Urinary metabolic biomarkers of diet quality in European children are associated with metabolic health

Nikos Stratakis[1]*†, Alexandros P Siskos[2]†, Eleni Papadopoulou[3]†, Anh N Nguyen[4], Yinqi Zhao[1], Katerina Margetaki[1], Chung-Ho E Lau[2,5], Muireann Coen[2,6], Lea Maitre[7,8,9], Silvia Fernández-Barrés[7,8,9], Lydiane Agier[10], Sandra Andrusaityte[11], Xavier Basagaña[7,8,9], Anne Lise Brantsaeter[3], Maribel Casas[7,8,9], Serena Fossati[7,8,9], Regina Grazuleviciene[11], Barbara Heude[12], Rosemary RC McEachan[13], Helle Margrete Meltzer[3], Christopher Millett[14,15], Fernanda Rauber[14,15,16], Oliver Robinson[5], Theano Roumeliotaki[17], Eva Borras[7,18], Eduard Sabidó[7,18], Jose Urquiza[7,8,9], Marina Vafeiadi[17], Paolo Vineis[5], Trudy Voortman[4], John Wright[13], David V Conti[1], Martine Vrijheid[7,8,9], Hector C Keun[2‡], Leda Chatzi[1‡]

[1]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, United States; [2]Cancer Metabolism & Systems Toxicology Group, Division of Cancer, Department of Surgery & Cancer and Division of Systems Medicine, Department of Metabolism, Digestion & Reproduction, Imperial College London, Hammersmith Hospital Campus, London, United Kingdom; [3]Norwegian Institute of Public Health, Oslo, Norway; [4]Department of Epidemiology, Erasmus University Medical Center, Rotterdam, Netherlands; [5]MRC Centre for Environment and Health, School of Public Health, Imperial College London, London, United Kingdom; [6]Oncology Safety, Clinical Pharmacology and Safety Sciences, R&D, AstraZeneca, Cambridge, United Kingdom; [7]ISGlobal, Barcelona, Spain; [8]Universitat Pompeu Fabra, Barcelona, Spain; [9]CIBER Epidemiologia y Salud Pública, Madrid, Spain; [10]Inserm, CNRS, University Grenoble Alpes, Team of environmental epidemiology applied to reproduction and respiratory health, IAB, Grenoble, France; [11]Department of Environmental Sciences, Vytautas Magnus University, Kaunas, Lithuania; [12]Centre for Research in Epidemiology and Statistics, Université de Paris, Inserm, Inra, Paris, France; [13]Bradford Institute for Health Research, Bradford Teaching Hospitals NHS Foundation Trust, Bradford, United Kingdom; [14]Public Health Policy Evaluation Unit, School of Public Health, Imperial College, London, United Kingdom; [15]Department of Preventive Medicine, School of Medicine, University of São Paulo, São Paulo, Brazil; [16]Center for Epidemiological Research in Nutrition and Health, University of São Paulo, São Paulo, Brazil; [17]Department of Social Medicine, Faculty of Medicine, University of Crete, Heraklion, Greece; [18]Centre for Genomic Regulation, The Barcelona Institute of Science and Technology, Barcelona, Spain

Abstract Urinary metabolic profiling is a promising powerful tool to reflect dietary intake and can help understand metabolic alterations in response to diet quality. Here, we used ^{1}H NMR spectroscopy in a multicountry study in European children (1147 children from 6 different cohorts) and identified a common panel of 4 urinary metabolites (hippurate, *N*-methylnicotinic acid, urea, and sucrose) that was predictive of Mediterranean diet adherence (KIDMED) and ultra-processed food consumption and also had higher capacity in discriminating children's diet quality than that

of established sociodemographic determinants. Further, we showed that the identified metabolite panel also reflected the associations of these diet quality indicators with C-peptide, a stable and accurate marker of insulin resistance and future risk of metabolic disease. This methodology enables objective assessment of dietary patterns in European child populations, complementary to traditional questionary methods, and can be used in future studies to evaluate diet quality. Moreover, this knowledge can provide mechanistic evidence of common biological pathways that characterize healthy and unhealthy dietary patterns, and diet-related molecular alterations that could associate to metabolic disease.

Editor's evaluation

This well executed study looks at the association of urinary metabolites to the types of diets consumed by European children. Using NMR they find four metabolites that are predictive of a Mediterranean diet. This presents both an approach additional to traditional questionnaire methods and potential insights into biological pathways and will be of interest to nutritionists and epidemiologists.

Introduction

Dietary habits are considered a key element for the prevention of chronic noncommunicable diseases (*Grosso et al., 2020*). In 2017, 22% of all deaths among adults were attributed to dietary risks with type 2 diabetes being among the top causes of diet-related deaths (*Collaborators, 2019*). However, it is notoriously difficult to measure diet accurately in large population studies and the absence of accurate dietary assessment methods is hampering the evidence linking diet and disease (*Posma et al., 2020*; *Ioannidis, 2018*). There is a need for novel approaches to better elucidate diet-related metabolic alterations and their association with disease risk.

Metabolomics is the systematic study of small-molecule metabolites in a biological system and has recently emerged as a powerful top-down approach providing a comprehensive phenotype of biological status. Urinary metabolic phenotypes carry rich information on environmental, lifestyle and nutritional exposures, physiological and metabolic status, and disease risks on an individual and population level (*Gibson et al., 2020*; *Collins et al., 2019*; *Rebholz et al., 2018*). Urine specimens have high concentrations of food-derived metabolites and studies have shown that urinary metabolic profiles could provide an objective measure of dietary intake (*O'Gorman and Brennan, 2017*).

Previous research identifying diet-related metabolic profiles has largely focused on selected food groups including fruits, vegetables, meat, and seafood (*Gibson et al., 2020*; *Lau et al., 2018*; *Guertin et al., 2014*; *Playdon et al., 2016*; *Scalbert et al., 2014*), while dietary patterns and food processing are far less studied (*Collins et al., 2019*; *Rebholz et al., 2018*; *Garcia-Perez et al., 2017*; *Martinez et al., 2016*). Ultra-processed foods (UPFs), which are industrial formulations undergoing a series of physical and chemical processes and typically lack intact healthy food components and include various additives, can result to cumulative intake of salt, added sugars, and fats (*Monteiro et al., 2019*). UPF consumption has been increasing worldwide (*Monteiro et al., 2013*; *Monteiro et al., 2018a*; *Monteiro et al., 2018b*; *Baker et al., 2020*) and, to our knowledge, no previous report exists on its metabolic signature. In contrast with the study of overall diets, including the UPF diet, exploring intakes of single foods, which is what is traditionally done in nutrition research, might be failing to provide a realistic image of dietary metabolic footprint. Moreover, most previous studies have focused on adults, and little is known about body's metabolic response to diet during childhood. This is important as age is a major source of variation in metabolite profiling, with large differences being observed between children and adults (*Ellul et al., 2019*).

Alterations in metabolite profiling could provide a mechanistic link between diet and disease development as early as in childhood. Alterations in the levels of several metabolic biomarkers and pathways have been associated with insulin resistance in childhood including branched-chain and aromatic amino acid metabolism, urea cycle, glucose, and carbohydrate metabolism (*Rauschert et al., 2017*; *Zhao et al., 2016*; *Martos-Moreno et al., 2017*). In addition, dietary patterns characterized by high consumption of UPFs, such soft drinks, sweet, and savoury snacks, have been related to higher risk of insulin resistance in children (*Karatzi et al., 2014*; *Romero-Polvo et al., 2012*). All previous studies

have included measurement of serum insulin levels as marker of insulin resistance. An alternative marker is C-peptide, a protein that is cosecreted with insulin on an equimolar basis from pancreatic β-cells and has been shown to strongly predict metabolic disease progression (*Patel et al., 2012*). C-peptide has a longer half-life than insulin and is recognized as a stable and accurate marker of endogenous insulin secretion, even in nonfasting conditions (*Hope et al., 2016*; *Vezzosi et al., 2007*; *Polonsky et al., 1986*). Previous studies in children have shown that higher carbohydrate intake is related to higher C-peptide concentrations (*Sunehag et al., 2002*; *Buyken et al., 2006*). However, there is little evidence on the metabolic signatures underlying the association of diet with C-peptide levels in children.

We conducted a multicountry study in European children within the Human Early-Life Exposome (HELIX) project (*Maitre et al., 2018*) aiming (1) to identify urinary metabolites associated with Mediterranean diet adherence and UPF consumption, and (2) to determine the extent to which these metabolites were associated with C-peptide, used as an early marker of metabolic health.

Methods
Study population
This study is embedded within the HELIX project (*Maitre et al., 2018*), a collaborative project across six established and ongoing longitudinal population-based birth cohort studies in Europe: Born in Bradford (BiB, UK) (*Wright et al., 2013*), Étude des Déterminants pré et postnatals du développement et de la santé de l'Enfant (EDEN, France) (*Heude et al., 2016*), Kaunas Cohort (KANC, Lithuania) (*Grazuleviciene et al., 2009*), INfancia y Medio Ambiente (INMA, Spain) (*Guxens et al., 2012*), Norwegian Mother, Father and Child Cohort Study (MoBa, Norway) (*Magnus et al., 2016*), and RHEA (RHEA, Greece) (*Chatzi et al., 2017a*). Participating cohorts covered singleton deliveries from 2003 to 2008. As part of HELIX, a subcohort of 1301 children (approximately 200 children in each cohort) were followed in 2014–2015 for a clinical examination, a computer-assisted interview with the parents, and the collection of biological samples. Data collection was standardized across cohorts and performed by trained staff. A full description of the HELIX follow-up methods and study population are provided by *Maitre et al., 2018*. Prior to the start of HELIX, all six cohorts on which HELIX is based had
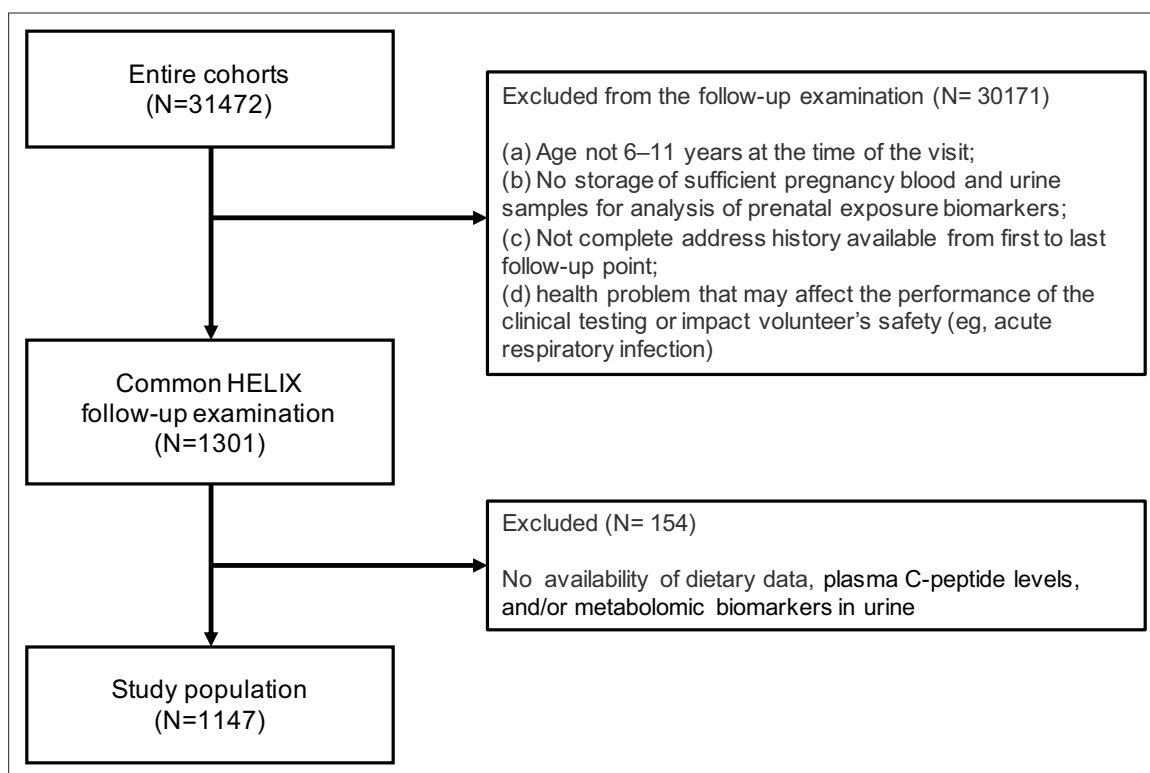

Figure 1. Participant flowchart.

undergone the required evaluation by national ethics committees and had obtained all the required permissions for their cohort recruitment and follow-up visits. Each cohort also confirmed that relevant informed consent and approval were in place for secondary use of data from preexisting data. The work in HELIX was covered by new ethics approvals from the local ethics committees at each site, and at enrolment in the HELIX subcohort, participants were asked to sign an informed consent form for the specific HELIX work including clinical examination and biospecimen collection and analysis. Additionally, the current study was approved by the University of Southern California Institutional Review Board.

Our study population consisted of 1147 children with available information on dietary intake, plasma C-peptide levels, and metabolomic biomarkers in urine collected during the HELIX follow-up at a mean age of 7.9 years (range: 5.4–12.0 years) (*Figure 1*).

Dietary assessment

Information about the children's habitual diet was collected via a semi quantitative food-frequency questionnaire (FFQ) covering the child's habitual diet, which was filled in by the parent attending the examination appointment. The FFQ, covering the past year, was developed by the HELIX research group, translated and applied to all cohorts (*Maitre et al., 2018*). It included 43 questions of intake of food items, which were aggregated in 16 main food groups: meat and meat products, fish and seafood, sweets, beverages, potatoes, vegetables, dairy products, fruits, bread and cereal, sweet bakery products, added fats, eggs, nuts, salty snacks, pulses, and dressings. It also included 15 specific questions to examining the degree of adherence to Mediterranean diet. Diet quality was assessed using two different approaches, (1) by assessing the degree of adherence to a Mediterranean diet based on the KIDMED index (Mediterranean Diet Quality Index for children and adolescents) (*Serra-Majem et al., 2004*) and (2) by assessing the proportion of UPF in the overall diet (*Monteiro et al., 2019*).

For the KIDMED index (*Serra-Majem et al., 2004*), items positively associated with the Mediterranean diet pattern (11 items) were assigned a value of +1, while those negatively associated with the Mediterranean diet pattern (4 items) were assigned a value of −1 (*Supplementary file 1a*). The scores for all 15 items were summed, resulting in a total KIDMED score ranging from −4 to 11, with higher scores reflecting greater adherence to a Mediterranean diet. We categorized the score into three groups: low (<1), moderate (1–4), and high (>4).

For UPF intake, we identified foods and drinks as 'ultra-processed' by using the NOVA classification, a food classification system based on the nature, extent, and purpose of industrial food processing (*Monteiro et al., 2019*) We identified the following 'ultra-processed' foods: cookies, pastries, sugar-sweetened, low-sugar and artificially sweetened beverages, cold meat cuts; ham, dairy desserts, sugar-sweetened and other breakfast cereals, crispbread and rusks; chocolate, sweets, margarine, dressings, and salty snacks. For some food items, our FFQ did not provide enough information on food processing to determine if a specific item belongs to one processing category or another. We discussed the classification of each food item with a team of nutritionists and used a conservative approach, such that the lower level of processing was chosen – for instance, we made the assumption that fries are homemade from fresh potatoes, and therefore, they were not classified as UPF. For each child, we calculated the daily proportion of all UPF in the total diet as the ratio between the sum of daily servings of UPF to the total daily sum of all food and drink servings. More details on the categorization of foods according to the NOVA classification are presented in *Supplementary file 1b*.

Urine metabolite profiling

Two urine samples, representing last night-time and first morning voids, were collected on the evening and morning before the clinical examination, kept in a fridge and transported in a temperature-controlled environment, and aliquoted and frozen within 3 hr of arrival at the clinics. They were subsequently pooled to generate a more representative sample of the last 24 hr for metabolomic analysis (*Lau et al., 2018*).

Urinary metabolic profiles were acquired using ^{1}H NMR spectroscopy according to *Lau et al., 2018*. In brief one-dimensional 600 MHz ^{1}H NMR spectra of urine samples from each cohort were acquired on the same Bruker Avance III spectrometer operating at 14.1 Tesla within a period of 1 month. The spectrometer was equipped with a Bruker SampleJet system, and a 5-mm broadband inverse configuration probe maintained at 300 K. Prior to analysis, cohort samples were randomized. Deuterated

3-(trimethylsilyl)-[2,2,3,3-d$_4$]-propionic acid sodium salt was used as internal reference. Aliquots of the study pooled quality control (QC) sample were used to monitor analytical performance throughout the run and were analysed at an interval of every 23 samples (i.e., 4 QC samples per well plate). The ^{1}H NMR spectra were acquired using a standard one-dimensional solvent suppression pulse sequence. Forty-four metabolites were identified and quantified as described in *Lau et al., 2018*. The urinary NMR showed excellent analytical performance, the mean coefficient of variation across the 44 NMR detected urinary metabolites was 11%. For the statistical analysis we have used creatinine-normalized metabolite concentrations (μmol/mmol of creatinine).

Plasma C-peptide

Blood was collected at the end of the clinical examination during the HELIX follow-up visit. The median postprandial interval (time between last meal and blood collection) was 3.3 hr (interquartile range [IQR: 2.8–4.0]).

For each cohort, concentration of C-peptide was assessed in child plasma at the CRG/UPF Proteomics Unit (Barcelona, Spain) using the xMAP and Luminex System multiplex platform according to the manufacturer's protocol. Blood samples were randomized and blocked by cohort prior to measurement to ensure a representation of each cohort in each plate (batch). For protein quantification, an 8-point calibration curve per plate was performed with protein standards provided in the Luminex kit and following the procedures described in the standard procedures described by the vendor. Commercial heat inactivated, sterile-filtered plasma from human male AB plasma (Sigma Cat # H3667) was used as constant controls to control for intra- and interplate variability. Four control samples were added per plate. Raw intensities obtained with the xMAP and Luminex system for each sample were converted to pg/ml using the calculated standard curves of each plate and accounting for the dilutions that were made prior measurement. The coefficient of variation for C-peptide was 16%. The LOD was determined and the lower and upper quantification limits (LOQ1 and LOQ2, respectively) were obtained from the calibration curves. C-peptide concentrations were log2-transformed to achieve normal distribution. Plate batch effect was corrected by subtracting for each individual and each protein the difference between the overall protein average minus the plate-specific protein average. Finally, values below LOQ1 and above LOQ2 were imputed using a truncated normal distribution using the truncdist R package.

Covariates

Adjustment factors were selected a priori based on literature (*Aranceta et al., 2003*; *Scaglioni et al., 2018*; *Patrick and Nicklas, 2005*) and included: maternal age (in years), maternal education level (low, middle, high), maternal prepregnancy body mass index (BMI, in kg/m^2), family affluence score (cohort-specific definition of low, middle, high), child sex, child age (in years), child BMI (in kg/m^2), child sedentary behavior (min/day of time spent watching TV, playing computer games or other sedentary games), child ethnicity (White European, Asian, other), and postprandial interval (in hours). We also included a cohort indicator as a fixed effect in the models, as this, in the context of an observational study, is expected to control for cohort effects (*Basagaña et al., 2018*). We imputed missing values for covariates (ranging from 0% to 4%) using the method of chained equations with the R package *mice*. Details about the imputation process in HELIX, diagnostics, and comparison between imputed and complete-case values have been reported in detail elsewhere (*Agier et al., 2019*).

Statistical analysis

As a first step in our analysis, we conducted a metabolome-wide association study to assess the associations of urinary metabolites with diet quality. Creatinine-normalized metabolite concentrations (μmol/mmol of creatinine) were log$_{10}$ transformed prior to statistical analyses to improve model fit. We fitted separate multivariable regression models for each metabolite with the KIDMED score or UPF intake (expressed as per 5% change of total daily food intake). To account for multiple hypothesis testing, we applied the Benjamini–Hochberg false discovery rate (FDR) correction; an FDR-corrected p value <0.05 denoted statistical significance. For metabolites identified to be associated with the diet quality indicators, we assessed between-cohort heterogeneity with the I^2 statistic and χ^2 test from Cochran's Q.

To examine the ability of the identified metabolite panels in discriminating children with low vs. high KIDMED scores (<1 vs. >4) and low vs. high UPF intakes (Quartile 1: <18% vs. Quartile 4: ≥29%), we plotted receiver operating characteristic (ROC) curves and estimated area under the ROC curve values, indicative of the discriminative performance of the metabolite models, based on tenfold cross validations. We also repeated the ROC analysis for a set of established sociodemographic factors (maternal education level, maternal prepregnancy BMI, family affluence score, child sedentary behavior, ethnicity, age, and sex) linked to childhood diet quality both previously (*Aranceta et al., 2003*; *Scaglioni et al., 2018*; *Patrick and Nicklas, 2005*) and in our study population (all p < 0.05) and compared the discriminative performance of this sociodemographic set with that of the metabolites.

Next, we examined the association of the KIDMED score and of UPF intake (as independent variables) with plasma C-peptide concentration (as dependent variable) using multivariable linear regression models. No departures from linearity in the associations of these diet quality indicators with C-peptide concentration were observed both visually and statistically (p for linearity >0.42) using generalized additive models. We examined the KIDMED score both as continuous (per score unit increase) and in categories of low (score <1, reference), moderate (score = 1–4), and high (score >4). Likewise, UPF intake was assessed both as continuous (per 5% change of total daily food intake) and in quartiles (Q1: <18%, reference; Q2: 18% to <23%; Q3: 23% to <29%; and Q4: ≥29% of total daily food intake). We also included a product term between KIDMED and UPF intake in the regression analysis to assess their interaction; to simplify interpretation of this model, we categorized the KIDMED score as low/moderate vs. high and UPF intake based on the median population intake (<23% vs. ≥23%). We conducted two sets of sensitivity analyses to assess the robustness of the results. First, we calculated cohort-specific effect estimates and assessed heterogeneity with the I^2 statistic and χ^2 test from Cochran's *Q*. Second, we examined potential effect modification by child sex and by child weight status (IOTF-defined normal weight vs. overweight/obese) on C-peptide by testing the multiplicative interaction term between the potential effect modifier and each diet quality measure. Finally, we fitted regression models with the metabolites found to be associated

Table 1. Characteristics of the study population.

Cohort of inclusion, *n* (%)	
BiB, UK	189 (16.5)
EDEN, France	149 (13)
INMA, Spain	202 (17.6)
KANC, Lithuania	194 (16.9)
MoBa, Norway	221 (19.3)
RHEA, Greece	192 (16.7)
Maternal characteristics	
Maternal age, mean (SD), years	30.7 (4.9)
Missing, *n* (%)	13 (1.1)
Prepregnancy BMI, mean (SD), kg/m^2	25 (5)
Missing, *n* (%)	21 (1.8)
Maternal educational level, *n* (%)	
Low	157 (13.7)
Medium	391 (34.1)
High	562 (49)
Missing, *n* (%)	37 (3.2)
Child characteristics	
Age at assessment, mean (SD), years	7.9 (1.6)
Sex, *n* (%)	
Male	626 (54.6)
Female	521 (45.4)
Ethnicity, *n* (%)	
White European	1,028 (89.6)
Asian	92 (8)
Other	27 (2.4)
Family affluence score, *n* (%)	
Low	126 (11)
Medium	448 (39.1)
High	569 (49.6)
Missing, *n* (%)	4 (0.4)
BMI, mean (SD), kg/m^2	16.9 (2.6)
Normal weight, *n* (%)*	906 (79)
Overweight/obese, *n* (%)*	237 (20.7)
Missing, *n* (%)	4 (0.4)
KIDMED score, mean (SD)	2.8 (1.7)
Low (<1), *n* (%)	104 (9.1)
Medium (1–4), *n* (%)	848 (73.9)
High (>4), *n* (%)	195 (17)

Table 1 continued on next page

Table 1 continued

Cohort of inclusion, *n* (%)	
Ultra-processed food intake, mean (SD), % of daily food intake	24.2 (8.7)

*Categories of normal weight and overweight/obese were derived using the International Obesity Taskforce criteria (**Cole and Lobstein, 2012**).

BiB = Born in Bradford cohort. EDEN = the Étude des Déterminants pré et postnatals du développement et de la santé de l'Enfant study. INMA = INfancia y Medio Ambiente cohort. KANC = Kaunas Cohort. KIDMED = Mediterranean Diet Quality Index for children and adolescents. MoBa = Norwegian Mother, Father and Child Cohort Study. RHEA = Rhea Mother Child Cohort study.

with diet quality indicators, and C-peptide in order to assess whether diet-related metabolites are related to β-cell function.

We performed analyses with both complete (missingness <4% in each covariate) and imputed data. Results were similar across raw and imputed data analyses, and hence, we present those using the imputed covariate data. For easier interpretation of effect estimates for the log-transformed C-peptide and metabolite values, we back-transformed regression coefficients and present results as percent change (% change = (back-transformed [beta] − 1) × 100).

Analyses were conducted using STATA version 14.2 (StataCorp LLC, TX) and R software version 3.5.3. Linear regression analyses were performed in STATA with the command '*regress*' and using '*mi estimate*' to account for the imputed covariate data (**StataCorp, 2021**). ROC analysis was performed in R with the caret package (**Kuhn, 2008**). Visualizations of the results were carried out using the ggplot2 package in R (**Wickham, 2016**).

Results
Study population
Among participating children (*n* = 1147), 626 (54.6%) were boys and 1028 (89.6%) were white (**Table 1**). The mean (SD) age at assessment was 7.9 (1.6) years. Median (IQR) C-peptide concentration was 1.26 (0.03, 1.95) ng/ml.

Seventeen percent of children (*n* = 195) had a high KIDMED score (>4), indicative of high adherence to the Mediterranean diet. Mean (SD) UPF intake in the overall study population was 24.2 (8.7)% of total daily food intake. KIDMED and UPF intake were negatively correlated (Spearman *r* = −0.44); children with high and low (<1) KIDMED score had a mean (SD) UPF intake of 18.8 (6.6)% and 33.4 (9.9)%, respectively. The two diet quality scores were associated with most recorded food intakes in opposite directions (**Supplementary file 1c, d**). Daily fruit and vegetable consumption, weekly fish consumption and not skipping breakfast were the major dietary habits differentiating children with low and high KIDMED score (**Supplementary file 1c**). Intakes of pastries and bakery products, dairy
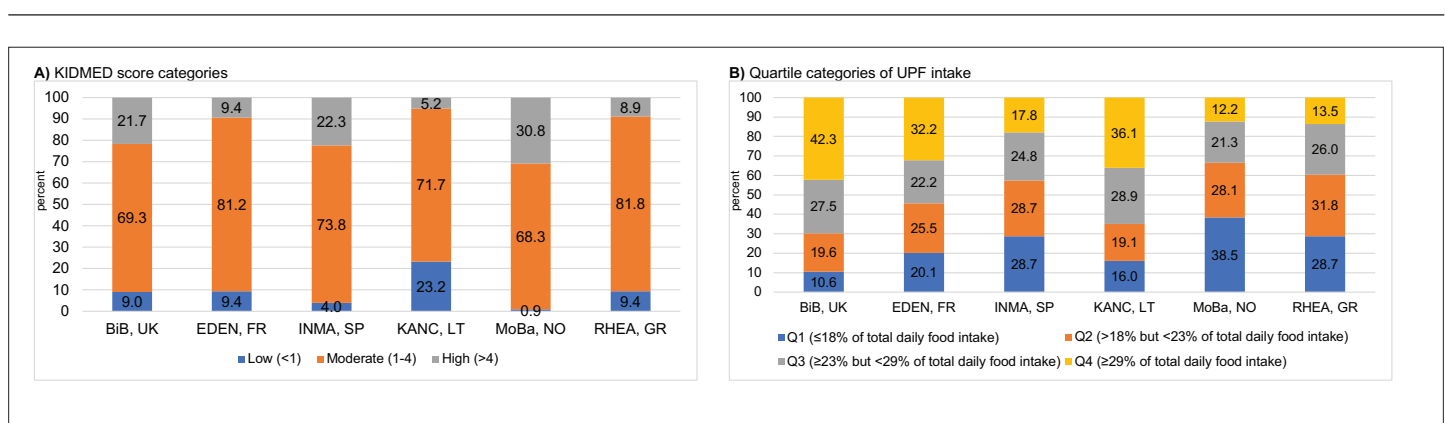
Figure 2. Levels of childhood adherence to the diet quality indicators of interest in each Human Early-Life Exposome (HELIX) subcohort. Panel (**A**) illustrates the levels of adherence to the Mediterranean diet which were defined as follows: low, KIDMED score, <1; moderate, KIDMED score, 1–4; and high, >4. Panel (**B**) illustrates the levels of ultra-processed food consumption (expressed as % of total daily food intake) which are based on quartile (*Q*) cutoffs according to the intake distribution of the overall HELIX study population. BiB, Born in Bradford cohort; EDEN, the Étude des Déterminants pré et postnatals du développement et de la santé de l'Enfant study; INMA, INfancia y Medio Ambiente cohort; KANC, Kaunas Cohort; KIDMED, Mediterranean Diet Quality Index for children and adolescents; MoBa, Norwegian Mother, Father and Child Cohort Study; RHEA, Rhea Mother Child Cohort study.
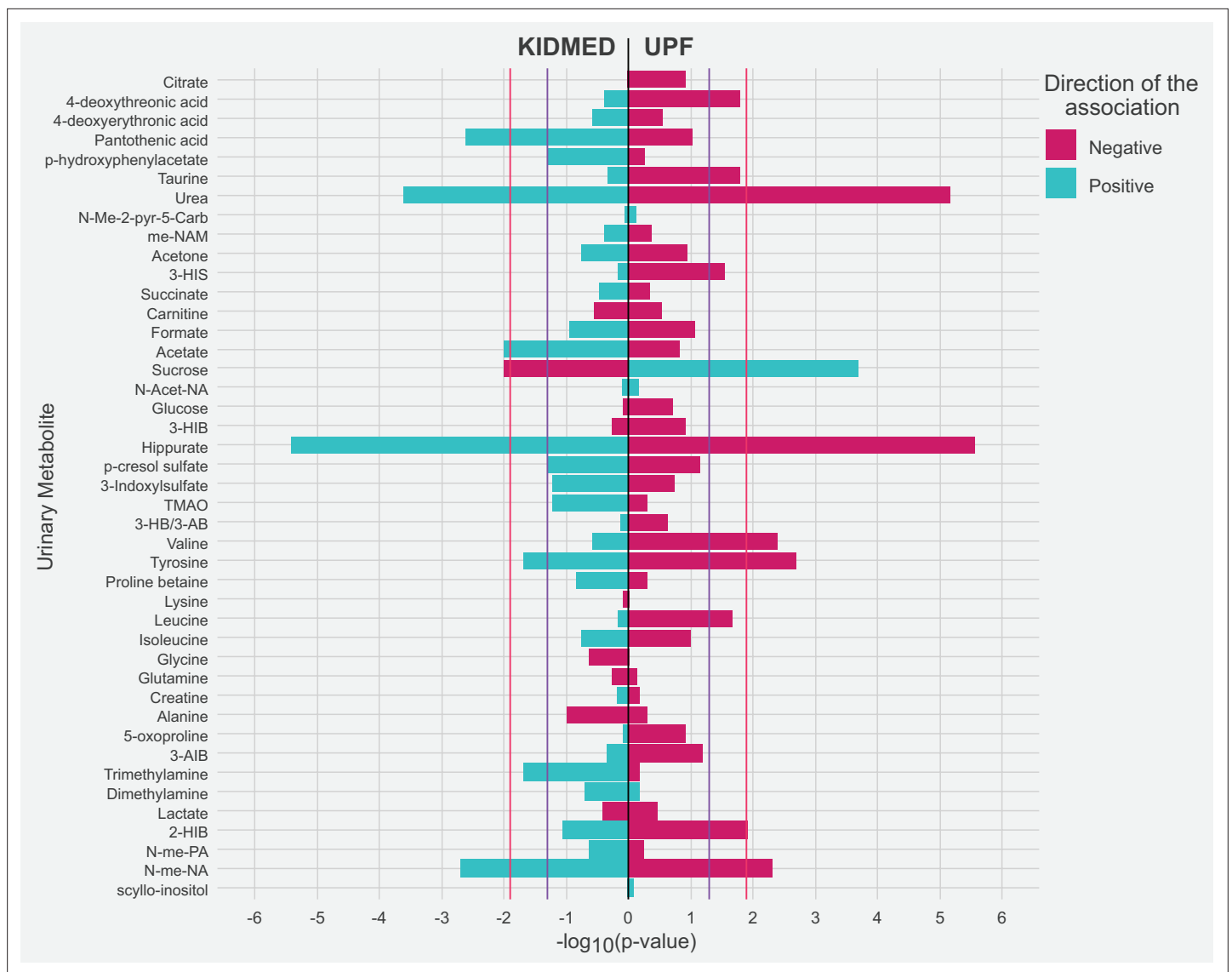
Figure 3. Adjusted associations of the diet quality indicators of interest with urinary metabolites in childhood. Linear regression models were adjusted for maternal age, maternal education level, maternal prepregnancy body mass index (BMI), family affluence status, child sex, child age, child BMI, child sedentary behavior, child ethnicity, and a cohort indicator. The purple line represents a p value of 0.05. The red line represents an false discovery rate (FDR)-adjusted p value of 0.05. 2-HIB, 2-hydroxyisobutyrate; 3-AIB, 3-aminoisobutyrate; 3-HB/3-AB, 3-hydroxybutyrate/3-aminoisobutyrate; 3-HIB, 3-hydroxyisobutyrate; 3-HIS, 3-hydroxyisovalerate; me-NAM, N1-methyl-nicotinamide; N-Acet-NA, N-acetyl neuraminic acid; N-Me-2-pyr-5-Carb, N-methyl-2-pyridone-5-carboxamide; N-me-NA, N-methylnicotinic acid; N-me-PA, N-methylpicolinic acid; TMAO, trimethylamine N-oxide.

desserts, margarine, and dressings were the major determinants of UPF intake (**Supplementary file 1d**). Children with the highest KIDMED scores were mostly from Norway and Spain (**Figure 2A**), while those with the highest UPF intake were mostly from Lithuania and the UK (**Figure 2B**).

Diet quality and urinary metabolome

Figure 3 and **Supplementary file 1e, f** show the associations of diet quality indicators with urinary metabolite levels in childhood. KIDMED and UPF intake exhibited an opposite pattern of association for most of the metabolites (30 out of 43). After controlling for FDR (FDR-corrected p value <0.05), we found that a panel of four metabolites related to both diet quality indicators. Specifically, a higher KIDMED score was associated with higher levels of hippurate, N-methylnicotinic acid, and urea and with lower levels of sucrose; UPF intake exhibited opposite associations with these four metabolites. A higher KIDMED score was also associated with higher acetate and pantothenic acid concentrations,
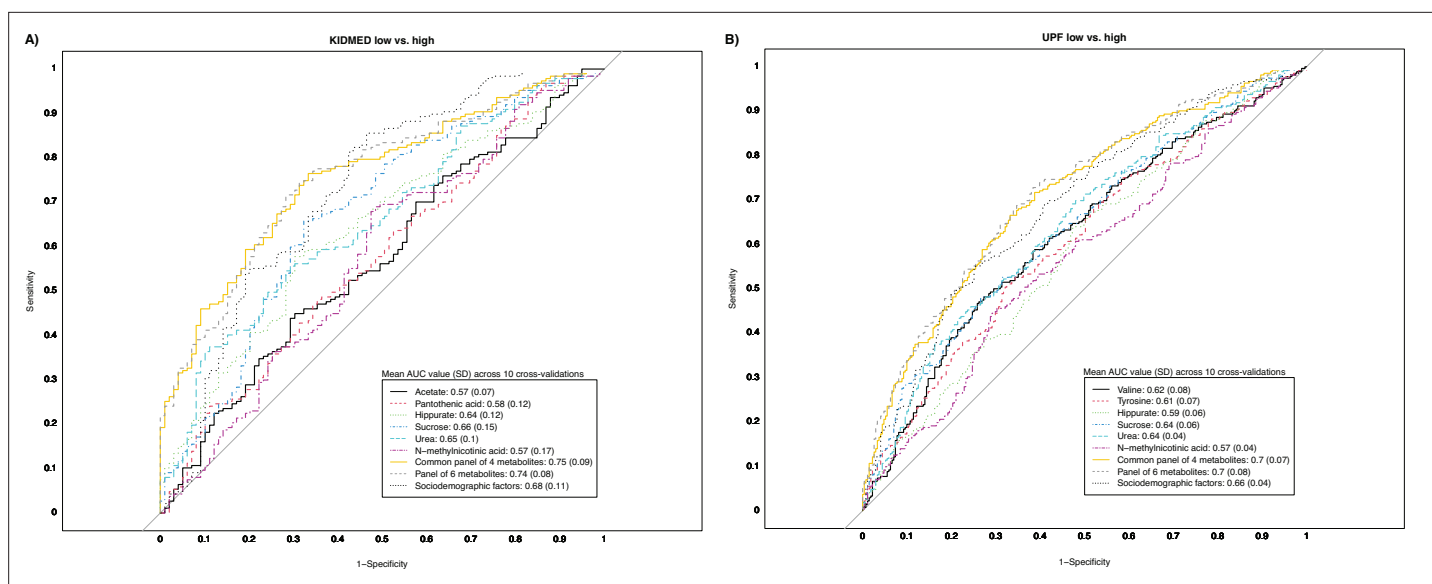
Figure 4. Receiver operating characteristic (ROC) curves reflecting the ability of urinary metabolites of interest in discriminating adherence to diet quality in childhood. Panel (**A**) illustrates the ability of urinary metabolites of interest in discriminating high adherence to the Mediterranean diet (KIDMED >4) from low adherence (KIDMED <1). Panel (**B**) illustrates the ability of urinary metabolites of interest in discriminating high ultra-processed food consumption (UPF ≥29% of total intake) from low consumption (UPF <18% of total daily food intake). ROC curves are based on models across the full study sample, and discriminative power is evaluated based on tenfold cross-validation. The mean area under the receiver operating characteristic curve (AUC) value (SD) across the ten cross-validations of each model is presented in the box. The common panel of four metabolites includes the metabolites associated with both diet quality indicators (hippurate, sucrose, urea, and *N*-methylnicotinid acid). The panel of six metabolites includes the metabolites associated with each diet quality indicator (common panel of four plus acetate and pantothenic acid for KIDMED, and plus valine and tyrosine for UPF). The panel of sociodemographic factors includes maternal education level, maternal prepregnancy BMI, family affluence score, child sedentary behavior, ethnicity, age, and sex.

while tyrosine and valine concentrations were inversely associated with UPF intake. There was no evidence of significant between-cohort heterogeneity in the associations between the diet quality indicators and these metabolites (<30%; p for heterogeneity >0.2). ROC curve analyses showed that the combination of four metabolites associated with both diet quality indicators performed better than individual metabolites in discriminating children with high and low KIDMED scores and UPF intake (*Figure 4*). The discriminative ability of this metabolite panel was not improved with the addition of urinary metabolites specifically linked to each diet quality indicator and was equal or even greater to that of established sociodemographic factors previously linked to diet quality in childhood. The regression formulas (scores) for predicting children's diet quality indicators based on the urinary metabolites are given in *Supplementary file 1g*.

Diet quality and C-peptide levels

Table 2 presents the associations of diet quality indicators with C-peptide concentration in childhood. We found that a higher KIDMED score was associated with lower C-peptide. Specifically, compared to a low KIDMED score (<1), children with a moderate score (1–4) had a 28% lower C-peptide concentration (percent change: −27.7, 95% CI: −49.6 to 3.9) and those with a high score (>4) had a 39% lower C-peptide concentration (percent change: −39.0, 95% CI: −60.6 to −5.7) (p-trend = 0.03). An opposite association was observed for UPF intake. Compared to children at the lowest quartile of UPF intake (<18% of total daily food intake), children at the second quartile (18% to <23% of total daily food intake) had a 24% higher C-peptide concentration (percent change: 24.3, 95% CI: −6.4 to 65.2), those at the third quartile (23% to < 29% of total daily food intake) had a 39% higher concentration (percent change: 38.5, 95% CI: 3.8–84.9), and those at the fourth quartile (≥29% of total daily food intake) had a 46% higher concentration (percent change: 46.0, 95% CI: 8.1–97.3) (p-trend = 0.01). There was no evidence of interaction between the diet quality indicators (*Supplementary file 1h*). When we examined the associations separately in each cohort, we found no significant between-cohort heterogeneity (I^2 <18%, p for heterogeneity >0.29) (*Figure 5*). We also found no evidence that

Table 2. Adjusted associations of diet quality with C-peptide levels in childhood*.

	C-peptide
	Percent change (95% CI)
KIDMED score (per unit increase)	−8.1 (−13.7, −2.2)
Low (<1)	Ref.
Moderate (1–4)	−27.7 (−49.6, 3.9)
High (>4)	−39.0 (−60.6, −5.7)
p-Trend	0.03
UPF intake (per 5% increase of total intake)	9.3 (2.8, 16.2)
Q1 (<18% of total intake)	Ref.
Q2 (18% to <23% of total intake)	24.3 (−6.4, 65.2)
Q3 (23% to <29% of total intake)	38.5 (3.8, 84.9)
Q4 (≥29% of total intake)	46.0 (8.1, 97.3)
p-Trend	0.01

*Effect estimates represent percent changes in log-2 transformed C-peptide levels and their 95% CIs derived from linear regression models adjusted for maternal age, maternal education level, maternal prepregnancy BMI, family affluence status, child sex, child age, child BMI, child sedentary behavior, child ethnicity, postprandial interval, and a cohort indicator. KIDMED = Mediterranean Diet Quality Index for children and adolescents. UPF = ultra-processed food.

the observed associations differed by the sex of the children or their weight status (*Supplementary file 1i*).

Diet-related urinary metabolites and C-peptide levels

When we examined the associations with the metabolite scores for each diet quality indicator, we found that the scores for KIDMED were associated with lower C-peptide, while opposite associations were observed with the metabolite scores for UPF (*Supplementary file 1j*). In the pairwise associations between the individual urinary metabolites linked to KIDMED or UPF intake and C-peptide, we found that higher levels of sucrose were associated with higher C-peptide levels.

Discussion

This is the first study of European children that identified a panel of urinary metabolites associated with diet quality, as assessed by a Mediterranean diet adherence score (KIDMED) and UPF intake. These metabolites include food constituents (or their metabolic products), vitamin-related compounds and metabolites related to amino acid, protein, and carbohydrate metabolism. For both diet quality indices, KIDMED score and UPF intake, there was a common panel of four metabolites exhibiting an opposite pattern of
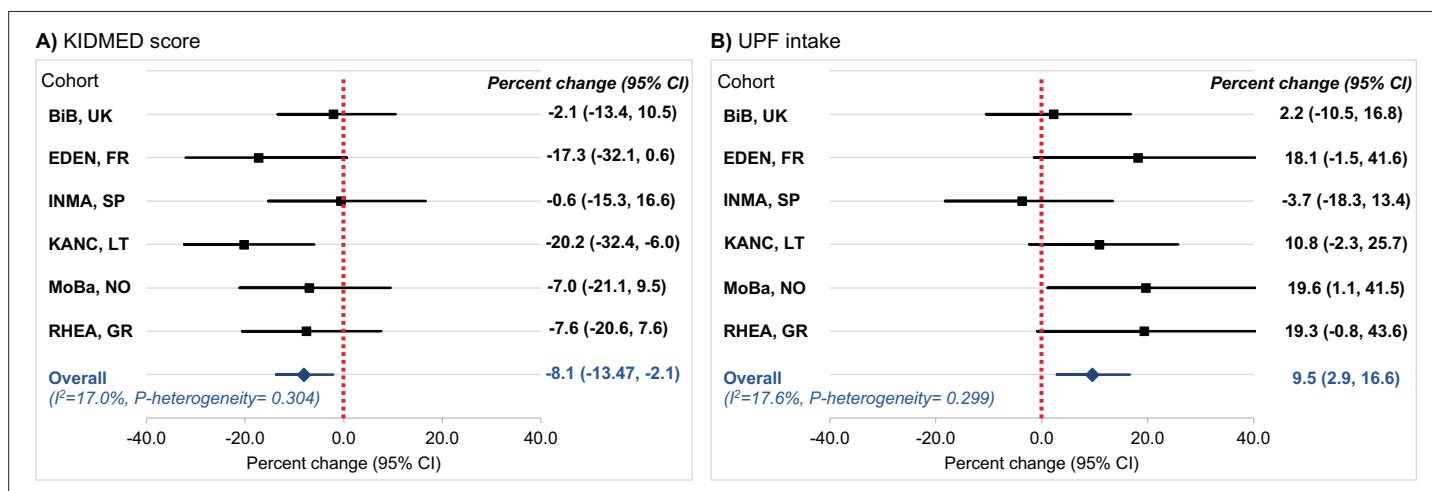
Figure 5. Cohort-specific associations of the diet quality indicators of interest with C-peptide in childhood. Panel (**A**) illustrates the associations for adherence to the Mediterranean diet, which was assessed via the KIDMED score (expressed per unit increase). Panel (**B**) illustrates the associations for ultra-processed food (UPF) intake (expressed per 5% increase of total daily food intake). Beta coefficients (95% confidence intervals, CIs) by cohort were obtained using linear regression models adjusted for maternal age, maternal education level, maternal prepregnancy body mass index (BMI), family affluence status, child sex, child age, child BMI, child sedentary behavior, child ethnicity, and postprandial interval. Combined estimates were obtained by using a fixed-effects meta-analysis. Squares represent the cohort-specific effect estimates; diamond represents the combined estimate; and horizontal lines denote 95% CIs. BiB, Born in Bradford cohort; EDEN, the Étude des Déterminants pré et postnatals du développement et de la santé de l'Enfant study; INMA, INfancia y Medio Ambiente cohort; KANC, Kaunas Cohort; MoBa, Norwegian Mother, Father and Child Cohort Study; RHEA, Rhea Mother Child Cohort study.

association: higher levels of hippurate, *N*-methylnicotinic acid, and urea, and lower levels of sucrose were concomitant with a higher KIDMED score and lower UPF intake. Moreover, these urinary metabolites that reflected these diet quality indicators, were also associated with C-peptide levels, a well-known marker of β-cell function and insulin resistance (*Patel et al., 2012*).

Previous studies assessing the metabolic response of the body to diet have largely focused on specific foods groups, *Collins et al., 2019* yet humans eat a combination of foods that may have additive or interactive effects on human physiology. In our study, we assessed two diet quality indices to better describe the complexity of overall diet. The Mediterranean diet is characterized by high intake of fruits, vegetables, legumes, nuts, and whole grain products, fish and low intakes of red meat and sweets, and it has long been appraised for its cardiometabolic benefits, even in children. (*Chatzi et al., 2017b*) Moreover, we assessed UPF intake to characterize the intake of industrial processed and prepared food items, with altered food structure, nutritional content, and taste (*Elizabeth et al., 2020*). UPF intake assessment diverts from the traditional strategy of studying nutrients, foods, or dietary patterns to identify the link between diet with health and disease, and it reflects the cumulative intake of artificial substances (i.e. flavorings, colorings, emulsifiers, and other additives) and processing byproducts. Effects of UPF consumption are likely to be the result of synergistic effects of many food ingredient compound and characteristics, and high consumption of UPF has been proposed to lead to metabolic dysregulation (*Srour et al., 2020*).

For both KIDMED and UPF, there was a common panel of four urinary metabolites (hippurate, *N*-methylnicotinic acid, urea, and sucrose), exhibiting an opposite pattern of association. In addition, KIDMED was positively associated with pantothenic acid and acetate, whereas UPF was negatively associated with two amino acids, valine and tyrosine. These metabolic signatures demonstrate a core common metabolic biomarker panel reflecting habitual dietary intake, but also show that the two diet quality indices and their metabolic signatures act complementary and can highlight different aspects of human metabolism and physiology. Our finding are consistent with previous metabolomics studies conducted in adults (*Garcia-Perez et al., 2017*; *Garcia-Perez et al., 2020*; *Almanza-Aguilera et al., 2017*). Four dietary interventions were developed (*Garcia-Perez et al., 2017*) with similar energy content and within the World Health Organization (WHO healthy eating guidelines), but with varying macro- and micronutrient intake. The urine metabolic profile for the diet most concordant with the guidelines, characterized by high intakes of dietary fiber (through fruits, vegetables, and whole grain cereal products) and low intakes of fat, sugar, and salt showed systematic differences for a total of 28 metabolites, including increased levels of acetate, hippurate, *N*-methylnicotinic acid, and urea, compared to the diet that diverted from the healthy guidelines (*Garcia-Perez et al., 2017*). The PREDIMED study also used NMR to define urinary biomarkers associated with a high adherence to a Mediterranean diet pattern in adults and the proposed biomarkers also included higher levels of urea (*Almanza-Aguilera et al., 2017*). To our knowledge, our study is the first to show the urinary metabolic footprint of total UPF intake, and also the first to show that urine NMR-derived scores of diet quality are reflective of a key biomarker of metabolic health, C-peptide, in healthy children. The ability to use a rapid, noninvasive biofluid screen to measure objective biomarkers of diet quality in children opens new avenues for exploring the significance of nutritional patterns to healthy development early in life.

Among the diet-related urinary metabolites, sucrose individually reflected the associations observed between diet quality and C-peptide (*World Health Organization, 2015*). Elevated levels of added sugars could lead to elevated glucose load in the human body, which in turn can lead to an increase in glycemic response and, thus, C-peptide production in healthy populations. To the best of our knowledge, there is only one previous study from Mexico examining the relation of diet to C-peptide in children (*Perng et al., 2017*). Similar to our findings, this study showed that adherence to a prudent dietary pattern characterized by food groups commonly consumed in Mediterranean diet (vegetables, fruit, fish, and legumes) was associated with low C-peptide levels in boys. Regarding other markers of glucose regulation, our findings are in line with previous studies in children and adults reporting that low adherence to a healthy dietary pattern or high consumption of specific UPFs (sugar-sweetened beverages and ultra-processed meat) were associated with impaired glucose regulation and insulin resistance (*Manios et al., 2010*; *Asghari et al., 2016*; *Chan She Ping-Delfos et al., 2015*; *McKeown et al., 2018*; *Ley et al., 2014*; *Fiorito et al., 2009*). Overall, our findings are consistent with the public health concerns raised by the WHO, relating poor overall quality with high intake level of sugars and

added sugars, and with poor metabolic health and a high risk of several noncommunicable diseases (**World Health Organization, 2015**; **Stanhope et al., 2013**; **WHO Study Group, 1990**; **WHO, 2003**).

Regarding the other urinary metabolites reflecting diet quality, hippurate is a normal component of urine, a metabolic product of phenolic compounds which are present in various dietary sources. It is also a biomarker of fruit/vegetable intake, as it has been confirmed in previous studies in healthy children, adolescents (**Lau et al., 2018**; **Krupp et al., 2012**) and adults (**Edmands et al., 2011**) N-methylnicotinic acid (trigonelline) is a product of the metabolism of niacin (vitamin B3) and a biomarker of various dietary sources like legumes (**Sri Harsha et al., 2018**) and fruits (**Lau et al., 2018**). Urea is the principal end product of amino acids and protein catabolism and a potential marker of protein intake from food (**Garcia-Perez et al., 2020**).

In addition, adherence to the Mediterranean diet was also positively associated with urinary levels of pantothenic acid and acetate. Both compounds have a central role in human biochemistry and the metabolism and synthesis of carbohydrates, proteins, and fats. Pantothenic acid (vitamin B5, necessary to form coenzyme-A) is present in many foods, and we have previously reported a positive association between consumption of dairy products and urinary pantothenic acid in the same study population (**Lau et al., 2018**). Further, we have previously shown that BMI is negatively associated with urinary levels of this metabolite (**Lau et al., 2018**) and our results suggest that higher adherence to the Mediterranean diet associates with pantothenic acid independently of the potential influence of BMI. Acetate has multiple dietary sources, it is produced by acetate-producing bacteria in foodstuff and urinary acetate is also modulated by human gut microbial metabolism. In our previous HELIX analysis on specific food group intakes, we have shown a positive association between potato consumption and urinary acetate levels (**Lau et al., 2018**).

Moreover, we found that UPF intake was negatively associated with two urinary amino acids, valine and tyrosine. Tyrosine is regarded as a conditionally essential amino acid in adults and essential in children. Foods high in dietary tyrosine include dairy, meat, eggs, beans, nuts, grains. Tyrosine is a precursor for neurotransmitters and hormones, increases dopamine availability which in turn could enhance cognitive performance (**Kühn et al., 2019**). Valine is an essential branch chain amino acid (BCAA) critical to energy homeostasis, protein and muscle metabolism (**Brosnan and Brosnan, 2006**; **Nie et al., 2018**). In many studies, it has been observed that elevated BCAAs are associated with insulin resistance and diabetes (**Lynch and Adams, 2014**). Also, in our previous HELIX study, **Lau et al., 2018** we found that urinary valine was associated with higher children's BMI. However, it remains to be eludicated whether these associations are causal (e.g., via mTOR activation) or consequential (e.g., due to reduced mitochondrial oxidation) in metabolic disease, (**Lynch and Adams, 2014**) and whether UPF intake plays a role in the etiology of the association of BCAAs with metabolic health.

Strengths and limitations

The main strengths of the study are the multicentric design which included children from six countries spanning north to south in Europe, the use of standardized data collection and biomarker measurement protocols across cohorts, and the fairly large sample size with biomarker data. Identified panels of urinary metabolites had similar or higher capacity in discriminating children's diet quality to that of established sociodemographic determinants. We chose ^{1}H NMR spectroscopy for urinary analysis as this is an inherently high reproducible, high throughput technique suitable for the identification and quantification of urine metabolites, which are typically of high concentrations, without complex sample preparation which could potentially introduce analytical biases. Also, urinary ^{1}H NMR spectroscopy has been applied in many other cohort studies and proposed as an objective method for dietary assessment (**Garcia-Perez et al., 2017**) potentially facilitating comparative studies in the future. We have used a pooled urine sample collection design which combined the last sample before bedtime with the first morning void sample of the following day, and we have shown in our preliminary work that this sample collection strategy has the advantage of reducing diurnal variations (**Lau et al., 2018**; **Maitre et al., 2017**).

Our study has also some limitations. As in any observational study, there is the possibility of unmeasured residual confounding. In our analyses, we took into account a number of sociodemographic and lifestyle factors in childhood (e.g., socioeconomic status, ethnicity) that are associated with both diet quality and glycemic response. We did not have data available on children's physical activity. Nevertheless, we adjusted all our models for sedentary behavior (including time spent in front of the

screen) which has been shown to associate to physical activity levels, as the time devoted to sedentary screen-time activities might affect availability of time devoted for exercise, or vice versa (*Serrano-Sanchez et al., 2011*; *Pearson et al., 2014*; *Aira et al., 2021*) Further, we did not have data available to control for energy intake. However, in all our models, we included BMI of the children, a measure strongly correlated to energy intake, (*Jakes et al., 2004*) and assessed UPF intake as proportion of total food intake. Moreover, the absence of heterogeneity across cohorts with different correlation structures and confounding patterns in their data (*Maitre et al., 2018*), provide evidence to support that unmeasured confounding is unlikely to have influenced the observed associations. Since the data collected are cross-sectional, and there is no temporality in the observed associations, further longitudinal studies examining metabolic and glycemic alterations in relation to diet quality are needed. Although ^{1}H NMR spectroscopy had the advantage of improving the specificity of the quantitation and provided explicit metabolite identification, it limited the number of metabolites being measured and provided partial coverage of the urine metabolome. Absolute quantification of some metabolites with exchangable protons such as urea could also be negatively impacted by the solvent suppression methods required for ^{1}H NMR spectroscopy of urine. Supplementing the current study with other complementary untargeted and targeted metabolomic approaches in future, such as mass spectrometry, would help enhance identification and robust quantification of urinary metabolites associated with diet quality in children.

In summary, this multicenter European study showed that urinary metabolic profiles related to food constituents (or their metabolic products), to amino acid and carbohydrate metabolism reflect adherence to the Mediterranean diet and UPF intake in childhood. Higher adherence to Mediterranean diet, lower UPF intake, and lower levels of the diet-related carbohydrate sucrose were associated with lower C-peptide levels, a marker of β-cell function. These results provide evidence to support efforts by public health authorities to recommend increased adherence to the Mediterranean diet and limiting UPF consumption in childhood. Further prospective studies examining the association of diet quality and related metabolomic profiles with C-peptide and other surrogates of insulin resistance are needed to replicate our findings.

Acknowledgements

Additional information

Competing interests

Muireann Coen: Is an employee of AstraZeneca. The author declares that no other competing interests exist. The other authors declare that no competing interests exist.

Funding

Author ORCIDs

Nikos Stratakis http://orcid.org/0000-0003-4613-0989
Alexandros P Siskos http://orcid.org/0000-0002-5635-7426
Anne Lise Brantsaeter http://orcid.org/0000-0001-6315-7134

Ethics

Prior to the start of HELIX, all six cohorts on which HELIX is based had undergone the required evaluation by national ethics committees and had obtained all the required permissions for their cohort recruitment and follow-up visits. Each cohort also confirmed that relevant informed consent and approval were in place for secondary use of data from preexisting data. The work in HELIX was covered by new ethics approvals from the local ethics committees at each site, and at enrolment in the HELIX subcohort, participants were asked to sign an informed consent form for the specific HELIX work including clinical examination and biospecimen collection and analysis. Additionally, the current study (study ID: HS-20-00390) was approved by the University of Southern California Institutional Review Board.

Decision letter and Author response

Decision letter https://doi.org/10.7554/eLife.71332.sa1
Author response https://doi.org/10.7554/eLife.71332.sa2

Additional files

Supplementary files

• Supplementary file 1. Additional results. (a) KIDMED items, and their scoring, used to assess adherence to the Mediterranean diet in HELIX children. (b) Items included in the FFQ by ultra-processed food (UPF) inclusion and information regarding the extent and purpose of food processing. (c) Intakes of food groups (in servings/week) by categories of the KIDMED score. (d) Intakes of food groups (in servings/week) by quartiles of ultra-processed food intake. (e) Associations of KIDMED score with urinary metabolites in childhood. (f) Associations of ultra-processed food consumption with urinary metabolites in childhood. (g) Regression formulas (scores) for predicting diet quality in childhood based on panels of urinary metabolites. (h) Interaction of diet quality indicators in association to C-peptide concentration in childhood. (i) Associations of diet quality with C-peptide concentration in childhood after stratifying by sex and by weight status, respectively. (j) Associations between urinary metabolites linked to diet quality and C-peptide in childhood.

• Transparent reporting form

Data availability

Due to the HELIX data policy and data use agreement, human subjects data used in this project cannot be freely shared. Researchers external to the HELIX Consortium who have an interest in using data from this project for reproducibility or in using data held in general in the HELIX data warehouse for research purposes can apply for access to data for a specific manuscript at the time. Interested researchers should fill in the application protocol found in ANNEX I at https://www.projecthelix.eu/files/helix_external_data_request_procedures_final.pdf and send this protocol to helixdata@isglobal.org. The applications are received by the HELIX Coordinator, and are processed and approved by the HELIX Project Executive Committee. The decision to accept or reject a proposal is taken by the HELIX Project Executive Committee. This decision will be based largely on potential overlap with other HELIX-related work, the adequacy of data protection plans, and the adequacy of authorship and acknowledgement plans. After approval by the HELIX PEC, the cohorts participating in HELIX will each be asked if they approve use of their data in the proposal and what their conditions are for participation. Each cohort will have the opportunity to opt out at this stage if any of their cohort-specific conditions are not met. These conditions are to be stated clearly by the cohort at this stage and the cohort should be open to discussing how these conditions could be fulfilled in subsequent proposal revisions. Further details on the content of the data warehouse (data catalog) including those data used for the present projects and procedures for external access are described on the project website http://www.projecthelix.eu/index.php/es/data-inventoryhttp://www.projecthelix.eu/index.php/es/data-inventory. Code used for data analysis has been described with references in the methods section under Statistical analysis. Source data for Figure 1 are provided in Supplementary File 1.

References

Agier L, Basagaña X, Maitre L, Granum B, Bird PK, Casas M, Oftedal B, Wright J, Andrusaityte S, de Castro M, Cequier E, Chatzi L, Donaire-Gonzalez D, Grazuleviciene R, Haug LS, Sakhi AK, Leventakou V, McEachan R, Nieuwenhuijsen M, Petraviciene I, et al. 2019. Early-life exposome and lung function in children in Europe: an analysis of data from the longitudinal, population-based HELIX cohort. *The Lancet. Planetary Health* **3**:e81–e92. DOI: https://doi.org/10.1016/S2542-5196(19)30010-5, PMID: 30737192

Aira T, Vasankari T, Heinonen OJ, Korpelainen R, Kotkajuuri J, Parkkari J, Savonen K, Uusitalo A, Valtonen M, Villberg J, Vähä-Ypyä H, Kokko SP. 2021. Physical activity from adolescence to young adulthood: patterns of change, and their associations with activity domains and sedentary time. *The International Journal of Behavioral Nutrition and Physical Activity* **18**:85. DOI: https://doi.org/10.1186/s12966-021-01130-x, PMID: 34193150

Almanza-Aguilera E, Urpi-Sarda M, Llorach R, Vázquez-Fresno R, Garcia-Aloy M, Carmona F, Sanchez A, Madrid-Gambin F, Estruch R, Corella D, Andres-Lacueva C. 2017. Microbial metabolites are associated with a high adherence to a Mediterranean dietary pattern using a 1H-NMR-based untargeted metabolomics approach. *The Journal of Nutritional Biochemistry* **48**:36–43. DOI: https://doi.org/10.1016/j.jnutbio.2017.06.001, PMID: 28692847

Aranceta J, Pérez-Rodrigo C, Ribas L, Serra-Majem L. 2003. Sociodemographic and lifestyle determinants of food patterns in Spanish children and adolescents: the enKid study. *European Journal of Clinical Nutrition* **57 Suppl 1**:S40–S44. DOI: https://doi.org/10.1038/sj.ejcn.1601813, PMID: 12947451

Asghari G, Yuzbashian E, Mirmiran P, Hooshmand F, Najafi R, Azizi F. 2016. Dietary Approaches to Stop Hypertension (DASH) Dietary Pattern Is Associated with Reduced Incidence of Metabolic Syndrome in Children and Adolescents. *The Journal of Pediatrics* **174**:178-184.. DOI: https://doi.org/10.1016/j.jpeds.2016.03.077, PMID: 27156186

Baker P, Machado P, Santos T, Sievert K, Backholer K, Hadjikakou M, Russell C, Huse O, Bell C, Scrinis G, Worsley A, Friel S, Lawrence M. 2020. Ultra-processed foods and the nutrition transition: Global, regional and national trends, food systems transformations and political economy drivers. *Obesity Reviews* **21**:e13126. DOI: https://doi.org/10.1111/obr.13126, PMID: 32761763

Basagaña X, Pedersen M, Barrera-Gómez J, Gehring U, Giorgis-Allemand L, Hoek G, Stafoggia M, Nieuwenhuijsen MJ, Brunekreef B, Slama R, ESCAPE Birth Outcomes working group. 2018. Analysis of multicentre epidemiological studies: contrasting fixed or random effects modelling and meta-analysis. *International Journal of Epidemiology* **47**:1343–1354. DOI: https://doi.org/10.1093/ije/dyy117, PMID: 29939274

Brosnan JT, Brosnan ME. 2006. Branched-chain amino acids: enzyme and substrate regulation. *The Journal of Nutrition* **136**:207S-11S. DOI: https://doi.org/10.1093/jn/136.1.207S, PMID: 16365084

Buyken AE, Kellerhoff Y, Hahn S, Kroke A, Remer T. 2006. Urinary C-peptide excretion in free-living healthy children is related to dietary carbohydrate intake but not to the dietary glycemic index. *The Journal of Nutrition* **136**:1828–1833. DOI: https://doi.org/10.1093/jn/136.7.1828, PMID: 16772444

Chan She Ping-Delfos WL, Beilin LJ, Oddy WH, Burrows S, Mori TA. 2015. Use of the Dietary Guideline Index to assess cardiometabolic risk in adolescents. *The British Journal of Nutrition* **113**:1741–1752. DOI: https://doi.org/10.1017/S0007114515001026, PMID: 25873301

Chatzi L, Leventakou V, Vafeiadi M, Koutra K, Roumeliotaki T, Chalkiadaki G, Karachaliou M, Daraki V, Kyriklaki A, Kampouri M, Fthenou E, Sarri K, Vassilaki M, Fasoulaki M, Bitsios P, Koutis A, Stephanou EG, Kogevinas M. 2017a. Cohort Profile: The Mother-Child Cohort in Crete, Greece (Rhea Study). *International Journal of Epidemiology* **46**:1392–1393k. DOI: https://doi.org/10.1093/ije/dyx084, PMID: 29040580

Chatzi L, Rifas-Shiman SL, Georgiou V, Joung KE, Koinaki S, Chalkiadaki G, Margioris A, Sarri K, Vassilaki M, Vafeiadi M, Kogevinas M, Mantzoros C, Gillman MW, Oken E. 2017b. Adherence to the Mediterranean diet during pregnancy and offspring adiposity and cardiometabolic traits in childhood. *Pediatric Obesity* **12 Suppl 1**:47–56. DOI: https://doi.org/10.1111/ijpo.12191, PMID: 28160450

Cole TJ, Lobstein T. 2012. Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. *Pediatric Obesity* **7**:284–294. DOI: https://doi.org/10.1111/j.2047-6310.2012.00064.x, PMID: 22715120

Collaborators GBDD. 2019. Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **393**:1958–1972. DOI: https://doi.org/10.1016/S0140-6736(19)30041-8, PMID: 30954305

Collins C, McNamara AE, Brennan L. 2019. Role of metabolomics in identification of biomarkers related to food intake. *The Proceedings of the Nutrition Society* **78**:189–196. DOI: https://doi.org/10.1017/S002966511900048X, PMID: 30767789

Edmands WMB, Beckonert OP, Stella C, Campbell A, Lake BG, Lindon JC, Holmes E, Gooderham NJ. 2011. Identification of human urinary biomarkers of cruciferous vegetable consumption by metabonomic profiling. *Journal of Proteome Research* **10**:4513–4521. DOI: https://doi.org/10.1021/pr200326k, PMID: 21770373

Elizabeth L, Machado P, Zinöcker M, Baker P, Lawrence M. 2020. Ultra-Processed Foods and Health Outcomes: A Narrative Review. *Nutrients* **12**:E1955. DOI: https://doi.org/10.3390/nu12071955, PMID: 32630022

Ellul S, Wake M, Clifford SA, Lange K, Würtz P, Juonala M, Dwyer T, Carlin JB, Burgner DP, Saffery R. 2019. Metabolomics: population epidemiology and concordance in Australian children aged 11-12 years and their parents. *BMJ Open* **9**:106–117. DOI: https://doi.org/10.1136/bmjopen-2017-020900, PMID: 31273021

Fiorito LM, Marini M, Francis LA, Smiciklas-Wright H, Birch LL. 2009. Beverage intake of girls at age 5 y predicts adiposity and weight status in childhood and adolescence. *The American Journal of Clinical Nutrition* **90**:935–942. DOI: https://doi.org/10.3945/ajcn.2009.27623, PMID: 19692492

Garcia-Perez I, Posma JM, Gibson R, Chambers ES, Hansen TH, Vestergaard H, Hansen T, Beckmann M, Pedersen O, Elliott P, Stamler J, Nicholson JK, Draper J, Mathers JC, Holmes E, Frost G. 2017. Objective assessment of dietary patterns by use of metabolic phenotyping: a randomised, controlled, crossover trial. *The Lancet. Diabetes & Endocrinology* **5**:184–195. DOI: https://doi.org/10.1016/S2213-8587(16)30419-3, PMID: 28089709

Garcia-Perez I, Posma JM, Chambers ES, Mathers JC, Draper J, Beckmann M, Nicholson JK, Holmes E, Frost G. 2020. Dietary metabotype modelling predicts individual responses to dietary interventions. *Nature Food* **1**:355–364. DOI: https://doi.org/10.1038/s43016-020-0092-z

Gibson R, Lau C-HE, Loo RL, Ebbels TMD, Chekmeneva E, Dyer AR, Miura K, Ueshima H, Zhao L, Daviglus ML, Stamler J, Van Horn L, Elliott P, Holmes E, Chan Q. 2020. The association of fish consumption and its urinary metabolites with cardiovascular risk factors: the International Study of Macro-/Micronutrients and Blood Pressure (INTERMAP). *The American Journal of Clinical Nutrition* **111**:280–290. DOI: https://doi.org/10.1093/ajcn/nqz293, PMID: 31782492

Grazuleviciene R, Danileviciute A, Nadisauskiene R, Vencloviene J. 2009. Maternal smoking, GSTM1 and GSTT1 polymorphism and susceptibility to adverse pregnancy outcomes. *International Journal of Environmental Research and Public Health* **6**:1282–1297. DOI: https://doi.org/10.3390/ijerph6031282, PMID: 19440446

Grosso G, Mateo A, Rangelov N, Buzeti T, Birt C. 2020. Nutrition in the context of the Sustainable Development Goals. *European Journal of Public Health* **30**:i19–i23. DOI: https://doi.org/10.1093/eurpub/ckaa034, PMID: 32391903

Guertin KA, Moore SC, Sampson JN, Huang W-Y, Xiao Q, Stolzenberg-Solomon RZ, Sinha R, Cross AJ. 2014. Metabolomics in nutritional epidemiology: identifying metabolites associated with diet and quantifying their potential to uncover diet-disease relations in populations. *The American Journal of Clinical Nutrition* **100**:208–217. DOI: https://doi.org/10.3945/ajcn.113.078758, PMID: 24740205

Guxens M, Ballester F, Espada M, Fernández MF, Grimalt JO, Ibarluzea J, Olea N, Rebagliato M, Tardón A, Torrent M, Vioque J, Vrijheid M, Sunyer J, INMA Project. 2012. Cohort Profile: the INMA--INfancia y Medio Ambiente--(Environment and Childhood) Project. *International Journal of Epidemiology* **41**:930–940. DOI: https://doi.org/10.1093/ije/dyr054, PMID: 21471022

Heude B, Forhan A, Slama R, Douhaud L, Bedel S, Saurel-Cubizolles M-J, Hankard R, Thiebaugeorges O, De Agostini M, Annesi-Maesano I, Kaminski M, Charles M-A, EDEN mother-child cohort study group. 2016. Cohort Profile: The EDEN mother-child cohort on the prenatal and early postnatal determinants of child health and development. *International Journal of Epidemiology* **45**:353–363. DOI: https://doi.org/10.1093/ije/dyv151, PMID: 26283636

Hope SV, Knight BA, Shields BM, Hattersley AT, McDonald TJ, Jones AG. 2016. Random non-fasting C-peptide: bringing robust assessment of endogenous insulin secretion to the clinic. *Diabetic Medicine* **33**:1554–1558. DOI: https://doi.org/10.1111/dme.13142, PMID: 27100275

Ioannidis JPA. 2018. The Challenge of Reforming Nutritional Epidemiologic Research. *JAMA* **320**:969–970. DOI: https://doi.org/10.1001/jama.2018.11025, PMID: 30422271

Jakes RW, Day NE, Luben R, Welch A, Bingham S, Mitchell J, Hennings S, Rennie K, Wareham NJ. 2004. Adjusting for energy intake--what measure to use in nutritional epidemiological studies? *International Journal of Epidemiology* **33**:1382–1386. DOI: https://doi.org/10.1093/ije/dyh181, PMID: 15333618

Karatzi K, Moschonis G, Barouti AA, Lionis C, Chrousos GP, Manios Y, Healthy Growth Study Group. 2014. Dietary patterns and breakfast consumption in relation to insulin resistance in children. The Healthy Growth Study. *Public Health Nutrition* **17**:2790–2797. DOI: https://doi.org/10.1017/S1368980013003327, PMID: 24477051

Krupp D, Doberstein N, Shi L, Remer T. 2012. Hippuric acid in 24-hour urine collections is a potential biomarker for fruit and vegetable consumption in healthy children and adolescents. *The Journal of Nutrition* **142**:1314–1320. DOI: https://doi.org/10.3945/jn.112.159319, PMID: 22623390

Kuhn M. 2008. Building Predictive Models in R Using the caret Package. *Journal of Statistical Software* **28**:i05. DOI: https://doi.org/10.18637/jss.v028.i05

Kühn S, Düzel S, Colzato L, Norman K, Gallinat J, Brandmaier AM, Lindenberger U, Widaman KF. 2019. Food for thought: association between dietary tyrosine and cognitive performance in younger and older adults. *Psychological Research* **83**:1097–1106. DOI: https://doi.org/10.1007/s00426-017-0957-4, PMID: 29255945

Lau C-HE, Siskos AP, Maitre L, Robinson O, Athersuch TJ, Want EJ, Urquiza J, Casas M, Vafeiadi M, Roumeliotaki T, McEachan RRC, Azad R, Haug LS, Meltzer HM, Andrusaityte S, Petraviciene I, Grazuleviciene R, Thomsen C, Wright J, Slama R, et al. 2018. Determinants of the urinary and serum metabolome in children from six European populations. *BMC Medicine* **16**:202. DOI: https://doi.org/10.1186/s12916-018-1190-8, PMID: 30404627

Ley SH, Sun Q, Willett WC, Eliassen AH, Wu K, Pan A, Grodstein F, Hu FB. 2014. Associations between red meat intake and biomarkers of inflammation and glucose metabolism in women. *The American Journal of Clinical Nutrition* **99**:352–360. DOI: https://doi.org/10.3945/ajcn.113.075663, PMID: 24284436

Lynch CJ, Adams SH. 2014. Branched-chain amino acids in metabolic signalling and insulin resistance. *Nature Reviews. Endocrinology* **10**:723–736. DOI: https://doi.org/10.1038/nrendo.2014.171, PMID: 25287287

Magnus P, Birke C, Vejrup K, Haugan A, Alsaker E, Daltveit AK, Handal M, Haugen M, Høiseth G, Knudsen GP, Paltiel L, Schreuder P, Tambs K, Vold L, Stoltenberg C. 2016. Cohort Profile Update: The Norwegian Mother

and Child Cohort Study (MoBa). *International Journal of Epidemiology* 45:382–388. DOI: https://doi.org/10.1093/ije/dyw029, PMID: 27063603

Maitre L, Lau CHE, Vizcaino E, Robinson O, Casas M, Siskos AP, Want EJ, Athersuch T, Slama R, Vrijheid M, Keun HC, Coen M. 2017. Assessment of metabolic phenotypic variability in children's urine using 1H NMR spectroscopy. *Scientific Reports* 7:46082. DOI: https://doi.org/10.1038/srep46082, PMID: 28422130

Maitre L, de Bont J, Casas M, Robinson O, Aasvang GM, Agier L, Andrušaitytė S, Ballester F, Basagaña X, Borràs E, Brochot C, Bustamante M, Carracedo A, de Castro M, Dedele A, Donaire-Gonzalez D, Estivill X, Evandt J, Fossati S, Giorgis-Allemand L, et al. 2018. Human Early Life Exposome (HELIX) study: a European population-based exposome cohort. *BMJ Open* 8:e021311. DOI: https://doi.org/10.1136/bmjopen-2017-021311, PMID: 30206078

Manios Y, Kourlaba G, Grammatikaki E, Koubitski A, Siatitsa P-E, Vandorou A, Kyriakou K, Dede V, Moschonis G. 2010. Development of a lifestyle-diet quality index for primary schoolchildren and its relation to insulin resistance: the Healthy Lifestyle-Diet Index. *European Journal of Clinical Nutrition* 64:1399–1406. DOI: https://doi.org/10.1038/ejcn.2010.172, PMID: 20808332

Martinez JA, Keun HC, Siskos AP. 2016. Effects of Mediterranean Diet on the Metabolome. Romagnolo DF, Selmin OI (Eds). *Mediterranean Diet: Dietary Guidelines and Impact on Health and Disease*. Springer International Publishing. p. 121–137. DOI: https://doi.org/10.1007/978-3-319-27969-5, PMID: 26413789

Martos-Moreno GÁ, Mastrangelo A, Barrios V, García A, Chowen JA, Rupérez FJ, Barbas C, Argente J. 2017. Metabolomics allows the discrimination of the pathophysiological relevance of hyperinsulinism in obese prepubertal children. *International Journal of Obesity* 41:1473–1480. DOI: https://doi.org/10.1038/ijo.2017.137, PMID: 28588306

McKeown NM, Dashti HS, Ma J, Haslam DE, Kiefte-de Jong JC, Smith CE, Tanaka T, Graff M, Lemaitre RN, Rybin D, Sonestedt E, Frazier-Wood AC, Mook-Kanamori DO, Li Y, Wang CA, Leermakers ETM, Mikkilä V, Young KL, Mukamal KJ, Cupples LA, et al. 2018. Sugar-sweetened beverage intake associations with fasting glucose and insulin concentrations are not modified by selected genetic variants in a ChREBP-FGF21 pathway: a meta-analysis. *Diabetologia* 61:317–330. DOI: https://doi.org/10.1007/s00125-017-4475-0, PMID: 29098321

Monteiro CA, Moubarac JC, Cannon G, Ng SW, Popkin B. 2013. Ultra-processed products are becoming dominant in the global food system. *Obesity Reviews* 14 Suppl 2:21–28. DOI: https://doi.org/10.1111/obr.12107, PMID: 24102801

Monteiro CA, Cannon G, Moubarac JC, Levy RB, Louzada MLC, Jaime PC. 2018a. The UN Decade of Nutrition, the NOVA food classification and the trouble with ultra-processing. *Public Health Nutrition* 21:5–17. DOI: https://doi.org/10.1017/S1368980017000234, PMID: 28322183

Monteiro CA, Moubarac JC, Levy RB, Canella DS, Louzada M, Cannon G. 2018b. Household availability of ultra-processed foods and obesity in nineteen European countries. *Public Health Nutrition* 21:18–26. DOI: https://doi.org/10.1017/S1368980017001379, PMID: 28714422

Monteiro CA, Cannon G, Levy RB, Moubarac J-C, Louzada ML, Rauber F, Khandpur N, Cediel G, Neri D, Martinez-Steele E, Baraldi LG, Jaime PC. 2019. Ultra-processed foods: what they are and how to identify them. *Public Health Nutrition* 22:936–941. DOI: https://doi.org/10.1017/S1368980018003762, PMID: 30744710

Nie C, He T, Zhang W, Zhang G, Ma X. 2018. Branched Chain Amino Acids: Beyond Nutrition Metabolism. *International Journal of Molecular Sciences* 19:E954. DOI: https://doi.org/10.3390/ijms19040954, PMID: 29570613

O'Gorman A, Brennan L. 2017. The role of metabolomics in determination of new dietary biomarkers. *The Proceedings of the Nutrition Society* 76:295–302. DOI: https://doi.org/10.1017/S0029665116002974, PMID: 28091346

Patel N, Taveira TH, Choudhary G, Whitlatch H, Wu WC. 2012. Fasting serum C-peptide levels predict cardiovascular and overall death in nondiabetic adults. *Journal of the American Heart Association* 1:e003152. DOI: https://doi.org/10.1161/JAHA.112.003152, PMID: 23316320

Patrick H, Nicklas TA. 2005. A review of family and social determinants of children's eating patterns and diet quality. *Journal of the American College of Nutrition* 24:83–92. DOI: https://doi.org/10.1080/07315724.2005.10719448, PMID: 15798074

Pearson N, Braithwaite RE, Biddle SJH, van Sluijs EMF, Atkin AJ. 2014. Associations between sedentary behaviour and physical activity in children and adolescents: a meta-analysis. *Obesity Reviews* 15:666–675. DOI: https://doi.org/10.1111/obr.12188, PMID: 24844784

Perng W, Fernandez C, Peterson KE, Zhang Z, Cantoral A, Sanchez BN, Solano-González M, Téllez-Rojo MM, Baylin A. 2017. Dietary Patterns Exhibit Sex-Specific Associations with Adiposity and Metabolic Risk in a Cross-Sectional Study in Urban Mexican Adolescents. *The Journal of Nutrition* 147:1977–1985. DOI: https://doi.org/10.3945/jn.117.256669, PMID: 28855420

Playdon MC, Sampson JN, Cross AJ, Sinha R, Guertin KA, Moy KA, Rothman N, Irwin ML, Mayne ST, Stolzenberg-Solomon R, Moore SC. 2016. Comparing metabolite profiles of habitual diet in serum and urine. *The American Journal of Clinical Nutrition* 104:776–789. DOI: https://doi.org/10.3945/ajcn.116.135301, PMID: 27510537

Polonsky KS, Licinio-Paixao J, Given BD, Pugh W, Rue P, Galloway J, Karrison T, Frank B. 1986. Use of biosynthetic human C-peptide in the measurement of insulin secretion rates in normal volunteers and type I diabetic patients. *The Journal of Clinical Investigation* 77:98–105. DOI: https://doi.org/10.1172/JCI112308, PMID: 3511094

Posma JM, Garcia-Perez I, Frost G, Aljuraiban GS, Chan Q, Van Horn L, Daviglus M, Stamler J, Holmes E, Elliott P, Nicholson JK. 2020. Nutriome-metabolome relationships provide insights into dietary intake and metabolism. *Nature Food* 1:426–436. DOI: https://doi.org/10.1038/s43016-020-0093-y, PMID: 32954362

Rauschert S, Kirchberg FF, Marchioro L, Koletzko B, Hellmuth C, Uhl O. 2017. Early Programming of Obesity Throughout the Life Course: A Metabolomics Perspective. *Annals of Nutrition & Metabolism* 70:201–209. DOI: https://doi.org/10.1159/000459635, PMID: 28301839

Rebholz CM, Lichtenstein AH, Zheng Z, Appel LJ, Coresh J. 2018. Serum untargeted metabolomic profile of the Dietary Approaches to Stop Hypertension (DASH) dietary pattern. *The American Journal of Clinical Nutrition* 108:243–255. DOI: https://doi.org/10.1093/ajcn/nqy099, PMID: 29917038

Romero-Polvo A, Denova-Gutiérrez E, Rivera-Paredez B, Castañón S, Gallegos-Carrillo K, Halley-Castillo E, Borges G, Flores M, Salmerón J. 2012. Association between dietary patterns and insulin resistance in Mexican children and adolescents. *Annals of Nutrition & Metabolism* 61:142–150. DOI: https://doi.org/10.1159/000341493, PMID: 23037180

Scaglioni S, De Cosmi V, Ciappolino V, Parazzini F, Brambilla P, Agostoni C. 2018. Factors Influencing Children's Eating Behaviours. *Nutrients* 10:E706. DOI: https://doi.org/10.3390/nu10060706, PMID: 29857549

Scalbert A, Brennan L, Manach C, Andres-Lacueva C, Dragsted LO, Draper J, Rappaport SM, van der Hooft JJJ, Wishart DS. 2014. The food metabolome: a window over dietary exposure. *The American Journal of Clinical Nutrition* 99:1286–1308. DOI: https://doi.org/10.3945/ajcn.113.076133, PMID: 24760973

Serra-Majem L, Ribas L, Ngo J, Ortega RM, García A, Pérez-Rodrigo C, Aranceta J. 2004. Food, youth and the Mediterranean diet in Spain. Development of KIDMED, Mediterranean Diet Quality Index in children and adolescents. *Public Health Nutrition* 7:931–935. DOI: https://doi.org/10.1079/phn2004556, PMID: 15482620

Serrano-Sanchez JA, Martí-Trujillo S, Lera-Navarro A, Dorado-García C, González-Henríquez JJ, Sanchís-Moysi J. 2011. Associations between screen time and physical activity among Spanish adolescents. *PLOS ONE* 6:e24453. DOI: https://doi.org/10.1371/journal.pone.0024453, PMID: 21909435

Sri Harsha PSC, Wahab RA, Garcia-Aloy M, Madrid-Gambin F, Estruel-Amades S, Watzl B, Andrés-Lacueva C, Brennan L. 2018. Biomarkers of legume intake in human intervention and observational studies: a systematic review. *Genes & Nutrition* 13:25. DOI: https://doi.org/10.1186/s12263-018-0614-6, PMID: 30214640

Srour B, Fezeu LK, Kesse-Guyot E, Allès B, Debras C, Druesne-Pecollo N, Chazelas E, Deschasaux M, Hercberg S, Galan P, Monteiro CA, Julia C, Touvier M. 2020. Ultraprocessed Food Consumption and Risk of Type 2 Diabetes Among Participants of the NutriNet-Santé Prospective Cohort. *JAMA Internal Medicine* 180:283–291. DOI: https://doi.org/10.1001/jamainternmed.2019.5942, PMID: 31841598

Stanhope KL, Schwarz JM, Havel PJ. 2013. Adverse metabolic effects of dietary fructose: results from the recent epidemiological, clinical, and mechanistic studies. *Current Opinion in Lipidology* 24:198–206. DOI: https://doi.org/10.1097/MOL.0b013e3283613bca, PMID: 23594708

StataCorp. 2021. Stata Multiple-Imputation Reference Manual. College Station, TX: StataCorp LLC.

Sunehag AL, Toffolo G, Treuth MS, Butte NF, Cobelli C, Bier DM, Haymond MW. 2002. Effects of dietary macronutrient content on glucose metabolism in children. *The Journal of Clinical Endocrinology and Metabolism* 87:5168–5178. DOI: https://doi.org/10.1210/jc.2002-020674, PMID: 12414888

Vezzosi D, Bennet A, Fauvel J, Caron P. 2007. Insulin, C-peptide and proinsulin for the biochemical diagnosis of hypoglycaemia related to endogenous hyperinsulinism. *European Journal of Endocrinology* 157:75–83. DOI: https://doi.org/10.1530/EJE-07-0109, PMID: 17609405

WHO. 2003. Diet, nutrition and the prevention of chronic diseases. World Health Organization Technical Report Series.

WHO Study Group. 1990. Diet, nutrition, and the prevention of chronic diseases. Report of a WHO Study Group. *World Health Organization Technical Report Series* 797:1–204 PMID: 2124402.,

Wickham H. 2016. Ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag. DOI: https://doi.org/10.1007/978-3-319-24277-4

World Health Organization. 2015. Guideline: Sugars Intake for Adults and Children. Geneva: WHO.

Wright J, Small N, Raynor P, Tuffnell D, Bhopal R, Cameron N, Fairley L, Lawlor DA, Parslow R, Petherick ES, Pickett KE, Waiblinger D, West J, Born in Bradford Scientific Collaborators Group. 2013. Cohort Profile: the Born in Bradford multi-ethnic family cohort study. *International Journal of Epidemiology* 42:978–991. DOI: https://doi.org/10.1093/ije/dys112, PMID: 23064411

Zhao X, Gang X, Liu Y, Sun C, Han Q, Wang G. 2016. Using Metabolomic Profiles as Biomarkers for Insulin Resistance in Childhood Obesity: A Systematic Review. *Journal of Diabetes Research* 2016:8160545. DOI: https://doi.org/10.1155/2016/8160545, PMID: 27517054