts, boiled in salted water	Spinach, boiled (79.01), broccoli, boiled (177.68), brussels sprouts (7.12)
Pulses	Peas, frozen, boiled in salted water; peas, canned, re-heated, drained; split peas, dried, boiled in unsalted water; baked beans, canned in tomato sauce; broad beans, boiled in salted water; green beans/French beans, boiled in salted water; lentils, red, split, dried, boiled in unsalted water; runner beans, boiled in salted water	Broad beans, boiled (155.63), green beans, boiled (37.08), common beans, black, boiled (73.66), common beans, white, boiled (53.26), common beans, others (619.06)
Soy and soy products	Tofu, soya bean steamed; soya mince, granules; vegeburger, retail, fried in vegetable oil; soya milk, plain	Soya beans, boiled (201.92), tofu (20.30), soy meat (13.19), soya milk (10.28)
Nuts	Hazelnuts; peanut butter, smooth; peanuts, roasted and salted	Hazelnuts (496.51), peanut (10.77), peanut butter (11.21)
Olive oil	Olive oil; fat spread (60% fat), with olive oil	Olive oil (61.23), olive oil spread (6.73)

<sup>a</sup>The item consumption includes all types of wines. This item corresponds to "wine, rose" in nutrients and (poly)phenol analysis. <sup>b</sup>Breakfast cereals included in EPIC-Norfolk FFQ: all brans, beanbuds, branflakes, cereal non-specific, cocopops, CommonSense Oat Bran Flakes, cornflakes, crunchy oat cereal, crunchy nut cornflakes, frosties, fruit n fibre, grapanuts, honey smacks, muesli, nutri-grain, oat and wheat bran, puffed wheat, rasin splitz, readybreak, rice crispies, ricicles, shredded wheat, shreddies, special K, start, sugar puffs, sultana bran, weetabix, weetaflakes, and weetos. <sup>c</sup>The item consumption includes all chocolates, single or squares. The item corresponds to "chocolate, fancy and filled" in nutrients and (poly)phenol analysis. <sup>d</sup>The item consumption includes all types of grapes. The item corresponds to "grapes, green" in nutrient and (poly)phenol analysis.

HPLC grade, Thermo Fisher Scientific, Loughborough, UK) into the waste plate. The elusion was conducted with 30  $\mu$ L of methanol containing 0.1% formic acid and 10 nM ammonium formate (HPLC grade, Sigma Aldrich, Steinheim, Germany) 3 times (90  $\mu$ L in total). There is an additional 35  $\mu$ L of water with 5  $\mu$ L of internal standard (taxifolin, concentration 0.25 mg  $\text{mL}^{-1}$ ) added to the collection plate, making the final volume 130  $\mu$ L.

A total of 110 (poly)phenol compounds were identified and quantified using authentic chemical standards. The UPLC-MS analysis of the samples and standard mixes was achieved using a triple-quadrupole mass spectrometer (SHIMADZU 8060, Shimadzu, Kyoto, Japan) coupled with a UPLC system (Shimadzu, Kyoto, Japan). The samples (5  $\mu$ L) were injected

using an autosampler (SIL-30AC, Shimadzu, Kyoto, Japan) through a Raptor Biphenyl column 2.1  $\times$  50 mm, 1.8  $\mu$ m (Restek, Bellefonte, USA) coupled with a compatible guard cartridge 5  $\times$  2.1 mm, 2.7  $\mu$ m (Restek, Bellefonte, USA) before reaching a heated ESI source. The mobile phases were water (HPLC grade, Sigma Aldrich, Steinheim, Germany) and acetonitrile (HPLC grade, Sigma Aldrich, Steinheim, Germany) both acidified with 0.1% formic acid (LC-MS grade, Thermo Fisher Scientific, Loughborough, UK) as solvents A and B, respectively. The gradient was 14 minutes joint with a 2-minute equilibration phase and applied under a 0.5  $\text{mL min}^{-1}$  flow rate at 30 °C. The MS parameters and multiple reaction monitoring (MRM) method parameters of the target compounds were detailed previously.<sup>36</sup> The peak area ratios of the target com-



pounds to the internal standard taxifolin were used in the quantification to minimise the influence of changes in device performance on the results. The LabSolutions software (SHIMADZU, Kyoto, Japan) was used in the peak integration and the Microsoft Excel (Excel 2020, Microsoft, USA) was used for concentration calculation.

### Statistical analysis

The statistical analysis was conducted using R (version 4.1.2).<sup>37</sup> The PPS of the study population was reported as the mean (standard deviation, SD). The PPS was calculated from the relative distribution of the intake of 20 (poly)phenol-rich foods obtained by FFQs. Since the PPS was normally distributed in the population, one-way ANOVA was conducted to compare the PPS between age groups and the *post-hoc* analysis was conducted using the Tukey HSD test. The correlation between age and PPS was assessed by Person's correlation coefficients. The correlation between estimated total (poly)phenol intake and PPS was assessed by Spearman's correlation coefficients. In order to assess the concordance in discerning between high and low adherence to the (poly)phenol-rich diet using PPS and estimated total (poly)phenol intake, the study participants were ranked into quartiles and depicted in an alluvial diagram.

The associations between PPS, the (poly)phenol-rich food, and dietary (poly)phenol intake, and nutrient intake were explored using a linear regression model with two covariates, energy intake levels and trial effect. The energy intake (kcal

$\text{d}^{-1}$ ) was collected using the EPIC-Norfolk FFQ. Participants from nine trials were labelled with the corresponding sequence number which was included as the categoric variable from 1 to 9 to avoid bias across trials and set as a trial effect. The *p*-values were adjusted for multiple comparisons by the false discovery rate (FDR) method.

The relationships between individual urine phenolic metabolite levels and the PPS and its components were evaluated using linear regression models. The metabolite levels were log-transformed and adjusted for batch effect using the ComBat method<sup>38</sup> with the sva package in R before entering the model. The ComBat method is an empirical Bayes method developed originally for removing batch effect in the microarray data in gene sequencing, and now it has been applied in metabolomics analysis.<sup>39</sup> The energy intake levels estimated from FFQs were adjusted as confounders in the linear regression model. The *p*-values were adjusted for multiple comparison by the FDR method.

## Results

### Characteristics of the study population

The included participants were  $42.1 \pm 18.4$  years old, with an average BMI of  $23.8 \pm 3.5 \text{ kg m}^{-2}$ . The baseline characteristics of the participants are shown in Table 2. There were 314 women (57.8%) and 229 men (42.2%). In general, the participants had a healthy lifestyle, with 70.8% having high levels of

**Table 2** Characteristics of the study population (*n* = 543)

Baseline characteristics	Men ( <i>n</i> = 229)	Women ( <i>n</i> = 314)	Total ( <i>n</i> = 543)	Missingness (%)
Age (years)	39.6 (18.1)	43.9 (18.5)	42.1 (18.4)	0
Age group (%)				
18–34	118 (51.5)	137 (43.6)	255 (41.8)	0
35–49	40 (17.5)	48 (15.3)	88 (14.4)	0
50–64	34 (14.8)	64 (20.4)	98 (16.1)	0
≥65	37 (16.2)	65 (20.7)	102 (16.7)	0
Ethnicity (%)				
White	146 (63.8)	236 (89.7)	382 (62.6)	0
Black	11 (4.8)	16 (6.1)	27 (4.4)	
Asian	59 (25.8)	55 (20.9)	114 (18.7)	
Mixed	13 (5.7)	7 (2.7)	20 (3.3)	
Physical activity level (%) <sup>a</sup>				
High	155 (64.0)	213 (65.1)	368 (70.8)	5.16
Moderate	54 (22.3)	74 (22.6)	128 (24.6)	
Low	10 (4.1)	9 (2.8)	19 (3.7)	
Smoking (%)				
Never smoker	151 (43.5)	252 (95.8)	403 (74.2)	0
Former smoker	59 (17.0)	54 (20.5)	113 (20.8)	
Current smoker	19 (5.5)	8 (3.0)	27 (5.0)	
BMI ( $\text{kg m}^{-2}$ )	23.7 (2.8)	23.8 (3.8)	23.8 (3.5)	0
Body fat (%)	17.6 (5.6)	30.0 (7.2)	24.8 (9.0)	0
IPAQ (MET per min) <sup>a</sup>	5416.6 (4593.2)	5468.0 (4398.3)	5446.1 (4477.9)	5.16
BMR (kcal $\text{d}^{-1}$ )	1659.8 (147.8)	1286.4 (147.2)	1443.8 (236.2)	0
Alcohol consumption (unit per weeks)	5.4 (6.4)	2.8 (3.6)	3.9 (5.1)	0
Energy intake (kcal $\text{d}^{-1}$ )	1735.3 (532.8)	1574.7 (514.9)	1725.0 (765.1)	3.31
Energy intake/BMR	1.1 (0.3)	1.2 (0.4)	1.2 (0.4)	3.31
Fruits (g $\text{d}^{-1}$ )	234.5 (197.6)	257.1 (196.0)	247.5 (196.8)	3.31
Vegetables (g $\text{d}^{-1}$ ) <sup>b</sup>	260.2 (152.3)	319.2 (327.0)	294.3 (269.0)	3.31
Potatoes (g $\text{d}^{-1}$ )	56.9 (41.4)	41.5 (34.2)	48.0 (38.1)	3.31
Egg intake (g $\text{d}^{-1}$ )	27.5 (29.3)	24.9 (22.2)	26.0 (25.5)	3.31
Fish intake (g $\text{d}^{-1}$ )	44.4 (44.8)	42.5 (38.13)	43.3 (41.1)	3.31
Meat intake (g $\text{d}^{-1}$ )	108.5 (74.51)	75.3 (59.81)	89.7 (68.5)	3.31

<sup>a</sup> Physical activity data available *n* = 520. <sup>b</sup> Not including potatoes.



physical activity, 95.0% of them being non-smokers, and an average alcohol consumption of  $3.9 \pm 5.1$  units per week. The average daily energy intake for women was  $1574.7 \pm 514.9$  kcal, and for men it was  $1735.3 \pm 532.8$  kcal. The daily intake of fruits and vegetables (not including potatoes) of the study population was  $247.5 \pm 196.8$  g and  $294.3 \pm 269.0$  g, respectively. Their daily egg, fish, and meat intake were  $26.0 \pm 25.5$  g,  $43.3 \pm 41.1$  g, and  $89.7 \pm 68.5$  g, respectively. The average fibre intake was  $16.1 \pm 7.0$  g d $^{-1}$ . The average intake of micro and macronutrients is detailed in the ESI, Table 1.†

### Intake of (poly)phenols and (poly)phenol-rich food items

The total (poly)phenol intake of the study population estimated by FFQs was  $1390.2 \pm 957.5$  mg day $^{-1}$ . Of all the food sources of (poly)phenols, the food items involved in the PPS contributed 99.7% of the total dietary (poly)phenols. Among them, tea and coffee contributed 33.7% and 44.2% of the total (poly)phenol intake, respectively. The contribution of these (poly)phenol-rich food items to the total (poly)phenol intake is shown in Table 3.

The consumption of the food items involved in the PPS in the study population is shown in Table 3. The non-alcoholic beverages were the most consumed category, with average tea intake of  $260.5 \pm 289.1$  g day $^{-1}$  and coffee intake of  $185.8 \pm 211.0$  g day $^{-1}$ . Vegetables, fruits, whole grains, and alcoholic beverages ( $253.5 \pm 211.0$  g d $^{-1}$ ,  $172.1 \pm 153.6$  g d $^{-1}$ ,  $86.6 \pm 78.4$  g d $^{-1}$ , and  $26.9 \pm 47.9$  g d $^{-1}$ , respectively) were less consumed food categories than non-alcoholic beverages, while nuts, chocolate and cocoa products, and olive oil showed the lowest intake ( $16.7 \pm 19.4$  g d $^{-1}$ ,  $3.1 \pm 5.0$  g d $^{-1}$ , and  $0.2 \pm 0.4$  g d $^{-1}$ , respectively).

### Distribution of the (poly)phenol-rich diet score

Table 4 shows the calculated PPS of the study population. The PPS ranged from 25 to 88, with a mean score of  $53.7 \pm$

**Table 4** Distribution of the (poly)phenol-rich diet score in the study population

Stratification variables	<i>n</i>	Mean (SD)	<i>t/F</i> value	<i>P</i> value
Overall	543	53.7 (11.7)		
Sex			4.852	<0.001
Men	229	50.9 (11.4)		
Women	314	55.7 (11.5)		
Age			7.752	<0.001
18–34	255	51.9 (12.2) <sup>a</sup>		
35–49	86	52.8 (10.5) <sup>b</sup>		
50–64	99	54.6 (11.2)		
≥65	103	53.7 (11.7) <sup>a, b</sup>		

Tukey HSD test a:  $p < 0.001$ ; b:  $p = 0.007$ .

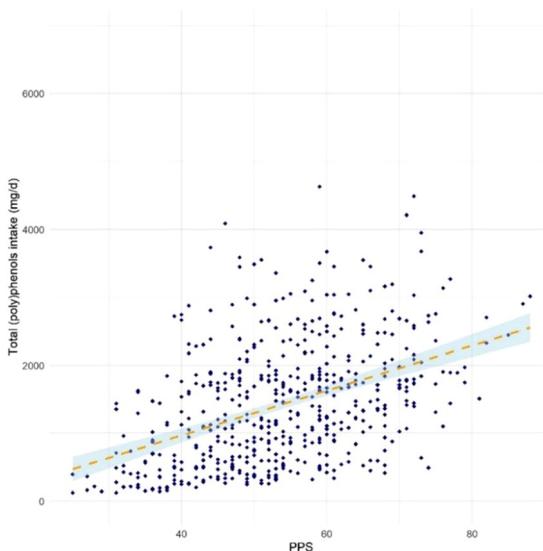
11.7. Women had a higher score than men ( $p < 0.001$ ). Participants in different age groups showed different adherence to a (poly)phenol-rich diet ( $F = 7.752$ ,  $p < 0.001$ ), with participants aged 65 years or older having significantly higher PPS than younger participants aged 18–34 and 35–49 ( $p < 0.001$  and  $p = 0.007$ , respectively). The (poly)phenol-rich diet score was positively correlated with age ( $r = 0.196$ ,  $p < 0.001$ ).

Fig. 1 shows the correlation between dietary total (poly)phenol intake and PPS in the study population. The PPS presented a moderate positive correlation with FFQ estimated total (poly)phenol intake ( $r = 0.43$ , 95% CI (0.36, 0.50),  $p < 0.001$ ). The agreements between PPS and FFQ estimated total (poly)phenol intake in ranking participants into quartiles are shown in Fig. 2. The figure shows that the two methods were comparable in differentiating participants in high and low adherence to the (poly)phenol-rich diet, with 35.4% of participants ranked in the same quartile and only 4.4% ranked in the opposite quartile (the 1st and 4th quartile).

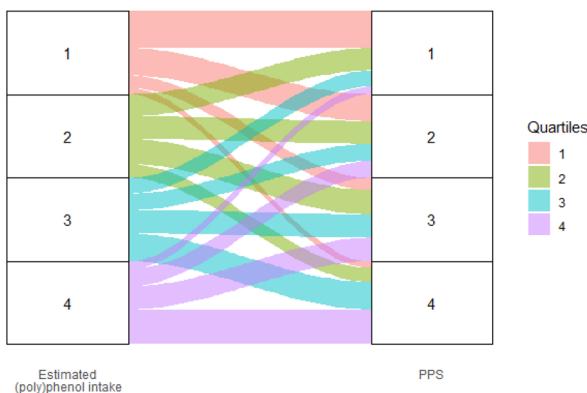
**Table 3** (Poly)phenol-rich food intake of the study population measured by food frequency questionnaires

(Poly)phenol-rich food items involved in PPS	Mean (SD) (g d $^{-1}$ )	Contribution to total (poly)phenol-rich food items (%)	Contribution to total (poly)phenol intake (%)
Tea	260.5 (289.3)	26.0	33.7
Coffee	185.8 (211.2)	18.5	44.2
Red wine	26.9 (47.9)	2.7	0.3
Whole grains	86.6 (78.5)	8.6	1.9
Chocolate and cocoa products	3.1 (5)	0.3	0.7
Berries	6.6 (8.8)	0.7	1.1
Apple and apple juice	66.5 (76.3)	6.6	6.5
Pear	19.2 (34.2)	1.9	0.2
Grape	12.4 (18.8)	1.2	0.8
Plum	2 (4.8)	0.2	0.6
Citrus fruit and juice	65.5 (84.2)	6.5	2.5
Potato and carrots	71.4 (49.2)	7.1	1.4
Onion	19.1 (19.1)	1.9	0.3
Pepper	7.6 (10.8)	0.8	0.0
Garlic	2.6 (2.9)	0.3	0.3
Green vegetables	63.3 (93.5)	6.3	2.9
Pulses	50.7 (59.5)	5.1	0.7
Soy and soy products	38.9 (103.7)	3.9	0.5
Nuts	13.7 (19.4)	1.4	1.3
Olive oil	0.2 (0.4)	0.0	0.0





**Fig. 1** Correlations between PPS and the FFQ estimated total (poly)phenol intake ( $n = 543$ ).



**Fig. 2** Agreements between PPS and the FFQ estimated total (poly)phenol intake in ranking participants into quartiles. Higher level of PPS were significantly associated with higher intake of micronutrients that are related to plant foods, such as potassium (stdBeta (95% CI): 0.70 (0.60, 0.81),  $p < 0.01$ ), magnesium (stdBeta (95% CI): 0.65 (0.55, 0.76),  $p < 0.01$ ), fibre (stdBeta (95% CI): 0.65 (0.57, 0.73),  $p < 0.01$ ), and total folate (stdBeta (95% CI): 0.60 (0.51, 0.68),  $p < 0.01$ ) and were also associated with lower intake of fat (stdBeta (95% CI): -0.65 (-0.57, -0.73),  $p < 0.01$ ), proteins (stdBeta (95% CI): -0.14 (-0.26, -0.02),  $p < 0.01$ ), vitamin D (stdBeta (95% CI): -0.10 (-0.18, -0.02),  $p < 0.01$ ), cholesterol (stdBeta (95% CI): -0.28 (-0.37, -0.19),  $p < 0.01$ ), and SFA (stdBeta (95% CI): -0.34 (-0.46, -0.22),  $p < 0.01$ ), which are related to animal-based diet. The detailed associations between the PPS, the intake of (poly)phenol-rich food items and nutrients are shown in ESI Fig. 2.†

### Associations between the (poly)phenol-rich diet score, (poly)phenol-rich food items and (poly)phenol intake estimated from FFQs

Fig. 3 shows the associations between the PPS, (poly)phenol-rich food items and the different classes and subclasses of dietary (poly)phenol estimated from the FFQs. Significant positive associations were observed in 24 out of 27 classes and

subclasses of (poly)phenol intake with PPS, with stdBeta ranging from 0.12 (0.04, 0.20) to 0.43 (0.36, 0.51). Among them flavonols and lignans showed a moderate association with PPS (stdBeta (95% confidential interval (CI)): 0.43 (0.34, 0.51) and 0.40 (0.32, 0.49), respectively, both  $p < 0.01$ ). Positive associations were seen between most (poly)phenol-rich food items and (poly)phenol intake. Associations were strong for citrus fruit and juice and flavanones (stdBeta (95% CI): 0.99 (0.99, 1.00),  $p < 0.01$ ), red wine and stilbenes and resveratrol (stdBeta (95% CI): 0.97 (0.95, 0.99) and 0.96 (0.93, 0.99), respectively, both  $p < 0.01$ ), tea and hydroxybenzoic acids, flavonoids and flavan-3-ols (stdBeta (95% CI): 0.98 (0.96, 1.00), 0.97 (0.94, 0.99) and 0.98 (0.96, 1.00), respectively, all  $p < 0.01$ ).

### Associations between the (poly)phenol-rich diet score and urinary (poly)phenol metabolites

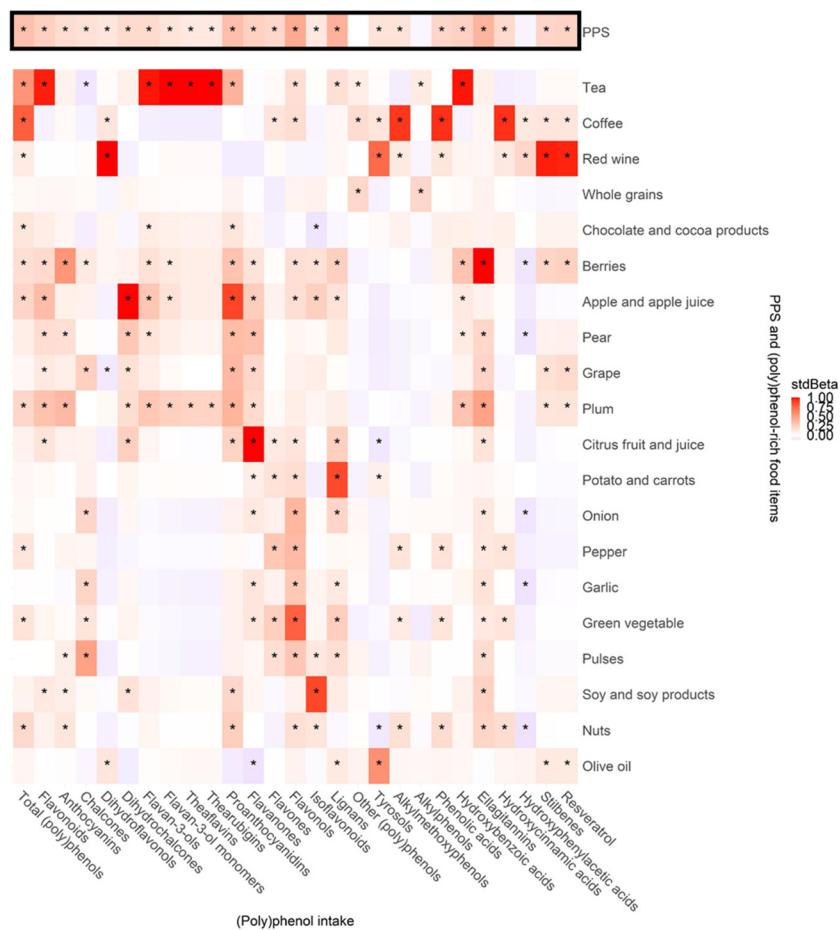
Among the analysed 110 (poly)phenol metabolites in 24 h urine samples, the PPS was significantly associated with 51 metabolites as shown in Table 5. These metabolites were from different classes and subclasses while most of them were phenolic acids ( $n = 39$ ), including 18 cinnamic acids, 13 hydroxybenzoic acids, six phenylpropanoic acids and two hippuric acids. The significant standardized regression coefficients (and 95% CI) between PPS and urinary metabolite levels were all positive, ranging from 0.14 (0.01, 0.29) for 4'-hydroxy-3'-methoxycinnamic acid (*trans*-ferulic acid) to 0.32 (0.18, 0.45) for cinnamic acid-4'-sulfate (*p*-coumaric acid-4'-sulfate), except for the stdBeta for 2,3,4-trihydroxybenzoic acid, which is negative (-0.16 (-0.30, -0.02)). In comparison, FFQ estimated total (poly)phenol intake was significantly associated with the level of 35 (poly)phenol metabolites (Fig. 4), with 28 of these metabolites being phenolic acids.

To explore the sources of (poly)phenol metabolites, the associations between urinary (poly)phenol metabolites and (poly)phenol-rich food items included in PPS are plotted in Fig. 4. Positive associations were observed between most (poly)phenol-rich food items and urinary metabolites. Coffee, the food associated with the greatest number of metabolites, was linked to 25 urinary metabolites, mainly cinnamic acids and phenylpropanoic acids (all  $p < 0.05$ ).

## Discussion

To our knowledge, PPS is the first dietary quality score that reflects adherence to a (poly)phenol-rich diet, which was characterised based on the habitual intake of 20 (poly)phenol-rich plant food items. This novel score was compared against the levels of (poly)phenol metabolites excreted over 24 h as objective measurements. The PPS was not only positively associated with the estimated intake of most of the classes and subclasses of (poly)phenols derived from FFQ but also with multiple urinary (poly)phenol metabolites, which suggested that PPS could reflect (poly)phenol intake or exposure levels





**Fig. 3** Association between PPS, (poly)phenol-rich food items and estimated (poly)phenol intake from FFQs. The heatmap was plotted according to the standardized regression coefficients (stdBeta). The colour scale indicates the effect (stdBeta) of each subclass of (poly)phenol intake on PPS or (poly)phenol-rich food item intake. Red and blue illustrate positive and negative effects, and colour intensity represents the degree of effect. The asterisks showed significance (\*: fdr-adjusted  $p < 0.05$ ). PPS, (poly)phenol-rich diet score. The associations were adjusted for energy intake and trial effect.

and identify participants with higher compliance to the diet rich in (poly)phenols in a free-living population.

(Poly)phenols exist in various plant-based foods including fruits, vegetables, tea, coffee, whole grains with high fibre, and cocoa products.<sup>40</sup> The major types of (poly)phenols consumed by the UK population include flavan-3-ols (mainly from tea), flavanones (mainly from citrus fruits), flavonols (mainly from tea, apple, and onions), hydroxycinnamic acids (mainly from fruits, vegetables, and coffee) and anthocyanins (mainly from berry fruits).<sup>41</sup> The PPS covers all the above major food sources of dietary (poly)phenols and the component food items were selected based on the most compelling findings from the surveys investigating (poly)phenol intake in the UK and comprehensive databases on the (poly)phenol content of foods. The analysis of data from the National Diet and Nutrition Survey (NDNS) in the UK showed that non-alcoholic beverages, tea and coffee were the major sources of flavonoids and hydroxycinnamic acids. They were also the main contributors to the total (poly)phenol intake in British adults, along with choco-

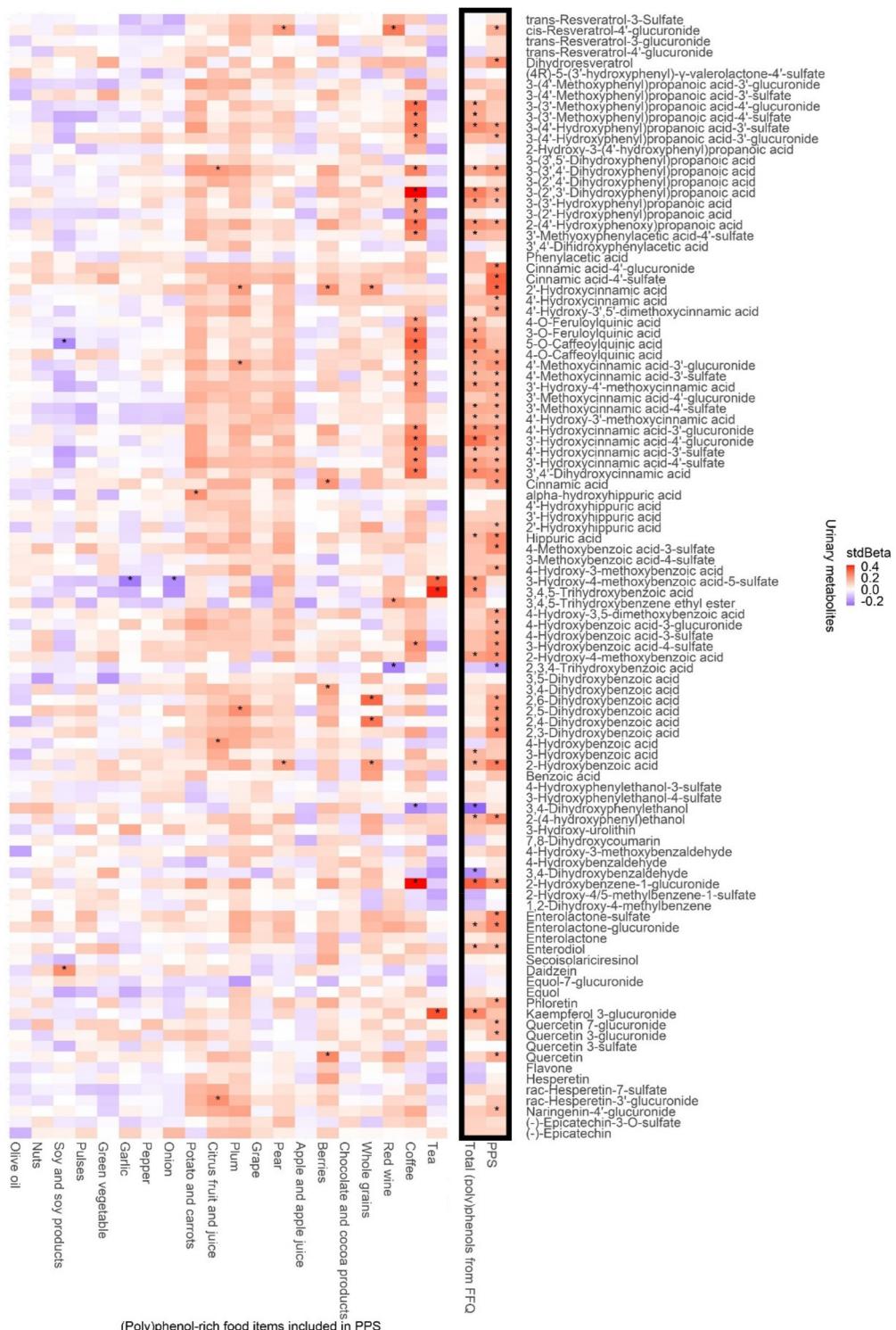
lates, fruits, and fruit juices.<sup>35</sup> In our research, tea and coffee were the highest contributors to total (poly)phenol intake among the selected 20 (poly)phenol-rich food items, which agrees with the NDNS estimated data.<sup>35</sup> Besides, in the previously published data of the same study cohort,<sup>29</sup> isoflavones were mainly obtained from soy and soy products, for instance, tofu and soy milk. Fruits, *i.e.*, oranges, apples, and berries, were reported to be the major food sources of flavanones, proanthocyanidins, and ellagitannins.<sup>29</sup> These food groups were also covered by the PPS. In addition to the major food sources of (poly)phenols from reported data, food items were also evaluated based on their (poly)phenol content calculated using an in-house (poly)phenol database.<sup>29</sup> This database includes data from Phenol-Explorer<sup>42</sup> and USDA<sup>18,19,20</sup> databases, as well as relevant published papers, providing comprehensive information on the (poly)phenol content of foods. It is worth pointing out that the list of (poly)phenol-rich foods included in the PPS was not solely decided on the total (poly)phenol content of foods. Some of the not so (poly)phenol-



Table 5 Associations between PPS and 24 h urinary (poly)phenol metabolites (*n* = 229)

Compound common name	Recommended name	Class	Subclass	stdBeta (95% CI)	P value
Naringenin-4'-glucuronide	Naringenin-4'-glucuronide	Flavonoids	Flavanones	0.15 (0.01, 0.29)	0.046
Quercetin-3-glucuronide	Quercetin 3-glucuronide	Flavonoids	Flavonols	0.17 (0.03, 0.31)	0.03
Quercetin-7-glucuronide	Quercetin 7-glucuronide	Flavonoids	Flavonols	0.16 (0.02, 0.30)	0.03
Quercetin	Quercetin	Flavonoids	Flavonols	0.19 (0.05, 0.33)	0.02
Phloretin	Phloretin	Flavonoids	Flavonols	0.20 (0.06, 0.34)	0.01
2-Hydroxybenzoic acid	2-Hydroxybenzoic acid	Phenolic acids	Hydroxybenzoic acids	0.27 (0.14, 0.41)	<0.01
2,3-Dihydroxybenzoic acid	2,3-Dihydroxybenzoic acid	Phenolic acids	Hydroxybenzoic acids	0.24 (0.10, 0.38)	<0.01
2,4-Dihydroxybenzoic acid	2,4-Dihydroxybenzoic acid	Phenolic acids	Hydroxybenzoic acids	0.21 (0.07, 0.35)	0.01
2,5-Dihydroxybenzoic acid	2,5-Dihydroxybenzoic acid	Phenolic acids	Hydroxybenzoic acids	0.22 (0.08, 0.36)	0.01
2,6-Dihydroxybenzoic acid	2,6-Dihydroxybenzoic acid	Phenolic acids	Hydroxybenzoic acids	0.21 (0.07, 0.34)	0.01
2,3,4-Trihydroxybenzoic acid	2,3,4-Trihydroxybenzoic acid	Phenolic acids	Hydroxybenzoic acids	-0.16 (-0.30, -0.02)	0.03
2-Hydroxy-4-methoxybenzoic acid	2-Hydroxy-4-methoxybenzoic acid	Phenolic acids	Hydroxybenzoic acids	0.23 (0.09, 0.37)	<0.01
Protocatechuic acid-4-sulfate	3-Hydroxybenzoic acid-4-sulfate	Phenolic acids	Hydroxybenzoic acids	0.22 (0.08, 0.36)	0.01
Protocatechuic acid-3-sulfate	4-Hydroxybenzoic acid-3-sulfate	Phenolic acids	Hydroxybenzoic acids	0.20 (0.06, 0.34)	0.01
Protocatechuic acid-3-glucuronide	4-Hydroxybenzoic acid-3-glucuronide	Phenolic acids	Hydroxybenzoic acids	0.17 (0.03, 0.31)	0.03
Syringic acid	4-Hydroxy-3,5-dimethoxybenzoic acid	Phenolic acids	Hydroxybenzoic acids	0.20 (0.06, 0.34)	0.01
Vanillic acid	4-Hydroxy-3-methoxybenzoic acid	Phenolic acids	Hydroxybenzoic acids	0.19 (0.05, 0.33)	0.01
Isovanillic acid-3-sulfate	4-Methoxybenzoic acid-3-sulfate	Phenolic acids	Hydroxybenzoic acids	0.24 (0.10, 0.38)	<0.01
Hippuric acid	Hippuric acid	Phenolic acids	Hippuric acids	0.25 (0.11, 0.38)	<0.01
2'-Hydroxyhippuric acid	2'-Hydroxyhippuric acid	Phenolic acids	Hippuric acids	0.15 (0.01, 0.29)	0.04
Cinnamic acid	Cinnamic acid	Phenolic acids	Cinnamic acids	0.22 (0.08, 0.36)	0.01
Caffeic acid	3',4'-Dihydroxycinnamic acid	Phenolic acids	Cinnamic acids	0.23 (0.09, 0.37)	<0.01
Caffeic acid-4'-sulfate	3'-Hydroxycinnamic acid-4'-sulfate	Phenolic acids	Cinnamic acids	0.18 (0.04, 0.32)	0.02
Caffeic acid-3'-sulfate	4'-Hydroxycinnamic acid-3'-sulfate	Phenolic acids	Cinnamic acids	0.15 (0.01, 0.29)	0.046
Caffeic acid-4'-glucuronide	3'-Hydroxycinnamic acid-4'-glucuronide	Phenolic acids	Cinnamic acids	0.20 (0.06, 0.34)	0.01
Caffeic acid-3'-glucuronide	4'-Hydroxycinnamic acid-3'-glucuronide	Phenolic acids	Cinnamic acids	0.23 (0.09, 0.37)	0.01
trans-Ferulic acid	4'-Hydroxy-3'-methoxycinnamic acid	Phenolic acids	Cinnamic acids	0.14 (0.01, 0.29)	0.049
Ferulic acid-4'-sulfate	3'-Methoxycinnamic acid-4'-sulfate	Phenolic acids	Cinnamic acids	0.18 (0.04, 0.32)	0.02
Ferulic acid-4'-glucuronide	3'-Methoxycinnamic acid-4'-glucuronide	Phenolic acids	Cinnamic acids	0.16 (0.02, 0.30)	0.04
Isoferulic acid	3'-Hydroxy-4'-methoxycinnamic acid	Phenolic acids	Cinnamic acids	0.18 (0.04, 0.32)	0.02
Isoferulic acid-3'-sulfate	4'-Methoxycinnamic acid-3'-sulfate	Phenolic acids	Cinnamic acids	0.16 (0.02, 0.30)	0.04
Isoferulic acid-3'-glucuronide	4'-Methoxycinnamic acid-3'-glucuronide	Phenolic acids	Cinnamic acids	0.23 (0.09, 0.37)	0.01
Cryptochlorogenic acid	4-O-Caffeoylquinic acid	Phenolic acids	Cinnamic acids	0.19 (0.05, 0.33)	0.01
Sinapic acid	4'-Hydroxy-3',5'-dimethoxycinnamic acid	Phenolic acids	Cinnamic acids	0.19 (0.05, 0.33)	0.01
<i>p</i> -Coumaric acid	4'-Hydroxycinnamic acid	Phenolic acids	Cinnamic acids	0.16 (0.02, 0.30)	0.03
<i>p</i> -Coumaric acid-4'-sulfate	Cinnamic acid-4'-sulfate	Phenolic acids	Cinnamic acids	0.32 (0.19, 0.46)	<0.01
<i>p</i> -Coumaric acid-4'-glucuronide	Cinnamic acid-4'-glucuronide	Phenolic acids	Cinnamic acids	0.28 (0.14, 0.41)	<0.01
<i>o</i> -Coumaric acid	2'-Hydroxycinnamic acid	Phenolic acids	Cinnamic acids	0.32 (0.18, 0.45)	<0.01
2-(4'-Hydroxyphenoxy)propanoic acid	2-(4'-Hydroxyphenoxy)propanoic acid	Phenolic acids	Phenylpropanoic acids	0.21 (0.07, 0.35)	0.01
3-(3'-Hydroxyphenyl)propanoic acid	3-(3'-Hydroxyphenyl)propanoic acid	Phenolic acids	Phenylpropanoic acids	0.15 (0.01, 0.30)	0.04
3-(2',3'-Dihydroxyphenyl)propanoic acid	3-(2',3'-Dihydroxyphenyl)propanoic acid	Phenolic acids	Phenylpropanoic acids	0.19 (0.05, 0.33)	0.02
Dihydrocaffeic acid	3-(3',4'-Dihydroxyphenyl)propanoic acid	Phenolic acids	Phenylpropanoic acids	0.22 (0.09, 0.36)	0.01
Dihydrocaffeic acid-3'-sulfate	3-(4'-Hydroxyphenyl)propanoic acid-3'-sulfate	Phenolic acids	Phenylpropanoic acids	0.21 (0.07, 0.35)	0.01
Dihydrocaffeic acid-3'-glucuronide	3-(4'-Hydroxyphenyl)propanoic acid-3'-glucuronide	Phenolic acids	Phenylpropanoic acids	0.17 (0.03, 0.31)	0.03
Enterodiol	Enterodiol	Lignans	Lignans	0.16 (0.02, 0.30)	0.03
Enterolactone-glucuronide	Enterolactone-glucuronide	Lignans	Lignans	0.25 (0.12, 0.39)	<0.01
Enterolactone-sulfate	Enterolactone-sulfate	Lignans	Lignans	0.25 (0.11, 0.39)	<0.01
Dihydroresveratrol	Dihydroresveratrol	Stilbenes	Resveratrol	0.24 (0.10, 0.38)	<0.01
cis-Resveratrol-4'-glucuronide	cis-Resveratrol-4'-glucuronide	Stilbenes	Resveratrol	0.18 (0.04, 0.32)	0.02
Catechol-1-glucuronide	2-Hydroxybenzene-1-glucuronide	Other (poly)phenols	Benzene diols and triols	0.24 (0.10, 0.38)	<0.01
Tyrosol	2-(4-Hydroxyphenyl)ethanol	Other (poly)phenols	Tyrosols	0.24 (0.10, 0.38)	<0.01





**Fig. 4** Association between (poly)phenol-rich food items and 24 h urinary (poly)phenol metabolite excretion. The heatmap was plotted according to the standardized regression coefficients (stdBeta). The colour scale indicates the effect (stdBeta) of each (poly)phenol-rich food item/total (poly)phenol intake/PPS on (poly)phenol metabolite excretion in 24 h urine. Red and blue illustrate positive and negative effects, and colour intensity represents the degree of effect. The asterisks show significance (\*: fdr-adjusted  $p < 0.05$ ). The 24 h urine metabolite excretion was adjusted for batch effect using the Combat method before entering the linear model. The energy intake was adjusted in the model.

dense foods, such as potatoes, were also included in the score since their intake is high in the UK population and therefore had a considerable contribution to the total (poly)phenol intake.<sup>35</sup> The exploration of published studies and databases enabled us to select 20 food groups and the specific food items were matched with the validated EPIC FFQ used in this work. These food items were quite general and widely included in different dietary assessment tools. This allows the potential use of the PPS across different cohorts and different dietary assessment tools.

PPS was positively associated with total dietary (poly)phenol intake estimated from FFQs and showed fair agreement in ranking participants into quartiles. This result indicated that the PPS might be a powerful tool to identify participants with high and low adherence to a (poly)phenol-rich diet. Compared to traditional dietary assessment using databases, the PPS calculation is much easier and less time-consuming, which makes it suitable for application in large epidemiological studies. The positive associations between PPS and multiple (poly)phenol classes and subclasses with a small range of stdBeta indicated that the association between PPS and total (poly)phenol intake was not driven by a certain type of (poly)phenol. This also suggests that PPS has a balanced representation of different subtypes of (poly)phenols.

Plant-based foods are included in many diet quality scores. Fruits and vegetables are widely included in most healthy dietary scores such as the DASH, MDS, HEI and PDI due to the strong evidence existing on their health benefits.<sup>43</sup> The major difference between the PPS and these scores is that the PPS includes individual fruits and vegetables rather than grouping them together, which adds the weights of fruits and vegetables to the final score. A total of six fruit and five vegetable items are included in the 20 PPS food items, therefore, providing a major contribution to the final score compared to other scores. Tea and coffee, which were included as one group in PDI, were separated in PPS due to their distinct (poly)phenol compositions and both being major sources of (poly)phenol intake in the UK. Separating the different items could provide better understanding of the contribution of these food items to the (poly)phenol-rich diet pattern and allow for more flexibility in studying the specific health effects of different components. Some other (poly)phenol-rich food items were also not included in other scores, such as chocolate and cocoa products.

Due to the tight linkage between PPS and plant foods, PPS showed a strong positive association with the nutrients that are commonly found in plant foods and presented a negative association with nutrients that mainly come from animal sources. The negative associations with animal-sourced food, nutrients, and bioactives is a shared feature for plant-rich dietary patterns like the PDI, (including the healthful plant-based diet index (hPDI) and unhealthful plant-based diet index (uPDI)),<sup>25</sup> and plant-based diets (PBDs) (including a vegan diet, lacto-ovo-vegetarian diet, and fish-vegetarian diet).<sup>44</sup> The PPS does not include animal-sourced items since plant-rich food is the only source of (poly)phenols.

Considering the beneficial effect of some animal source bioactives, for instance, omega-3 fatty acids,<sup>45</sup> the intake of fish, egg, and meat may be required to be included as covariate factors when assessing its effect on health outcomes.

The PPS was designed theoretically rather than empirically as there are currently no gold standards for estimating (poly)phenol intake. This score was designed based on relative intake rather than absolute intake, therefore, it should only be used to rank participants in the same population for (poly)phenol intake and not to compare across different populations. Absolute cut-off values were not applied because currently there is limited evidence to propose an adequate (poly)phenol intake amount from these foods that would exert health benefits.<sup>46</sup> Besides, the relative scoring system guarantees a balanced distribution in the final score, which is beneficial for the analysis relating to health outcomes. If participants were all scored as low or high consumers by the absolute values, the score would not be able to reflect the variance in the intake.

The algorithm used to calculate the PPS follows the same methodology as the DASH dietary score, which is calculated with quintile criteria for each food group, and a score from 1 to 5 represents (poly)phenol intake from the lowest to the highest intake. The final score of PPS ranged from 20 to 100, reflecting the overall ranking of (poly)phenol-rich food consumption of the participants in the study population. It should also be noted that equal weightage to coffee, tea, and many fruits and vegetables was given in the score, even though coffee and tea contributed to nearly 80% of the total (poly)phenol intake. Many other food sources, such as soy and soy products, nuts and seeds, are contributing to subclasses of (poly)phenols other than the major ones such as hydroxycinnamic acids and flavan-3-ols. Therefore, higher weightage was attached to these food items when using PPS to rank the individuals' adherence to (poly)phenol-rich diet than when calculating total (poly)phenol intake. There is still very limited understanding of the differential effects of various types of (poly)phenols on health especially those consumed in lower amounts. Foods are ingested as a complex mixture of different components and if we only focus on the foods that are major sources of (poly)phenols, namely non-alcoholic beverages, we may be underestimating the effect of other subclasses of (poly)phenols when evaluating relationships between (poly)phenol rich diets and health.

In this study, multiple (poly)phenol metabolites in 24 h urine samples were significantly associated with the PPS. The number of metabolites significantly associated with PPS was higher than the number associated with the FFQ estimated total (poly)phenol intake and individual food intake. In addition, the pattern of metabolites associated with estimated total (poly)phenol intake was driven mainly by tea and coffee consumption, as most metabolites associated with tea and coffee intake are also associated with the total (poly)phenol intake. In contrast, the metabolites associated with the PPS cover a wider range of classes and subclasses of (poly)phenols including cinnamic acids, hydroxybenzoic acids, phenylacetic



acids, and hippuric acids together with lignans, flavonoids, tyrosols, benzenes, and resveratrol, which suggested that the PPS is a good indicator of a (poly)phenol-rich dietary pattern with multiple food sources of (poly)phenols. The observed associations between metabolites and (poly)phenol-rich food intake align with the compositional profiles of the respective foods. For example, tea consumption correlated with kaempferol-3-glucuronide and gallic acid (3,4,5-trihydroxybenzoic acid), coffee intake correlated with multiple cinnamic acids, red wine with *cis*-resveratrol-4'-glucuronide, soy with daidzein, *etc.* The limited number of associations of many food items with metabolites could be because these food items were less frequently consumed and were less likely to be captured in the 24 h urine test. This suggests that PPS may have advantages over estimated (poly)phenol intake in reflecting the ingested amount and pattern of (poly)phenols. Indeed, our previous research in the same study population found that the FFQ-estimated total (poly)phenol intake was poorly correlated to the total and individual subclasses of urinary (poly)phenol metabolites.<sup>29</sup> Therefore, being easier to calculate and more closely associated with (poly)phenol exposure levels, PPS could be a better tool than estimated total intake in reflecting adherence to (poly)phenol-rich diets in epidemiological studies.

Compared to the traditional widely used dietary assessment methods, biomarkers are a more objective approach to reflect exposure levels because they could prevent the errors derived from misreporting. However, to date there are still very few (poly)phenol metabolites that have been validated to predict intake levels of certain (poly)phenols.<sup>47–49</sup> Many low molecular weight phenolic metabolites with a high abundance in urine and plasma such as phenolic acids could come from both dietary and non-dietary sources,<sup>50,51</sup> or endogenous metabolism.<sup>52</sup> Apart from that, the high inter-individual variability in the metabolism of (poly)phenols could also hinder the validity of a biomarker because of the inconsistent dose-response in the general population. However, multiple (poly)phenol metabolites were found to have positive associations with (poly)phenol and (poly)phenol-rich food intake in previous studies.<sup>53,54</sup> In addition, despite the high inter-individual variability in gut microbiome composition and metabolism abilities in free-living populations, many of the metabolites derived from gut microbial metabolism were significantly associated with PPS.

The current study has several limitations. Firstly, although being widely applied in UK studies,<sup>55–58</sup> the EPIC-Norfolk FFQ was not designed or validated to assess (poly)phenol intake but nutrients and food groups.<sup>30</sup> Several food sources of (poly)phenols were not included in the questionnaire, such as blueberries, many common spices, and some nuts and seeds. Additionally, some food items with largely different (poly)phenol content were not distinguished in the questions. For instance, dark and milk chocolate, red and white wine, and black and green tea. The above imprecision factors might influence the accuracy of PPS to estimate (poly)phenol intake in our study population and underestimate the effect of high PPS on health outcomes.<sup>59</sup> Due to the various dietary patterns

in different countries, the major contributors to dietary (poly)phenol intake could vary between populations.<sup>60,61</sup> Research on a universal PPS version based on the local catering culture of various continents is still required. Besides, the PPS was scored according to the relative intake distribution in the study population, so the relationships between the score and health could not be pooled and compared across studies. However, this is a common limitation for other diet scores such as DASH and PDI. Furthermore, different (poly)phenols may have different health effects while the PPS includes all types of (poly)phenols. The effects of a (poly)phenol-rich diet on health might be modulated by individual (poly)phenol subclasses, which need to be taken into account when interpreting the relationship between the PPS and health. Future studies could develop diet scores specifically focused on some (poly)phenol groups.

In conclusion, the PPS provides a novel way of ranking participants based on (poly)phenol-rich food intake obtained from validated FFQs to estimate adherence to (poly)phenol-rich diets. The tight linkage between PPS and nutrients, (poly)phenol intake, and urinary metabolites also reflects its potential capacity of holistically characterizing a (poly)phenol-rich diet quality. Future studies are required to evaluate the link between high adherence to (poly)phenol-rich diets using the PPS and cardiometabolic health.

## Author contributions

Conceptualisation: Y. X., Y. L., A. R. M., and R. G.; data curation: Y. X., Y. L., and J. H.; data analysis and writing: Y. X. and Y. L.; methodology development: Y. X., Y. L., A. R. M. and R. G.; editing manuscript and supervision, A. R. M. and R. G.; and project administration and funding: A. R. M.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

We extend our highest gratitude to the participants of all clinical studies. We also thank Anthony Sullivan, Chris Titman, and Neil Loftus from Shimadzu UK Ltd for their support with the LC-MS/MS instrument. Y. L. and Y. X. are funded by the King's-China Scholarship Council (K-CSC) joint scholarship (ID: 10.13039/501100004543).

## References

- 1 A. Afshin, P. J. Sur, K. A. Fay, L. Cornaby, G. Ferrara, J. S. Salama, E. C. Mullany, K. H. Abate, C. Abbafati and Z. Abebe, *Health effects of dietary risks in 195 countries*,

1990–2017: a systematic analysis for the Global Burden of Disease Study 2017, *Lancet*, 2019, **393**, 1958–1972.

2 G. B. D. D. Collaborators, Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017, *Lancet*, 2019, **393**, 1958–1972.

3 D. D. Wang, Y. Li, S. N. Bhupathiraju, B. A. Rosner, Q. Sun, E. L. Giovannucci, E. B. Rimm, J. E. Manson, W. C. Willett, M. J. Stampfer and F. B. Hu, Fruit and Vegetable Intake and Mortality: Results From 2 Prospective Cohort Studies of US Men and Women and a Meta-Analysis of 26 Cohort Studies, *Circulation*, 2021, **143**, 1642–1654.

4 P. B. Mellen, T. F. Walsh and D. M. Herrington, Whole grain intake and cardiovascular disease: a meta-analysis, *Nutr. Metab. Cardiovasc. Dis.*, 2008, **18**, 283–290.

5 J. S. de Munter, F. B. Hu, D. Spiegelman, M. Franz and R. M. van Dam, Whole grain, bran, and germ intake and risk of type 2 diabetes: a prospective cohort study and systematic review, *PLoS Med.*, 2007, **4**, e261.

6 Q. Sun, D. Spiegelman, R. M. van Dam, M. D. Holmes, V. S. Malik, W. C. Willett and F. B. Hu, White rice, brown rice, and risk of type 2 diabetes in US men and women, *Arch. Intern. Med.*, 2010, **170**, 961–969.

7 G. Grosso, J. Yang, S. Marventano, A. Micek, F. Galvano and S. N. Kales, Nut consumption on all-cause, cardiovascular, and cancer mortality risk: a systematic review and meta-analysis of epidemiologic studies, *Am. J. Clin. Nutr.*, 2015, **101**, 783–793.

8 G. M. A. L. Barabási and J. Loscalzo, The unmapped chemical complexity of our diet, *Nat. Food*, 2020, **1**, 33–37.

9 L. Hooper, C. Kay, A. Abdelhamid, P. A. Kroon, J. S. Cohn, E. B. Rimm and A. Cassidy, Effects of chocolate, cocoa, and flavan-3-ols on cardiovascular health: a systematic review and meta-analysis of randomized trials, *Am. J. Clin. Nutr.*, 2012, **95**, 740–751.

10 F. Poti, D. Santi, G. Spaggiari, F. Zimetti and I. Zanotti, Polyphenol Health Effects on Cardiovascular and Neurodegenerative Disorders: A Review and Meta-Analysis, *Int. J. Mol. Sci.*, 2019, **20**, 351–377.

11 B. H. Parmenter, K. D. Croft, J. M. Hodgson, F. Dalgaard, C. P. Bondonno, J. R. Lewis, A. Cassidy, A. Scalbert and N. P. Bondonno, An overview and update on the epidemiology of flavonoid intake and cardiovascular disease risk, *Food Funct.*, 2020, **11**, 6777–6806.

12 G. Grosso, A. Micek, J. Godos, A. Pajak, S. Sciacca, F. Galvano and E. L. Giovannucci, Dietary Flavonoid and Lignan Intake and Mortality in Prospective Cohort Studies: Systematic Review and Dose-Response Meta-Analysis, *Am. J. Epidemiol.*, 2017, **185**, 1304–1316.

13 A. Cassidy, E. J. O'Reilly, C. Kay, L. Sampson, M. Franz, J. P. Forman, G. Curhan and E. B. Rimm, Habitual intake of flavonoid subclasses and incident hypertension in adults, *Am. J. Clin. Nutr.*, 2011, **93**, 338–347.

14 R. Zamora-Ros, D. K. Barupal, J. A. Rothwell, M. Jenab, V. Fedirko, I. Romieu, K. Aleksandrova, K. Overvad, C. Kyro, A. Tjonneland, A. Affret, M. His, M. C. Boutron-Ruault, V. Katzke, T. Kuhn, H. Boeing, A. Trichopoulou, A. Naska, M. Kritikou, C. Saieva, C. Agnoli, M. Santucci de Magistris, R. Tumino, F. Fasanelli, E. Weiderpass, G. Skeie, S. Merino, P. Jakobsen, M. J. Sanchez, M. Dorronsoro, C. Navarro, E. Ardanaz, E. Sonestedt, U. Ericson, L. Maria Nilsson, S. Boden, H. B. Bueno-de-Mesquita, P. H. Peeters, A. Perez-Cornago, N. J. Wareham, K. T. Khaw, H. Freisling, A. J. Cross, E. Riboli and A. Scalbert, Dietary flavonoid intake and colorectal cancer risk in the European prospective investigation into cancer and nutrition (EPIC) cohort, *Int. J. Cancer*, 2017, **140**, 1836–1844.

15 R. Zamora-Ros, N. G. Forouhi, S. J. Sharp, C. A. Gonzalez, B. Buijsse, M. Guevara, Y. T. van der Schouw, P. Amiano, H. Boeing, L. Bredsdorff, F. Clavel-Chapelon, G. Fagherazzi, E. J. Feskens, P. W. Franks, S. Grioni, V. Katzke, T. J. Key, K. T. Khaw, T. Kuhn, G. Masala, A. Mattiello, E. Molina-Montes, P. M. Nilsson, K. Overvad, F. Perquier, J. R. Quiros, I. Romieu, C. Sacerdote, A. Scalbert, M. Schulze, N. Slimani, A. M. Spijkerman, A. Tjonneland, M. J. Tormo, R. Tumino, A. D. van der, C. Langenberg, E. Riboli and N. J. Wareham, The association between dietary flavonoid and lignan intakes and incident type 2 diabetes in European populations: the EPIC-InterAct study, *Diabetes Care*, 2013, **36**, 3961–3970.

16 Y. Xu, M. Le Sayec, C. Roberts, S. Hein, A. Rodriguez-Mateos and R. Gibson, Dietary Assessment Methods to Estimate (Poly)phenol Intake in Epidemiological Studies: A Systematic Review, *Adv. Nutr.*, 2021, **12**, 1781–1801.

17 J. A. Rothwell, J. Perez-Jimenez, V. Neveu, A. Medina-Remón, N. M'Hiri, P. Garcia-Lobato, C. Manach, C. Knox, R. Eisner, D. S. Wishart and A. Scalbert, Phenol-Explorer 3.0: a major update of the Phenol-Explorer database to incorporate data on the effects of food processing on polyphenol content, *Database*, 2013, **2013**, bat070.

18 S. Bhagwat and D. B. Haytowitz, USDA Database for the Isoflavone Content of Selected Foods, 2015, DOI: [10.15482/USDA.ADC/1324538](https://doi.org/10.15482/USDA.ADC/1324538).

19 S. Bhagwat and D. B. Haytowitz, USDA Database for the Proanthocyanidin Content of Selected Foods, 2015, DOI: [10.15482/USDA.ADC/1324621](https://doi.org/10.15482/USDA.ADC/1324621).

20 S. Bhagwat and D. B. Haytowitz, USDA Database for the Flavonoid Content of Selected Foods, 2016, DOI: [10.15482/USDA.ADC/1324465](https://doi.org/10.15482/USDA.ADC/1324465).

21 P. M. Waijers, E. J. Feskens and M. C. Ocke, A critical review of predefined diet quality scores, *Br. J. Nutr.*, 2007, **97**, 219–231.

22 N. C. f. H. Statistics, *Healthy eating index-2015, scores—US Department of Agriculture*, Center for Nutrition Policy and Promotion, 2013.

23 A. Trichopoulou, T. Costacou, C. Bamia and D. Trichopoulos, Adherence to a Mediterranean diet and survival in a Greek population, *N. Engl. J. Med.*, 2003, **348**, 2599–2608.

24 T. T. Fung, S. E. Chiue, M. L. McCullough, K. M. Rexrode, G. Logroscino and F. B. Hu, Adherence to a DASH-style diet and risk of coronary heart disease and stroke in women, *Arch. Intern. Med.*, 2008, **168**, 713–720.



25 A. Satija, S. N. Bhupathiraju, D. Spiegelman, S. E. Chiuve, J. E. Manson, W. Willett, K. M. Rexrode, E. B. Rimm and F. B. Hu, Healthful and Unhealthful Plant-Based Diets and the Risk of Coronary Heart Disease in U.S. Adults, *J. Am. Coll. Cardiol.*, 2017, **70**, 411–422.

26 G. S. Aljuraiban, R. Gibson, L. M. Oude Griep, N. Okuda, L. M. Steffen, L. Van Horn and Q. Chan, Perspective: The Application of A Priori Diet Quality Scores to Cardiovascular Disease Risk-A Critical Evaluation of Current Scoring Systems, *Adv. Nutr.*, 2020, **11**, 10–24.

27 Y. Li, Y. Xu, X. Ma, M. Le Sayec, H. Wu, P. Dazzan, C. Nosarti, C. Heiss, R. Gibson and A. Rodriguez-Mateos, (Poly)phenol intake, plant-rich dietary patterns and cardiometabolic health: a cross-sectional study, *Food Funct.*, 2023, **14**(9), 4078–4091.

28 A. V. D. Rodriguez-Mateos, C. G. Krueger, D. Shanmuganayagam, J. Reed, L. Calani, P. Mena, D. Del Rio and A. Crozier, Bioavailability, bioactivity and impact on health of dietary flavonoids and related compounds: an update, *Arch. Toxicol.*, 2014, **88**, 1803–1853.

29 Y. Xu, Y. Li, X. Ma, W. Alotaibi, M. Le Sayec, A. Cheok, E. Wood, S. Hein, P. Young Tie Yang, W. L. Hall, C. Nosarti, P. Dazzan, R. Gibson and A. Rodriguez-Mateos, Comparison between dietary assessment methods and biomarkers in estimating dietary (poly)phenol intake, *Food Funct.*, 2023, **14**(3), 1369–1386.

30 S. A. Bingham, A. A. Welch, A. McTaggart, A. A. Mulligan, S. A. Runswick, R. Luben, S. Oakes, K. T. Khaw, N. Wareham and N. E. Day, Nutritional methods in the European Prospective Investigation of Cancer in Norfolk, *Public Health Nutr.*, 2001, **4**, 847–858.

31 A. A. Mulligan, R. N. Luben, A. Bhaniani, D. J. Parry-Smith, L. O'Connor, A. P. Khawaja, N. G. Forouhi, K. T. Khaw and E. P.-N. F. Study, A new tool for converting food frequency questionnaire data into nutrient and food group values: FETA research methods and availability, *BMJ Open*, 2014, **4**, e004503.

32 R. A. McCance, E. M. Widdowson and B. Holland, *McCance and Widdowson's The composition of foods*, Royal Society of Chemistry, Cambridge, 5th revised and extended edn, 1991.

33 R. A. McCance and E. M. Widdowson, *McCance and Widdowson's The composition of foods*, Royal Society of Chemistry, Cambridge, 6th summary edn, 2002.

34 J. Perez-Jimenez, V. Neveu, F. Vos and A. Scalbert, Identification of the 100 richest dietary sources of polyphenols: an application of the Phenol-Explorer database, *Eur. J. Clin. Nutr.*, 2010, **64**(Suppl 3), S112–S120.

35 N. Ziauddeen, A. Rosi, D. Del Rio, B. Amoutzopoulos, S. Nicholson, P. Page, F. Scazzina, F. Brighenti, S. Ray and P. Mena, Dietary intake of (poly)phenols in children and adults: cross-sectional analysis of UK National Diet and Nutrition Survey Rolling Programme (2008–2014), *Eur. J. Nutr.*, 2019, **58**, 3183–3198.

36 M. Dominguez-Fernandez, Y. Xu, P. Young Tie Yang, W. Alotaibi, R. Gibson, W. L. Hall, L. Barron, I. A. Ludwig, C. Cid and A. Rodriguez-Mateos, Quantitative Assessment of Dietary (Poly)phenol Intake: A High-Throughput Targeted Metabolomics Method for Blood and Urine Samples, *J. Agric. Food Chem.*, 2021, **69**, 537–554.

37 R Core Team, R: A language and environment for statistical computing, <https://www.R-project.org/>.

38 W. E. Johnson, C. Li and A. Rabinovic, Adjusting batch effects in microarray expression data using empirical Bayes methods, *Biostatistics*, 2007, **8**, 118–127.

39 W. Han and L. Li, Evaluating and minimizing batch effects in metabolomics, *Mass Spectrom. Rev.*, 2022, **41**, 421–442.

40 J. A. Rothwell, V. Knaze and R. Zamora-Ros, Polyphenols: dietary assessment and role in the prevention of cancers, *Curr. Opin. Clin. Nutr. Metab. Care*, 2017, **20**, 512–521.

41 G. Williamson, The role of polyphenols in modern nutrition, *Nutr. Bull.*, 2017, **42**, 226–235.

42 Phenol-Explorer database, *Journal*.

43 G. Aljuraiban, Q. Chan, R. Gibson, J. Stamler, M. L. Daviglus, A. R. Dyer, K. Miura, Y. Wu, H. Ueshima, L. Zhao, L. Van Horn, P. Elliott, L. M. Oude Griep and I. R. Group, Association between plant-based diets and blood pressure in the INTERMAP study, *BMJ Nutr. Prev. Health*, 2020, **3**, 133–142.

44 G. Marrone, C. Guerrero, D. Palazzetti, P. Lido, A. Marolla, F. Di Daniele and A. Noce, Vegan Diet Health Benefits in Metabolic Syndrome, *Nutrients*, 2021, **13**, 817–840.

45 A. A. Sangouni, Z. Orang and H. Mozaffari-Khosravi, Effect of omega-3 supplementation on cardiometabolic indices in diabetic patients with non-alcoholic fatty liver disease: a randomized controlled trial, *BMC Nutr.*, 2021, **7**, 86.

46 C. Del Bò, S. Bernardi, M. Marino, M. Porrini, M. Tucci, S. Guglielmetti, A. Cherubini, B. Carrieri, B. Kirkup, P. Kroon, R. Zamora-Ros, N. H. Liberona, C. Andres-Lacueva and P. Riso, Systematic Review on Polyphenol Intake and Health Outcomes: Is there Sufficient Evidence to Define a Health-Promoting Polyphenol-Rich Dietary Pattern?, *Nutrients*, 2019, **11**, 1355–1410.

47 J. I. Ottaviani, R. Fong, J. Kimball, J. L. Ensunsa, A. Britten, D. Lucarelli, R. Luben, P. B. Grace, D. H. Mawson, A. Tym, A. Wierzbicki, K. T. Khaw, H. Schroeter and G. G. C. Kuhnle, Evaluation at scale of microbiome-derived metabolites as biomarker of flavan-3-ol intake in epidemiological studies, *Sci. Rep.*, 2018, **8**, 9859.

48 J. I. Ottaviani, R. Fong, J. Kimball, J. L. Ensunsa, N. Gray, A. Vogiatzoglou, A. Britten, D. Lucarelli, R. Luben, P. B. Grace, D. H. Mawson, A. Tym, A. Wierzbicki, A. D. Smith, N. J. Wareham, N. G. Forouhi, K. T. Khaw, H. Schroeter and G. G. C. Kuhnle, Evaluation of (-)-epicatechin metabolites as recovery biomarker of dietary flavan-3-ol intake, *Sci. Rep.*, 2019, **9**, 13108.

49 M. R. Ritchie, M. S. Morton, A. M. Thompson, N. Deighton, A. Blake, J. H. Cummings and C. M. Steel, Investigation of the reliability of 24 h urine excretion as a biomarker of isoflavone exposure over time and over a wide range of isoflavone intakes, *Eur. J. Clin. Nutr.*, 2004, **58**, 1286–1289.



50 M. Bojic, C. A. Sedgeman, L. D. Nagy and F. P. Guengerich, Aromatic hydroxylation of salicylic acid and aspirin by human cytochromes P450, *Eur. J. Pharm. Sci.*, 2015, **73**, 49–56.

51 C. Manach, A. Scalbert, C. Morand, C. Remesy and L. Jimenez, Polyphenols: food sources and bioavailability, *Am. J. Clin. Nutr.*, 2004, **79**, 727–747.

52 G. Ebinger, Y. Michotte and P. Herregodts, The significance of homovanillic acid and 3,4-dihydroxyphenylacetic acid concentrations in human lumbar cerebrospinal fluid, *J. Neurochem.*, 1987, **48**, 1725–1729.

53 W. M. Edmands, P. Ferrari, J. A. Rothwell, S. Rinaldi, N. Slimani, D. K. Barupal, C. Biessy, M. Jenab, F. Clavel-Chapelon, G. Fagherazzi, M. C. Boutron-Ruault, V. A. Katzke, T. Kuhn, H. Boeing, A. Trichopoulou, P. Lagiou, D. Trichopoulos, D. Palli, S. Grioni, R. Tumino, P. Vineis, A. Mattiello, I. Romieu and A. Scalbert, Polyphenol metabolome in human urine and its association with intake of polyphenol-rich foods across European countries, *Am. J. Clin. Nutr.*, 2015, **102**, 905–913.

54 H. Noh, H. Freisling, N. Assi, R. Zamora-Ros, D. Achaintre, A. Affret, F. Mancini, M. C. Boutron-Ruault, A. Flogel, H. Boeing, T. Kuhn, R. Schubel, A. Trichopoulou, A. Naska, M. Kritikou, D. Palli, V. Pala, R. Tumino, F. Ricceri, M. Santucci de Magistris, A. Cross, N. Slimani, A. Scalbert and P. Ferrari, Identification of Urinary Polyphenol Metabolite Patterns Associated with Polyphenol-Rich Food Intake in Adults from Four European Countries, *Nutrients*, 2017, **9**, 796–810.

55 P. Jones, J. E. Cade, C. E. L. Evans, N. Hancock and D. C. Greenwood, Does adherence to the World Cancer Research Fund/American Institute of Cancer Research cancer prevention guidelines reduce risk of colorectal cancer in the UK Women's Cohort Study?, *Br. J. Nutr.*, 2018, **119**, 340–348.

56 G. Lietz, K. L. Barton, P. J. Longbottom and A. S. Anderson, Can the EPIC food-frequency questionnaire be used in adolescent populations?, *Public Health Nutr.*, 2002, **5**, 783–789.

57 K. Papier, T. Y. Tong, P. N. Appleby, K. E. Bradbury, G. K. Fensom, A. Knappel, A. Perez-Cornago, J. A. Schmidt, R. C. Travis and T. J. Key, Comparison of Major Protein-Source Foods and Other Food Groups in Meat-Eaters and Non-Meat-Eaters in the EPIC-Oxford Cohort, *Nutrients*, 2019, **11**, 824–840.

58 R. Zamora-Ros, V. Cayssials, M. Jenab, J. A. Rothwell, V. Fedirko, K. Aleksandrova, A. Tjonneland, C. Kyro, K. Overvad, M. C. Boutron-Ruault, F. Carbonnel, Y. Mahamat-Saleh, R. Kaaks, T. Kuhn, H. Boeing, A. Trichopoulou, E. Valanou, E. Vasilopoulou, G. Masala, V. Pala, S. Panico, R. Tumino, F. Ricceri, E. Weiderpass, M. Lukic, T. M. Sandanger, C. Lasheras, A. Agudo, M. J. Sanchez, P. Amiano, C. Navarro, E. Ardanaz, E. Sonestedt, B. Ohlsson, L. M. Nilsson, M. Rutegard, B. Bueno-de-Mesquita, P. H. Peeters, K. T. Khaw, N. J. Wareham, K. Bradbury, H. Freisling, I. Romieu, A. J. Cross, P. Vineis and A. Scalbert, Dietary intake of total polyphenol and polyphenol classes and the risk of colorectal cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort, *Eur. J. Epidemiol.*, 2018, **33**, 1063–1075.

59 S. A. Bingham, C. Gill, A. Welch, A. Cassidy, S. A. Runswick, S. Oakes, R. Lubin, D. I. Thurnham, T. J. Key, L. Roe, K. T. Khaw and N. E. Day, Validation of dietary assessment methods in the UK arm of EPIC using weighed records, and 24-hour urinary nitrogen and potassium and serum vitamin C and carotenoids as biomarkers, *Int. J. Epidemiol.*, 1997, **26**(Suppl 1), S137–S151.

60 A. Vogiatzoglou, A. A. Mulligan, M. A. Lentjes, R. N. Luben, J. P. Spencer, H. Schroeter, K. T. Khaw and G. G. Kuhnle, Flavonoid intake in European adults (18 to 64 years), *PLoS One*, 2015, **10**, e0128132.

61 R. Zamora-Ros, V. Knaze, J. A. Rothwell, B. Hemon, A. Moskal, K. Overvad, A. Tjonneland, C. Kyro, G. Fagherazzi, M. C. Boutron-Ruault, M. Touillaud, V. Katzke, T. Kuhn, H. Boeing, J. Forster, A. Trichopoulou, E. Valanou, E. Peppa, D. Palli, C. Agnoli, F. Ricceri, R. Tumino, M. S. de Magistris, P. H. Peeters, H. B. Bueno-de-Mesquita, D. Engeset, G. Skeie, A. Hjartaker, V. Menendez, A. Agudo, E. Molina-Montes, J. M. Huerta, A. Barricarte, P. Amiano, E. Sonestedt, L. M. Nilsson, R. Landberg, T. J. Key, K. T. Khaw, N. J. Wareham, Y. Lu, N. Slimani, I. Romieu, E. Riboli and A. Scalbert, Dietary polyphenol intake in Europe: the European Prospective Investigation into Cancer and Nutrition (EPIC) study, *Eur. J. Nutr.*, 2016, **55**, 1359–1375.

