Nutrient scoring for the DEGS1-FFQ – from food intake to nutrient intake

Ronja Thieleking
Max Planck Institute for Human Cognitive and Brain Sciences
Lennard Schneidewind
Max Planck Institute for Human Cognitive and Brain Sciences
Arsene Kanyamibwa
Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland
Hendrik Hartmann
Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland
Annette Horstmann
Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland
A. Veronica Witte
Clinic for Cognitive Neurology, University Medical Center Leipzig, Leipzig, Germany
Evelyn Medawar (medawar@cbs.mpg.de)
Max Planck Institute for Human Cognitive and Brain Sciences

Research Article

Keywords: dietary assessment, nutrient scoring, fiber intake, FFQ

Posted Date: May 10th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1525712/v2

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Nutrition on January 13th, 2023. See the published version at https://doi.org/10.1186/s40795-022-00636-2.

Abstract

Background: While necessary for studying dietary decision-making or public health, estimates of nutrient supply based on self-reported food intake are barely accessible or fully lacking and remain a challenge in human research. In particular, detailed information on dietary fiber is limited. In this study we introduce an automated openly available approach to assess self-reported nutrient intake for research purposes for a popular, validated German food frequency questionnaire (FFQ).

Methods: To this end, we i) developed and shared a code for assessing nutrients (carbohydrates, fat, protein, sugar, fiber...) for 53 items of the quantitative, validated German DEGS1-FFQ questionnaire implementing expert-guided nutritional values of diverse sources with several raters. In a sample of individuals ($n_{GUT\text{-}BRAIN}$ = 61 (21 female) overweight, omnivorous), we ii) cross-validated nutrient intake of the last 7 days and the last 24 hours and iii) computed test-retest reliability across two timepoints. Further, iv) we report newly computed nutrient intake for two independent cross-sectional cohorts with continuous weight status and different dietary habits (n_{Mensa}= 134 (79 female, 1 diverse), n_{GREADT} = 76 male). Exploratively, we correlated computed nutrient intake with v) anthropometric and vi) blood-based biomarkers.

Results: In overweight adults (n= 61 (21 female), mean age 28.2±6.5 years, BMI 27.4±1.6 kg/m^2) nutrient intakes were mostly normally distributed and within or surpassing recommended reference nutrient ranges for both last 7 days and last 24 hours. Reliability between last 7 days and 24 hours per visit was moderate (Pearson's $r_{all} \geq 0.34$, $p_{all} < 0.001$, $r_{max} = 0.54$) and absolute agreement across two timepoints was moderate for 7 days ($kappa_{all} > 0.40$, $p_{all} < 0.001$) and poor for 24 hours ($kappa_{all} > 0.08$, $p_{all} < 0.001$). Associations of dietary components to anthropometric markers showed distinct sex differences, with overall higher intake by males compared to females and opposite associations of fiber intake and BMI in males compared to females. Links between nutrient intake relative to calorie intake and anthropometrics as well as serum markers remain inconclusive.

Conclusion: We provide an openly available tool to systematically assess nutrient intake, including fiber, based on self-report by a common German FFQ. The computed nutrient scores resembled overall plausible and reliable measures of nutrient intake given the known limitations of FFQs regarding over- or underreporting. Our open code nutrient scoring can help to examine dietary intake in experimental studies, including dietary fiber and its subclasses, and can be readily adapted to other FFQs. Further validation of computed nutrients with biomarkers and nutrient-specific metabolites in serum, urine or feces will help to interpret self-reported dietary intake.

Introduction

Benefits and drawbacks of dietary habit assessment

Nutrition science relies on tracking dietary intake of individuals using more or less sophisticated self-reported dietary diaries, energy chambers or other observational measures [1]. Advantages of commonly used food frequency questionnaires (FFQ) based on self-report are low costs, low time investment and the possibility of self-administration. Disadvantages are non-uniformity across studies (due to differences in number and variety of food items, time frame of food intake), self-report of non-expert study participants leading to under-/over-/misreporting of food intake and lack of detail on specific food items [2]. In particular, in real-life settings self-reported measures are criticized for being too imprecise to carry valuable evidence, especially for providing robust data for nutritional epidemiological research and dietary recommendations for society [3], however methodological improvements such as ecological momentary assessment might overcome some of these limitations [4]. Self-administered questionnaires save valuable interviewing resources, but usually require more preparation time and pre-testing than an interview-administered FFQ. Participants may only report commonly eaten items or miss some of the questions, therefore checking for completeness and plausibility is necessary [5]. Computer-based methods for the recall of dietary intake offers the possibility to check the answers automatically and to implement a variety of quality controls to assist the overall standardization and accuracy of the collected data [6, 7]. In addition, data entry errors and time effort are reduced.

Nutrient scoring for FFQs with a focus on fiber

Depending on the research question, dietary patterns or certain aspects of the diet, i.e. single macro- or micronutrients, may be of important to assess reliably. Yet, extracting macro- and micronutrient levels from dietary data requires either using nutrient reference databases and developing a scoring method for the FFQ at hand or feeding the data manually into commercial software, resulting in high effort and error-prone methodology. Data on the nutrient level is crucial for assessing associations or effects of nutrient intake on health or behaviour. For example, dietary fiber is known to be a beneficial dietary component related to better health status [8], lower all-cause mortality [9], colorectal cancer [10], inflammatory bowel disease [11], and depression [12]. A systematic review of 185 prospective studies and 58 clinical trials with 4,635 adult participants suggested highest risk reduction with a daily fiber intake of 25 g to 29 g, in a dose-dependent response when combining dietary fiber and whole-grain foods [8]. Plant-based (vegetarian, vegan) diets are estimated to be on average higher in dietary fiber compared to animal-based diets [13, 14]. Strict plant-based diets have been shown to reach those beneficial fiber intake ranges [15], whereas gradual increases in fiber were shown depending on dietary adherence (meat-eaters, fish-eaters, vegetarians, vegans) [16]. Moreover, diets high in fat and sugar are in parallel likely to be low in fiber [17]. Indeed, measuring actual dietary fiber intake is difficult due to the definition of substances that fall under this category and how these can be accurately measured [18]. Dietary fiber can be defined as “non-digestible carbohydrates and lignin that are intrinsic and intact in plants” [10], whereas other sources define dietary fiber as non-digestible plant polysaccharides [19]. There are also different approaches to the categorization of dietary fiber, since it can be classified by its source [10, 20] or by its subtypes [20]. Most commonly, for fiber-specific FFQs, such as the EAT5 FFQ [21], dietary fiber intake FFQ (DFI-FFQ) [19] or others [22, 23], fiber is measured in gram per day [15, 19], or fiber intake relative to total energy intake, i.e. gram per 1000 kcal per day [10]. When analyzing a larger data set, it is reasonable to split the sample

in quintiles based on fiber intake [10, 15], whereas smaller datasets are commonly split in tertiles or at the median. In addition, fiber intake was shown to be gender-specific [19]. More recent questionnaire development efforts also include the distinction of soluble, insoluble and prebiotic dietary fibers (FiberTAG, [24]). Overall, fiber-specific FFQs and fiber scoring could serve as a cost-efficient and quick tool to detect insufficient intake. Yet, they remain underdeveloped and niche due to low accuracy and low validity.

Recall bias in FFQ data

Furthermore, FFQs are subject to random measurement errors due to recall bias caused by misreporting of food intake. The degree of subjective bias is linked to the characteristics of the respondent. Reliable predictors for potential underreporting of food intake are age, BMI, and level of education. Levels of underreporting were 31% in the Second National Health and Nutrition Examination Survey and 46% for women and 29% for men in the National Diet and Nutrition Survey of British Adults [5]. Also, individuals with higher BMI underreported to a higher degree than those with lower BMI [25]. Another issue limiting the validity of FFQs is the use of food consumption databases to calculate the energy and nutrient intake according to the reported food. These databases tend to ignore natural variation of food, often contain limited information regarding the composition of foods and may contain food groups with mixed qualities [26, 27]. In contrast to an interview-administered FFQ ideally administered by a nutrition expert, a self-administered one uses closed-ended questions when asking for the consumption frequency and quantity. This way, the coding time and transcription error is minimized and the answers to the questions are straightforward to assess [5]. Since closed-ended questions ask for standardized portions and frequencies, some details of dietary intake may not be captured. Many food items are part of mixed dishes, therefore the respondent has to quantify the consumption of a food from single foods and mixed dishes at once. Further, when the FFQ asks for the daily intake of milk, ideally it should also ask for the consumption as plain beverage and as add-on for cereals and coffee [26].

Selecting and designing FFQs for the aim of the study

Using FFQs to assess dietary intake thus comes with several considerations to take. Before designing or selecting an FFQ for a study protocol, its purpose should be defined. For example, the FFQ of an intervention study using dietary intake as a measure of interest or compliance needs to be sensitive and specific enough to detect the sometimes subtle changes in dietary intake prompted by. When preexisting questionnaires are modified, the original purpose, target population, year of application and previous validations have to be evaluated. If the initial FFQ has been designed several years ago, it may not cover today’s common food items. Moreover, cultural translatability may be limited, for instance if the original questionnaire has been developed for a British population, it may not be suited for an Indian population. The aim of the study also influences the characteristics of the food list used in the FFQ. In some cases a detailed list with single items can be preferable, but it may cause an overestimation of food intake, whereas a shorter list consisting mainly of food groups can underestimate food intake and variation. The order of the presented food items or groups should also aim to reduce cognitive load and recall bias of

the participant. Hence linked items should be grouped together, for example muesli and milk or bread and butter.

The most appropriate way to validate FFQs seems to combine dietary diary records and biomarkers [7]. Since biochemical measurements of nutrients provide unbiased estimates of dietary intake, they are not subject to recall bias. However, they are limited to certain nutrients and have inherent error sources which are linked to the biochemical assays themselves and the individual characteristics and metabolism of the participants. Most biomarkers do not allow assessment of true absolute dietary intake [5].

Despite being a validated and widely used tool for assessing dietary intake (e.g. Esche et al. 2018; Paprott et al. 2016; Scheidt-Nave et al. 2012; Huhn et al. 2018), the semi-quantitative DEGS1-FFQ is missing an automated nutrient scoring methodology for making nutrient intake assessment in Germany more feasible.

Aim

The commonly used German DEGS1-FFQ which was first used within the DEGS1 study [26, 32] by the Robert Koch Institute (Berlin, Germany) is a tool to measure the approximate intake of 53 single food items consumed based on self-report of frequency and quantity. The main outcome is mean daily portion in grams for each of the 53 items. However, this measure is of limited use due to its numerous outcome variables and not suitable for investigating more specific intake of macro- and micronutrients and dietary fiber. The aim is to translate self-reported dietary intake (using the German DEGS1-FFQ) into nutrient intake per day for various macro- and micronutrients and other nutrients. We further assessed test-retest reliability for computed nutrient intake and assayed potential relations to anthropometric measures and biomarkers.

Methods

Study sample

The herein analyzed data has been taken from a within-subject cross-over design in a human dietary intervention study (title: GUT-BRAIN, registered under; NCT03829189). Data for the validation of the nutrient scoring includes two baseline assessments (first baseline and second baseline after wash-out period). The main sample consists of $n_{GUT-BRAIN}$ = 61 (21F) omnivorous participants with a mean age of 28.2 ± 6.5 years and an average BMI of 27.4 ± 1.6 kg/m^2 (range: 25–31 kg/m^2). Participants gave their written informed consent before taking part in the study and were compensated with 9€ per hour.

DEGS1-Food Frequency Questionnaire

Food intake was recorded using the validated German DEGS1-FFQ [26], adapted from the last 28 days to different time periods: a) the last 7 days (FFQ7d) and b) the last 24 hours (FFQ24h) (n_{FFQ7d} = n_{FFQ24h} = 61 individuals, 110 data points). With the DEGS1-FFQ, the consumption of 53 food items over a certain period is assessed. This period can vary from 4 weeks down to 24 hours. The questionnaire gathers

information about the frequency and amount of these 53 food items. The original scoring of the DEGS1-FFQ results in the mean daily portion of each food item in grams.

Participants in the GUT-BRAIN study completed both questionnaires at two visits respectively. Visits were at least 28 days apart. Each participant filled out the questionnaire online via browser-based LimeSurvey v3.0. The questionnaire as well as the .lsq-file for LimeSurvey can be accessed via gitlab.gwdg.de. The FFQ24h was filled out by the participants at home just before the visit and the FFQ7d directly at the research institute.

Additional data samples

For cross-validation, we extended the data by two independent samples. The first sample, titled “Mensa”, consisted of n_{Mensa} = 134 (79 female, 43 male, 1 diverse, 11 NA) German university cafeteria visitors with a mean BMI of 22.5 ± 3.1 kg/m^2 (range: 17.5–40.6 kg/m^2). The study was an observational online study investigating post-meal ratings of hunger and well-being. Participants in the “Mensa” study were omnivorous dieters only. The DEGS1-FFQ in the “Mensa” study covered a time frame of 14 days dietary intake [33]. Secondly, in a cross-sectional sample of adult men, including omnivorous and vegetarian dieters, dietary intake was assessed with the DEGS1-FFQ for the last 4 weeks (n_{GREADT} = 76M) [34]. Anthropometric data (BMI: 23.6 ± 2.7 kg/m^2, range: 18.6–36.4 kg/m^2, age: 26.6 ± 4.4 years, range: 18–40) and serum markers (glucose, HbA1c, lipids, inflammatory markers, general health markers) are available. The “GREADT” study was designed with two groups with significantly different dietary fat and sugar intake as measured by the 26-item German Version of the Dietary Fat and Free Sugar-Short Questionnaire (DFS) [35].

Nutrient database

All reference nutrient values were extracted from the German Nutrient Reference Database “Bundeslebensmittelschlüssel” (BLS, version 3.02, Max Rubner-Institut, Karlsruhe, Germany) or, in rare cases, from individual sources directly from food suppliers. The BLS is a database comprising extensive tables on food composition (macro- and micronutrients) of single food items.

DEGS1-FFQ dietary scoring

As mentioned above, the original scoring of the DEGS1-FFQ (Haftenberger et al. 2010) provides the mean daily portion for each of the 53 food items in grams. The calculation of the mean daily portions is based on the amount and frequency that were indicated for each food item.

However, for 16 of the 53 food items, participants give more specific information which the original scoring disregards. These “type” questions provide information on the way a food item was processed (cooked or fried), if a food item was high or low in fat or if a certain drink was consumed undiluted or diluted. Our aim was to incorporate these details to obtain more precise information on the participants food intake. Therefore, we evaluated the answers of the “type” questions in order to correctly calculate the macro- and micronutrients of the daily food intake.

In order to convert the daily food intake into nutrient values, we created a reference table. This reference table includes nutrient values for all 53 food items as well as for the variations of the food items addressed by the “type” questions. The macronutrient values cover energy (kcal/100g), protein, fat, carbohydrates, dietary fiber and overall sugar as well as the fiber subclasses cellulose, lignin, water-solulable and -insolulable fiber. Micronutrients include tyrosine, tryptophan, saturated fatty acids, short-, medium- and long-chain fatty acids as well as Omega-3 and Omega-6 fatty acids. All nutrient values except for energy are provided in mg/100g.

In order to comprehend how we created the reference table, we published the process here: https://gitlab.gwdg.de/omega-lab/ffq-nutrient-scoring. There, you will also find the LimeSurvey questionnaire file as well as the R-code to conduct the nutrient scoring.

The steps from the raw DEGS1-FFQ data to macro- and micronutrients consumed per day, are the following:

- calculation of the mean daily portions for each of the 53 food items based on the original scoring
- identifying information on the different types of preparation, food types, or ways of consumption specified in the 16 “type” questions
- calculation of the macro- and micronutrients for each food item by combining all acquired information on the 53 food items
- summing up all nutrients to acquire total macro- and micronutrient intake per day

Biomarker analysis

Blood was obtained in fasting state (12.5 ± 2.2h fasted) from every participant of the GUT-BRAIN study at the same time for each session (glucose: S-Monovette 2.7ml FE; HbA1c: S-Monovette 2.7ml K3E; all other: S-Monovette 9ml Z-Gel). Blood samples were centrifuged at 3500 rpm at 7°C for 6 minutes and serum was aliquoted within 1h of obtainment. Processed aliquots were stored at -80°C within 1h of collection and further analyzed in one batch per marker. Analyses were conducted at Synevo Studien Service Labor GmbH c/o IMD Institut für Medizinische Diagnostik Berlin-Potsdam GbR, Berlin, Germany and the Institute for Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics (ILM) Leipzig University, Leipzig, Germany.

Anthropometric markers

All measurements were taken at each testing day of the GUT-BRAIN study. Body-mass-index (BMI) was measured weighing participants in light clothes and no shoes in a fasted state and measuring height against the wall, calculating weight(kg)/height(m)^2. Waist-to-hip-ratio (WHR) was assessed measuring waist circumference between the lower costal arch and the iliac crest, hip circumference was measured above the head of the greater trochanter (thigh bone) using a tape measure. Body fat mass percentage was derived using hand-to-foot bioimpedance analysis with BIACORPUS RX 4004M (Medi Cal Healthcare GmbH, Karlsruhe, Germany). Four gel electrodes were placed over the ipsilateral wrist and ankle Blood

pressure was measured three times after 10 minutes of resting in lying position with OMROM M500 HEM-7213-D (OMRON HEALTHCARE Co., Ltd., Kyoto, Japan). The mean values for each diastolic and systolic phase were calculated.

Statistical analysis

All analyses were performed with R (version 3.6.1). Figures were created with the packages ggplot2 (version 3.3.0) and corrplot (version 0.9.0).

Main analysis

Intra-variability of two FFQs on different timescales, i.e. FFQ7d and FFQ24h, was assessed using Pearson's correlation coefficients (after checking for normal distribution) with a significance level of $\alpha = 0.05$ across all participants and both timepoints (Fig. 1). Moreover, reliability of nutrient scoring across two timepoints for identical nutrient outcomes was assessed using Intraclass Correlation Coefficient (ICC) with ICC() function in psych-package (version 2.1.9) for both FFQ 7 days and 24 hours using two-way random, single measures for absolute agreement (ICC2) and guidelines for interpretation according to [36]. To test correlations between nutrient intake and biomarkers Pearson's correlation matrices were created using corrplot-package (version 0.90).

Exploratory analysis

We used correlations (Spearman's rank correlation coefficient if not stated otherwise, $\alpha = 0.05$) to test whether main dietary intake variables (based on DEGS1-FFQ data) correlate with anthropometrics (BMI, body fat, WHR) and blood pressure as a measure of cardiovascular risk (Fig. 1). Further, linear mixed models (lme4-package, version 1.1.27.1.) were used to account for multiple datapoints from the same subject as well as for age, sex, and other confounding factors.

Results

Computed nutrient intake

For food intake over the last 7 days as the sum of all 53 food group items, computed mean caloric intake matched reference values of 2000 kcal quite well, as well as protein (0.8 g/kg body weight/d based on German Society for Nutrition (DGE) guidelines) and fat intake (30% of overall intake based on DGE, for 2400 kcal 80–80 g/d) (Table 1, Fig. 2). Sugar intake in our sample was almost twice as high than the recommended 50 g/d (based on German Nutrition Counselling Network (DEBI)). Carbohydrate intake was slightly lower and fiber intake almost 50% lower than recommended (minimum of 30 g/d based on DGE). Nutrient intake for 7 days FFQ was overall higher in males than in females. Regarding the last 24 hours, computed nutrient intake was on average higher than for the last 7 days, yet relative intake of nutrients reflected reference nutrient values (Table 2, Fig. 2), except for sugar which was 300% of the recommended intake and fiber which was again almost 50% lower than recommended. Although nutrient

levels were mostly close to recommended levels on average, there was a large inter-individual variability for all nutrients for both 7 days and 24 hours FFQ, but this was more pronounced for 24 hours FFQ data.

Table 1
Nutrient intake descriptives based on FFQ 7 days (n = 61). F: female, M: male, BL1/BL2: baseline visits.

FFQ 7 days				
	BL1		BL2	
	F	M	F	M
	(n = 21)	(n = 40)	(n = 13)	(n = 36)
Energy [kcal]				
Mean (SD)	1490 (558)	1730 (516)	1500 (466)	1740 (488)
Median [Min, Max]	1460 [501, 3100]	1800 [628, 2730]	1300 [919, 2350]	1770 [830, 2780]
Protein [g]				
Mean (SD)	51.9 (22.4)	69.9 (26.1)	58.6 (23.1)	68.1 (21.9)
Median [Min, Max]	47.4 [7.85, 89.8]	66.7 [16.7, 140]	47.5 [33.0, 106]	66.1 [27.0, 130]
Fat [g]				
Mean (SD)	61.9 (33.1)	97.4 (54.4)	76.2 (32.1)	93.1 (39.3)
Median [Min, Max]	58.4 [4.54, 126]	79.2 [17.6, 281]	68.6 [24.0, 140]	91.9 [25.4, 225]
Sugar [g]				
Mean (SD)	102 (41.3)	92.0 (50.7)	86.0 (30.2)	78.8 (37.7)
Median [Min, Max]	98.0 [46.6, 223]	86.6 [25.9, 290]	78.7 [46.7, 160]	73.7 [19.1, 180]
Carbohydrates [g]				
Mean (SD)	196 (82.8)	210 (70.9)	182 (56.6)	204 (64.3)
Median [Min, Max]	187 [76.8, 450]	214 [80.3, 398]	181 [112, 297]	194 [91.0, 339]
Fiber [g]				
Mean (SD)	15.8 (6.87)	16.9 (6.11)	15.9 (9.32)	16.0 (6.78)
Median [Min, Max]	16.0 [1.54, 27.9]	15.5 [6.23, 30.5]	13.7 [6.34, 40.0]	15.2 [4.16, 32.7]
Sat. FA [mg]				
Mean (SD)	20200 (10600)	26100 (9840)	22800 (9950)	26800 (8380)

FFQ 7 days				
Median [Min, Max]	20900 [2530, 45200]	24700 [8150, 51800]	18800 [8890, 43200]	26700 [13100, 45100]
Tyrosine [mg]				
Mean (SD)	1870 (856)	2570 (1090)	2110 (859)	2440 (812)
Median [Min, Max]	1860 [228, 3200]	2320 [561, 5970]	1700 [1200, 3820]	2350 [1020, 4770]
Tryptophan [mg]				
Mean (SD)	610 (259)	831 (309)	696 (266)	810 (259)
Median [Min, Max]	585 [85.0, 1060]	813 [199, 1650]	575 [393, 1210]	790 [291, 1560]
Omega-3 [mg]				
Mean (SD)	1130 (768)	1230 (559)	1720 (1540)	1430 (710)
Median [Min, Max]	1080 [39.1, 3600]	1050 [529, 2670]	1210 [418, 6190]	1240 [342, 3000]
Omega-6 [mg]				
Mean (SD)	7110 (3740)	8400 (2870)	6400 (2320)	9060 (3690)
Median [Min, Max]	7590 [343, 14500]	8080 [2460, 14700]	6170 [3390, 10400]	9570 [3260, 17100]

Table 2
Nutrient intake descriptives based on FFQ 24 hours (n = 61). F: female, M: male, BL1/BL2: baseline visits.

FFQ 24 hours				
	BL1		**BL2**	
	F	**M**	**F**	**M**
	(n = 21)	**(n = 40)**	**(n = 13)**	**(n = 36)**
Energy [kcal]				
Mean (SD)	2230 (1050)	3120 (1760)	2020 (1070)	2570 (1030)
Median [Min, Max]	2310 [421, 4110]	2870 [729, 10600]	1990 [597, 4530]	2450 [844, 5410]
Protein [g]				
Mean (SD)	83.7 (56.0)	121 (82.2)	74.7 (37.8)	97.0 (46.5)
Median [Min, Max]	76.1 [10.3, 259]	102 [27.6, 440]	65.5 [16.0, 147]	92.7 [28.4, 245]
Fat [g]				
Mean (SD)	120 (142)	177 (152)	107 (88.3)	120 (91.3)
Median [Min, Max]	68.9 [8.61, 621]	121 [35.5, 656]	62.7 [9.75, 253]	95.5 [13.4, 411]
Sugar [g]				
Mean (SD)	165 (121)	190 (154)	118 (85.8)	121 (72.9)
Median [Min, Max]	150 [12.4, 533]	158 [28.3, 803]	107 [27.5, 356]	99.5 [23.3, 357]
Carbohydrates [g]				
Mean (SD)	291 (165)	378 (222)	244 (127)	295 (116)
Median [Min, Max]	261 [19.9, 701]	326 [55.0, 1230]	218 [76.4, 541]	280 [125, 598]
Fiber [g]				
Mean (SD)	22.2 (18.3)	25.4 (18.6)	17.6 (9.29)	21.7 (9.89)
Median [Min, Max]	17.8 [1.92, 88.1]	19.8 [2.10, 101]	15.2 [3.50, 37.9]	19.2 [6.23, 45.5]
Sat. FA [mg]				
Mean (SD)	29600 (16600)	48600 (32600)	32500 (22400)	39900 (19600)

FFQ 24 hours				
Median [Min, Max]	30500 [4930, 66400]	37800 [14100, 160000]	23500 [5010, 82400]	36700 [7100, 101000]
Tyrosine [mg]				
Mean (SD)	2980 (2050)	4510 (3130)	2690 (1380)	3480 (1700)
Median [Min, Max]	2690 [306, 8990]	3770 [1050, 16100]	2190 [421, 5070]	3270 [869, 8910]
Tryptophan [mg]				
Mean (SD)	980 (661)	1470 (991)	893 (471)	1170 (592)
Median [Min, Max]	891 [121, 3130]	1280 [369, 5250]	745 [174, 1760]	1100 [289, 3010]
Omega-3 [mg]				
Mean (SD)	1850 (2180)	2590 (2860)	1890 (2000)	1980 (1500)
Median [Min, Max]	1330 [62.5, 8440]	1470 [77.9, 13800]	922 [106, 7220]	1710 [234, 7440]
Omega-6 [mg]				
Mean (SD)	11300 (10100)	14100 (11700)	8290 (7500)	12400 (8360)
Median [Min, Max]	8650 [518, 42600]	11700 [1850, 62500]	5880 [1040, 29700]	9690 [1330, 37400]

Reliability across two timepoints of assessment of nutrient intake

The intra-class correlation coefficient (ICC) was computed to assess the agreement between two assessments with the same instrument, i.e. test-retest reliability, of dietary intake in 61 individuals. For FFQ7d, there was moderate absolute agreement between the two data assessments, for all macronutrients ($kappa_{all} \geq 0.40$, $kappa_{max} = 0.73$, all $p < 0.001$), with highest agreement for protein and lowest for fat intake. For FFQ24h, agreement between the two timepoints was poor overall ($kappa_{all} \geq 0.08$, $kappa_{max} = 0.34$), $p_{all} < 0.02$), and highest for carbohydrates and lowest for fiber intake. Intra- compared to inter-individual variability of nutrient intakes was found to be lower overall and both were higher for FFQ24h than for FFQ7d (**SI** Fig. 1 **+ 2**).

Reliability between 7 days and 24 hours FFQ nutrient intake

In general higher nutrient scores at 7d correlated with higher nutrient scores at 24h, respectively, resembling that individuals who reported high nutrient intake on the FFQ7d also reported high nutrient

intake on the FFQ24h. Intra-variability between 7 days and last 24 hours on the individual's level for all macro- and micronutrients was moderate (Pearson's $r_{all} \geq 0.34$, $p_{all} < 0.001$, $r_{max} = 0.54$, Fig. 3).

Extension of nutrient scoring by two independent samples

An additional cross-sectional sample from German university cafeteria visitors showed similar nutrient values and deviations from the reference values (Table 3). Calorie intake was lower than expected reference values, protein and fat matched reference ranges well, sugar intake was 60% higher than recommended. Carbohydrate was slightly lower and fiber intake 30% lower than recommended.

Table 3
Computed nutrient intake for a cross-sectional sample of university cafeteria visitors for FFQ 14 days (n = 134).

DEGS1-FFQ 14 days nutrient intake by gender				
	female	male	diverse	overall
	(n = 79)	(n = 43)	(n = 1)	(n = 134 incl. 11 NA's for gender)
BMI (kg/m^2)				
Mean (SD)	22.6 (3.40)	22.5 (2.37)	22.0 (NA)	22.5 (3.05)
Median [Min, Max]	22.0 [18.0, 40.6]	22.7 [17.5, 27.2]	22.0 [22.0, 22.0]	22.3 [17.5, 40.6]
Energy [kcal]				
Mean (SD)	1450 (572)	1640 (597)	1510 (NA)	1500 (577)
Median [Min, Max]	1360 [350, 3070]	1530 [758, 3040]	1510 [1510, 1510]	1400 [350, 3070]
Protein [g]				
Mean (SD)	51.9 (22.2)	68.5 (31.4)	59.0 (NA)	56.9 (26.9)
Median [Min, Max]	48.1 [10.9, 161]	58.3 [28.2, 154]	59.0 [59.0, 59.0]	51.2 [10.9, 161]
Fat [g]				
Mean (SD)	70.2 (44.5)	89.9 (51.0)	35.7 (NA)	74.3 (46.6)
Median [Min, Max]	54.8 [12.7, 219]	87.0 [21.9, 293]	35.7 [35.7, 35.7]	60.1 [12.7, 293]
Sugar [g]				
Mean (SD)	89.5 (48.4)	80.3 (58.0)	57.2 (NA)	84.4 (50.5)
Median [Min, Max]	85.4 [18.2, 341]	69.5 [23.6, 360]	57.2 [57.2, 57.2]	77.0 [18.2, 360]
Carbohydrates [g]				
Mean (SD)	181 (68.7)	198 (78.5)	175 (NA)	187 (71.7)
Median [Min, Max]	177 [44.2, 404]	197 [93.7, 472]	175 [175, 175]	185 [44.2, 472]
Fiber [g]				

DEGS1-FFQ 14 days nutrient intake by gender				
Mean (SD)	19.1 (10.1)	18.2 (9.97)	22.1 (NA)	18.7 (9.83)
Median [Min, Max]	17.6 [4.11, 60.5]	15.1 [6.08, 46.9]	22.1 [22.1, 22.1]	17.0 [4.11, 60.5]
Cellulose [mg]				
Mean (SD)	22400 (17100)	22800 (11400)	12200 (NA)	21800 (14800)
Median [Min, Max]	18000 [4380, 97900]	20800 [8060, 56600]	12200 [12200, 12200]	17500 [4380, 97900]
Lignin [mg]				
Mean (SD)	1880 (882)	2490 (1270)	2170 (NA)	2060 (1080)
Median [Min, Max]	1760 [370, 6510]	2080 [949, 6780]	2170 [2170, 2170]	1830 [370, 6780]
Soluble Fiber [mg]				
Mean (SD)	606 (255)	817 (380)	721 (NA)	672 (320)
Median [Min, Max]	574 [120, 1840]	700 [313, 1820]	721 [721, 721]	597 [120, 1840]
Insoluble Fiber [mg]				
Mean (SD)	1400 (3410)	3510 (13500)	1220 (NA)	2040 (8100)
Median [Min, Max]	708 [332, 30400]	1410 [388, 89900]	1220 [1220, 1220]	874 [213, 89900]
Sat. FA [mg]				
Mean (SD)	7810 (7060)	8300 (3310)	7280 (NA)	7790 (5780)
Median [Min, Max]	6040 [2380, 41300]	7660 [1840, 16400]	7280 [7280, 7280]	6610 [1820, 41300]

Another cross-sectional sample including men only (n = 76) with a mean BMI of 23.6 ± 2.7 kg/m^2 (range: 18.6–36.4) and mean age of 26.6 ± 4.4 years (range: 18–40) showed similar nutrient values and deviations from the reference values for an FFQ recall period of four weeks (Table 4). The sample was further grouped into high and low dietary fat and sugar consumers (HFS vs. LFS) based on DFS scores. Overall, calorie intake, protein, and fat matched reference ranges well. Fiber was slightly lower than recommended (< 30g) for both groups. Groups differed significantly in overall calorie intake, protein, fat, sugar, carbohydrates, and many other micronutrients, with higher intake in HFS vs. LFS group. Notably, both groups surpassed recommended intake of sugar (about 1.5-2.5x higher than recommended), yet the HFS group did so by far. Serum markers were only significantly different for HbA1c, with higher levels in

the HFS group (HFS: 5.21 ± 0.2%; LFS: 5.08 ± 0.3%, p = 0.03), a long-term marker of glucose metabolism (**SI** Table 1).

Table 4

Computed nutrient intake for a cross-sectional sample of adult men for FFQ 30 days. HFS = High Fat and Sugar Group, LFS = Low Fat and Sugar Group. P-values are indicated for standard 2-sample t.test for numeric variables and for chi-squared tests of independence for categorical variables.

DEGS FFQ Monthly Nutrient Intake by DFS Group			
	HFS	**LFS**	
	(n = 35)	**(n = 41)**	
factor(Diet)			p-value
OMN	32 (91.4%)	27 (65.9%)	**0.0168**
VEG	3 (8.6%)	14 (34.1%)	
mean daily portion [g]			
Mean (SD)	4580 (1380)	4300 (1730)	0.437
Median [Min, Max]	4500 [2210, 8260]	4030 [1280, 8370]	
Energy [kcal]			
Mean (SD)	2530 (650)	1710 (694)	**< 0.001**
Median [Min, Max]	2430 [1320, 4050]	1620 [666, 3540]	
Protein [g]			
Mean (SD)	95.4 (26.3)	75.6 (42.5)	**0.0154**
Median [Min, Max]	92.9 [43.2, 156]	62.7 [23.0, 213]	
Fat [g]			
Mean (SD)	120 (46.6)	74.5 (41.6)	**< 0.001**
Median [Min, Max]	117 [48.3, 245]	65.7 [19.2, 210]	
Sugar [g]			
Mean (SD)	130 (60.0)	85.8 (39.5)	**< 0.001**
Median [Min, Max]	112 [47.5, 342]	82.6 [25.2, 171]	
Carbohydrates [g]			
Mean (SD)	306 (86.2)	204 (87.3)	**< 0.001**
Median [Min, Max]	294 [181, 584]	201 [46.2, 461]	
Fiber [g]			
Mean (SD)	27.8 (10.4)	25.5 (13.0)	0.392

DEGS FFQ Monthly Nutrient Intake by DFS Group			
Median [Min, Max]	25.0 [11.1, 56.4]	22.2 [8.69, 62.1]	
Cellulose [mg]			
Mean (SD)	5040 (2320)	4850 (2830)	0.742
Median [Min, Max]	4520 [2290, 12300]	3860 [1010, 13100]	
Lignin [mg]			
Mean (SD)	1320 (640)	1430 (952)	0.532
Median [Min, Max]	1230 [318, 2840]	1130 [194, 5320]	
Soluble Fiber [mg]			
Mean (SD)	8310 (2850)	6990 (3430)	0.0706
Median [Min, Max]	7730 [3210, 16400]	6160 [2200, 15200]	
Insoluble Fiber [mg]			
Mean (SD)	19300 (7450)	17800 (9340)	0.436
Median [Min, Max]	17200 [7830, 40400]	15700 [5700, 44700]	
Sat. FA [mg]			
Mean (SD)	41300 (15300)	26100 (13600)	**< 0.001**
Median [Min, Max]	37800 [18600, 87000]	23600 [4780, 62000]	
Tryptophan [mg]			
Mean (SD)	1120 (315)	894 (522)	**0.0213**
Median [Min, Max]	1070 [537, 1940]	742 [276, 2680]	
Tyrosine [mg]			
Mean (SD)	3460 (1040)	2860 (1780)	0.0733
Median [Min, Max]	3330 [1570, 5590]	2350 [796, 9100]	
SCFA [mg]			
Mean (SD)	1790 (796)	1350 (1090)	**0.0452**
Median [Min, Max]	1720 [567, 3690]	983 [47.2, 4990]	
MCFA [mg]			
Mean (SD)	1600 (682)	1040 (756)	**0.00115**
Median [Min, Max]	1450 [759, 3470]	829 [123, 3380]	

DEGS FFQ Monthly Nutrient Intake by DFS Group			
LCFA [mg]			
Mean (SD)	81700 (29600)	52200 (24300)	**< 0.001**
Median [Min, Max]	76700 [38800, 172000]	46900 [13000, 109000]	
Omega-3 [mg]			
Mean (SD)	5950 (11500)	1340 (1070)	**0.0237**
Median [Min, Max]	2110 [607, 61500]	990 [331, 5510]	
Omega-6 [mg]			
Mean (SD)	13300 (5220)	8610 (4680)	**< 0.001**
Median [Min, Max]	12100 [5490, 28400]	7050 [2600, 20900]	

Regarding differences for dietary adherence groups (omnivorous n = 59, vegetarian n = 17), most vegetarians were in the LFS group (82%) (**SI** Table 2). Compared to omnivorous dieters, vegetarians reported significantly lower overall energy, fat, sugar, saturated fats, long-chain fatty acids and omega-3 intake (all $p < 0.05$). Despite lower overall energy intake, lignin was significantly higher ($p = 0.02$) and overall fiber intake was only slightly higher in the vegetarian diet group (vegetarians: 29 ± 11.7 g/day, non-vegetarians: 25.9 ± 11.9 g/day, $p = 0.34$). The diet groups did not differ in anthropometric measures, yet cholesterol, a biomarker reflective of metabolic health, was lower for vegetarians compared to non-vegetarians ($p = 0.04$) (**SI** Table 3).

Overall macronutrient composition of nutrient intake was comparable for FFQ data from different time periods (24h, 7d, 14d, 30d) and consisted across samples of 53–60% carbohydrates, 16–25% fat, 16–20% protein, 4–7% fiber (Fig. 4).

Correlation of computed nutrient intake with anthropometric markers

We tested if computed nutrient intake was related to anthropometric measures and serum markers in the overweight, omnivorous, main sample (for sample descriptives see **SI** Fig. 3 **+ 4**). Of note, sex-standardized body fat mass was highly anti-correlated with sex-standardized fat-free mass in males ($r = -0.76$), yet the inverse was true for females ($r = 0.35$). In addition, females showed high accordance of BMI and sex-standardized body fat mass ($r = 0.77$, Fig. 5a-b). Due to those differences in anthropometrics by sex, we considered sex-stratified analyses in further steps only.

Firstly, all computed nutrient intakes were highly positively correlated, except for a negative correlation between sugar and fat intake in females ($r = -0.14$, **Figure 5b**). Fiber intake adjusted for energy intake was not significantly different in females compared to males (females: 10.7±3.9 g/d; males: 9.6±2.4 g/d, p =

0.15). All nutrients, including fiber and except for saturated fats, were positively linked to BMI in males ($0.04 < r < 0.29$, **Figure 5a**). On the contrary, in females total fiber, insoluble fiber, cellulose and lignin intake were linked to lower BMI ($-0.36 < r < -0.09$, **Figure 5b**).

Regarding pair-wise correlations of nutrients with BMI, only those of total fiber and two of its subclasses (cellulose, insoluble fiber) ($r_{all} < -0.14$, $p_{all} < 0.05$), as well as of omega-6 intake ($r = 0.14$, $p < 0.05$) were statistically significant linked (Fig. 6). Further, higher fiber intake was significantly linked to lower body fat mass and mean diastolic blood pressure ($r_{all} < -0.16$, $p_{all} < 0.05$), whereas lean body mass, WHR, and lower mean systolic blood pressure were not ($p > 0.05$).

To account for interdependent data points from the same individuals and for overall energy intake, we ran linear mixed models with subject as random factor and age and sex as covariates. Here, fiber intake in grams showed a significant negative association with body fat mass ($b = -0.11$, $p < 0.02$), yet this turned insignificant when adjusting fiber intake for total calorie intake ($b = -0.05$, $p = 0.6$). Models for all other anthropometric markers (BMI, WHR, lean body mass, blood pressure) did not explain significant amounts of variance.

When merging samples and looking at the whole weight range from normal-weight to obese (n = 182, BMI: 18.6–36.4 kg/m^2 M ± SD: 25.9 ± 2.8; WHR: 0.65–0.98, M ± SD: 0.81 ± 0.05; data from GUT-BRAIN and GREADT), negative associations between nutrient intake and BMI remained strongest for fiber and its subclasses compared to all other nutrients ($-0.34 < r < -0.29$) (Fig. 7a). Moreover, WHR showed a small negative link to higher fiber intake ($r < -0.09$). Sex-stratified results for men only (n = 155) showed similar links for fiber intake with BMI and stronger links with WHR compared to the mixed sex sample (Fig. 7b). Linear mixed models adjusted for age, sex, and timepoint with subject as random factor did not show significant links between fiber intake adjusted for energy intake and anthropometric markers.

Discussion

We translated self-reported food intake (using the commonly used German DEGS1-FFQ) of different samples into detailed nutrient intake per day for various macro- and micronutrients and dietary components, assessed test-retest and between-timeframe reliability, and made the scoring scripts openly available. Although we assume a large proportion of under- and overreporting on the individual's level, the nutrient values were normally distributed and mostly met recommended reference intake on average, except for higher sugar and less fiber intake then recommended. This is in line with studies across different European countries showing that sugar intake can make up to 115 g/d or 20% of overall energy intake [37] and fiber intake only 14–21 g/d not reaching recommended levels [38]. As expected, males consumed higher amounts of nutrients than females, and differences in nutrient intake were present across groups of different dietary adherence or eating habits, such as between omnivorous and vegetarian dieters or dietary fat and sugar indices.

Reliability between last 7 days and 24 hours FFQ was medium for all macro- and micronutrients. In the validation study of the (original) DEGS1-FFQ on the food group level, reliability was assessed between 28 days FFQ and two 24h recalls by phone and ranged from low to good across 53 food items. Most validation studies use multiple 24 hour recalls per participant to assess dietary intake of different days. This way, day-to-day variations of food consumption can be recorded [39] and the imprecisions of FFQs at the individual level can be adjusted [40, 41]. Combining FFQ and 24hR has been shown to increase the correlation with actual dietary intake. While repeated administration of 24hR improves consistency with actual dietary intake and decreases random error, the benefit is marginal, with four 24hRs sufficing to reach a maximum correlation of 90% [40]. While the combination of FFQs with 24hR is beneficial, the use of 24hR to validate FFQs is not advised [42]. In comparison to the item-wise correlation, we computed reliability for nutrient intake summed over all food groups for 7 days and 24 hours assessed with the same online tool. The overall better performance of our nutrient scoring compared to the original food group scoring can be attributed to the reduction of outcome nutrient variables compared to food groups and the assumed higher consistency for overall nutrient intake patterns compared to single food item intake. The provided nutrient scoring of the DEGS1-FFQ can thus be considered valid in terms of assessing relatively similar nutrients and food groups when reporting 7 days or, to a lesser degree, 24h.

Assessing reliability over two timepoints within one month, we found moderate agreement for FFQ7d ($kappa_{all} \geq 0.40$, $kappa_{max} = 0.73$), similar to previous validation studies [43, 44], with highest reliability for protein and lowest for fat. For FFQ24h, reliability was poor ($kappa_{all} \geq 0.08$, $kappa_{max} = 0.34$), with highest reliability for carbohydrates and lowest for fiber. Indeed, the difference between FFQ7d and FFQ24h could reflect individual variance in eating habits that are more consistent over a time course of one week than on single days, including weekend days, when actual food intake might change quite drastically. To achieve highest correlation the administration of four 24hR has been recommended [40]. Therefore, the use of only two FFQ24h in our study might have led to underpowered results and lower consistency. Overall, FFQs relating to shorter time periods (e.g. 24h) may be helpful to assess diet as a confounder variables, e.g. for microbial sampling, yet reliability is poor. FFQs with longer time periods (e.g. 7 days or more) have higher reliability and should be used for assessing more habitual dietary habits or intake or adapted to the length of the intervention period.

We further evaluated the sensibility of the computed nutrient data and extended our nutrient scoring for DEGS1-FFQ data to two additional samples. The male-only sample with vegetarians and omnivores showed similar nutrient results as the sample with a wider BMI range and omnivores, with higher reported intake for the high fat and sugar group (for comparison: total energy intake: *GUT-BRAIN*: 1490–1740 kcal; *Mensa*: 1450-1640kcal, *GREADT*: 1710–2530 kcal), indicating that our results reflect common variance in actual eating habits. In particular for the male sample, the grouping into high vs. low dietary fat and sugar intake, showed distinct group differences for nutrient classes, sugar intake was much higher than recommended in both groups and fiber similarly lower than recommended. Further grouping into high vs. low fat and sugar intake showed distinct group differences for reported energy and nutrient intake (i.e. protein, fat, sugar, carbohydrates, saturated fats and others).

The DFS grouping might be limited in terms of comparing plant-based derived fiber intake as the questionnaire is by design focusing on processed and animal-based products. Sensible DFS grouping was confirmed by higher HbA1c levels in the high DFS group, reflecting a risk for diabetes development, coronary heart disease or stroke in the long-term [45].

Associations of anthropometrics and biomarkers with nutrient intake showed mixed results. First, all nutrients were collinearly linked. As expected, in the omnivorous, overweight sample higher total fiber intake across 7 days was weakly linked to lower BMI, body fat mass and mean diastolic blood pressure, pointing towards healthier diets high in fiber intake relating to lower weight in that sample. A link between fiber and lower weight has been shown before for diets restricted in animal-based foods [46] and systematically reviewed for whole-grain and fiber-rich foods [8]. All other nutrients were positively linked to BMI. However, when running linear models accounting for interdependency of datapoints and with fiber adjusted for caloric intake, these associations were not significant. When looking at the larger merged sample spanning from normal-weight to obese, higher total fiber and its subclasses (insoluble, soluble, cellulose, lignin) intake remained negatively linked to BMI and WHR. Yet again, linear mixed models showed no significant association. The fiber-BMI link was more distinct in sex-stratified analyses. All nutrients, including fiber, were linked to higher BMI in males. In females, however, higher fiber intake was strongly linked to lower BMI. Sex-specificity was further shown by an anti-correlation of body fat and fat-free mass in males, with the inverse relation in females, who showed high accordance of BMI and body fat mass.

Some studies reported inverse findings compared to ours: whole-grain intake was associated with lower BMI for both sexes, yet fiber in particular was inversely correlated with BMI only in men, not in women [47, 48] and likewise with immune function [49]. In contrast to previous studies [50, 51], fiber intake was not different between males and females in our sample, therefore fiber intake differences might have not played a strong role in sex-specific metabolic mechanisms. Proposed mechanisms of fiber intake in women may be metabolic benefits, i.e. reduced lipids in the blood, mediated by estradiol levels [52] and even blunted hormonal signaling during the reproductive cycle [53]. The picture on sex-specific associations of fiber intake on anthropometrics seems rather inconclusive and more studies are needed to disentangle sex-specific effects of fiber intake on metabolic, immune or reproductive markers.

Although we found links between nutrient intake and anthropometric markers, most associations were non-significant when adjusting for computed total energy intake. Although this might not reflect real total intake, proportions of nutrients may be relevant and telling within the same nutrient scoring pipeline within one dataset. We recommend adjusting nutrient-of-interest variables for overall calorie intake as has been done previously [54–56].

A strength of this analysis is the pooling of data from two human studies with deeply phenotyped samples with research questions focusing on eating habits resulting in a larger dataset with 189 data points. By increasing power, we could show that negative links between nutrient intake and BMI and WHR

are strongest for fiber and its subclasses compared to all other nutrients. Sex-stratified analyses showed stronger, but opposite links for fiber and anthropometric measures in men.

These results encourage researchers to regard calculated nutrient intake as a putative measure of interest to be extracted from FFQ data. Such calculation pipelines are rarely if at all available. Therefore we publish all scripts open access and open code. Overall, we propose FFQs along with automated nutrient scoring as a powerful tool to assess dietary intake. Our automated pipeline may contribute to developing nutrient scoring further and to advance nutrition sciences. In which context fiber intake may be a powerful tool for weight management and dietetic treatments as proposed before [57, 58] remains to be investigated further. Interestingly, as the sample consists of young to middle-aged, healthy overweight individuals, the observed associations already indicate moderate effects already in a healthy population. Linear mixed model estimates suggested that every 10 grams of dietary fiber per day was associated with around 0.5 to 1.0 kg/m^2 lower BMI. Dietary intake, in particular high fiber diets, have a large potential in preventing obesity-related states and comorbidities on a societal level [8, 59]. We suggest to increase educational efforts on fiber content of foods (as it is oftentimes not printed on food packaging, or available in experimental datasets, e.g. Food-pics database [60]) and to ameliorate policy making in the food sector (public and private) [61] and nutrition communication [62] to enhance fiber-rich diets and food items.

Overall, nutritional epidemiology will benefit from more advanced nutrient assessment and future studies revealing more insights on the impact of nutrient intake to provide more reliable and comparable evidence, which may inform public policy-making in the long-term.

Limitations

Firstly, due to inherent structure of the FFQ used, imprecision in the results might remain. For instance, only 53 food items are covered in the DEGS1, which means that a variety of different food items is not taken into account leading to gaps in data acquisition (e.g. legumes/ soy products,...). Another inaccuracy might stem from the fact that only certain FFQ questions are accompanied by a visual prompt, such as a picture of the portion size of a certain food (as provided by the Robert Koch Institute). As a result, the lack of a benchmark when estimating food intake might have led to deviations in the assessment.

Secondly, large ranges of values were present for some nutrients. This can be attributed to occasional over- or underreporting, however, we refrained from excluding datapoints to show truly reported values showcasing potential reporting biases. This shows the variability of self-reported FFQ data at the nutrient level and leaves room for defining data curation strategies depending on the research question, such as curating implausible data entries by consensus decision of different raters.

Thirdly, as the main sample is partly from a dietary intervention study, characteristics of this sample may reflect a selection bias in favour of omnivorous eaters with some awareness of the study goals to influence eating behaviour. Yet, we cross-validated data from the main sample with two other samples,

albeit also from studies focusing on eating behaviour. Overall, self-reported dietary data is never blinded and neutral, since participants may reflect on social desirability and therefore report in a biased way. Yet, self-reported FFQ data in combination with recalls are a valid tool to assess nutrient intake [40] and relative nutrient intake can be compared.

Conclusions

Our newly developed nutrient scoring allows to extract specific nutrient information of interest that can be further used to address specific research questions and to reduce dimensionality of FFQ outcomes, thus improving comparability for (self-reported) nutrient intake across studies. Reliability of computed nutrient values is similar to previously reported dietary intake for 24 hours and 7 days reports. We believe that by making the scripts and descriptives for nutrient scoring available, we can provide the nutrition research community with more precise proxies for dietary intake, especially with respect to the German DEGS1-FFQ, but also by allowing to adopt the openly shared methodology to other questionnaires.

Declarations

Ethics approval and consent to participate

The study was conducted according to the guidelines of the Declaration of Helsinki and the institutional ethics board of the Medical Faculty of the University of Leipzig, Germany, raised no concerns regarding the study protocol (228/18-ek and 400/18-ek) and gave approval for the current study. All participants provided written informed consent. Recruitment took place via the institute's database and advertisements.

Consent for publication

All authors agreed to the final version of this manuscript.

Availability of data and materials

All scripts are available here: https://gitlab.gwdg.de/omega-lab/ffq-nutrient-scoring/

Competing Interests

The authors declare that they have no competing interests.

Funding

Funding has been provided by the Max Planck Society and the German Environmental Foundation (EM).

Authors' contributions

study conception: EM, RT, AH, AVW

data collection: EM, RT, LS, HH

dietary scoring: RT

data analysis: EM, RT, LS, AK

manuscript first draft: EM, LS

manuscript final version: AH, HH, AK, EM, RT, LS, AVW

Acknowledgements

We thank Tatjana Schütz for helpful discussions and Frauke Beyer for advice on statistics and data illustration. We also thank Laura Hesse, Charlotte Wiegank and Emmy Töws as the medical staff of the GUT-BRAIN study. A special thanks goes to Lukas Recker for writing R-scripts according to the manual of the DEGS1-FFQ. Moreover, Emmy Töws and Iryna Manuilova contributed mainly to extracting reference nutrients values from the BLS. All blood markers were analysed by the Institute for Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics (ILM) Leipzig University, Leipzig, Germany and the Synevo Studien Service Labor GmbH c/o IMD Institut für Medizinische Diagnostik Berlin-Potsdam GbR, Berlin.

References

1. Shim J-S, Oh K, Kim HC. Dietary assessment methods in epidemiologic studies. Epidemiol Health. 2014;36:e2014009. doi:10.4178/epih/e2014009.
2. Briefel RR, Flegal KM, Winn DM, Loria CM, Johnson CL, Sempos CT. Assessing the nation's diet: limitations of the food frequency questionnaire. J Am Diet Assoc. 1992.
3. Ioannidis JPA. The challenge of reforming nutritional epidemiologic research. JAMA - Journal of the American Medical Association. 2018;320:969–70.
4. Maugeri A, Barchitta M. A systematic review of ecological momentary assessment of diet: implications and perspectives for nutritional epidemiology. Nutrients. 2019;11:2696.
5. Cade J, Thompson R, Burley V, Warm D. Development, validation and utilisation of food-frequency questionnaires – a review. Public Health Nutr. 2002;5:567–87.
6. Crispim SP, Nicolas G, Casagrande C, Knaze V, Illner AK, Huybrechts I, et al. Quality assurance of the international computerised 24 h dietary recall method (EPIC-Soft). Br J Nutr. 2014;111:506–15.
7. Timon CM, Van Den Barg R, Blain RJ, Kehoe L, Evans K, Walton J, et al. A review of the design and validation of web- and computer-based 24-h dietary recall tools. Nutrition Research Reviews. 2016;29:268–80.

8. Reynolds A, Mann J, Cummings J, Winter N, Mete E, Te Morenga L. Carbohydrate quality and human health: a series of systematic reviews and meta-analyses. Lancet. 2019. doi:10.1016/S0140-6736(18)31809-9.
9. Yang Y, Zhao L-G, Wu Q-J, Ma X, Xiang Y-B. Association Between Dietary Fiber and Lower Risk of All-Cause Mortality: A Meta-Analysis of Cohort Studies. Am J Epidemiol. 2015;181:83–91. doi:10.1093/aje/kwu257.
10. Mai V, Flood A, Peters U, Lacey J V., Schairer C, Schatzkin A. Dietary fibre and risk of colorectal cancer in the Breast Cancer Detection Demonstration Project (BCDDP) follow-up cohort. Int J Epidemiol. 2003;32:234–9.
11. Milajerdi A, Ebrahimi-Daryani N, Dieleman LA, Larijani B, Esmaillzadeh A. Association of Dietary Fiber, Fruit, and Vegetable Consumption with Risk of Inflammatory Bowel Disease: A Systematic Review and Meta-Analysis. Advances in Nutrition. 2021;12:735–43. doi:10.1093/advances/nmaa145.
12. Fatahi S, Matin SS, Sohouli MH, Găman MA, Raee P, Olang B, et al. Association of dietary fiber and depression symptom: A systematic review and meta-analysis of observational studies. Complementary Therapies in Medicine. 2021;56:102621.
13. Clarys P, Deriemaeker P, Huybrechts I, Hebbelinck M, Mullie P. Dietary pattern analysis: A comparison between matched vegetarian and omnivorous subjects. Nutrition Journal. 2013;12. doi:10.1186/1475-2891-12-82.
14. Davies GJ, Crowder M, Dickerson JW. Dietary fibre intakes of individuals with different eating patterns. Hum Nutr Appl Nutr. 1985;39:139–48.
15. Cade JE, Burley VJ, Greenwood DC. Dietary fibre and risk of breast cancer in the UK Women's Cohort Study. Int J Epidemiol. 2007;36:431–8.
16. Davey GK, Spencer EA, Appleby PN, Allen NE, Knox KH, Key TJ. EPIC–Oxford: lifestyle characteristics and nutrient intakes in a cohort of 33 883 meat-eaters and 31 546 non meat-eaters in the UK. Public Health Nutr. 2003;6:259–68.
17. Breen C, Ryan M, McNulty B, Gibney MJ, Canavan R, O'Shea D. High saturated-fat and low-fibre intake: A comparative analysis of nutrient intake in individuals with and without type 2 diabetes. Nutr Diabetes. 2014;4:104. doi:10.1038/nutd.2014.2.
18. Ward HA, Keogh R, Lentjes M, Luben RN, Wareham NJ, Khaw KT. Fibre intake in relation to serum total cholesterol levels and CHD risk: A comparison of dietary assessment methods. Eur J Clin Nutr. 2012;66:296–304. doi:10.1038/ejcn.2011.184.
19. Healey G, Brough L, Murphy R, Hedderley D, Butts C, Coad J. Validity and reproducibility of a habitual dietary fibre intake short food frequency questionnaire. Nutrients. 2016;8.
20. Levi F, Pasche C, Lucchini F, La Vecchia C. Dietary fibre and the risk of colorectal cancer. Eur J Cancer. 2001;37:2091–6.
21. Leong C, Taylor RW, Haszard JJ, Fleming EA, Tannock GW, Szymlek-Gay EA, et al. Relative validity and reproducibility of a food frequency questionnaire to assess nutrients and food groups of

relevance to the gut microbiota in young children. Nutrients. 2018;10. doi:10.3390/nu10111627.

22. Bolca S, Huybrechts I, Verschraegen M, De Henauw S, Van de Wiele T. Validity and Reproducibility of a Self-Administered Semi-Quantitative Food-Frequency Questionnaire for Estimating Usual Daily Fat, Fibre, Alcohol, Caffeine and Theobromine Intakes among Belgian Post-Menopausal Women. Int J Environ Res Public Health. 2009;6:121–50. doi:10.3390/ijerph6010121.
23. Persson P-G, Carlsson S, Grill V, Hagman U, Lundgren A-C, Östenson C-G, et al. Food frequency questionnaire versus 7-day weighed dietary record information on dietary fibre and fat intake in middle-aged Swedish men. Scand J Soc Med. 1998;26:75–80. doi:10.1177/14034948980260010301.
24. Neyrinck AM, Nazare J-A, Rodriguez J, Jottard R, Dib S, Sothier M, et al. Development of a Repertoire and a Food Frequency Questionnaire for Estimating Dietary Fiber Intake Considering Prebiotics: Input from the FiberTAG Project. Nutrients. 2020;12:2824. doi:10.3390/nu12092824.
25. Freedman LS, Commins JM, Moler JE, Arab L, Baer DJ, Kipnis V, et al. Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for energy and protein intake. American Journal of Epidemiology. 2014;180:172–88.
26. Haftenberger M, Heuer T, Heidemann C, Kube F, Krems C, Mensink GB. Relative validation of a food frequency questionnaire for national health and nutrition monitoring. 2010.
27. Naska A, Lagiou A, Lagiou P. Dietary assessment methods in epidemiological research: Current state of the art and future prospects. F1000Research. 2017;6.
28. Esche J, Krupp D, Mensink GBM, Remer T. Dietary potential renal acid load is positively associated with serum uric acid and odds of hyperuricemia in the German adult population. J Nutr. 2018;148:49–55.
29. Paprott R, Mühlenbruch K, Mensink GBM, Thiele S, Schulze MB, Scheidt-Nave C, et al. Validation of the German Diabetes Risk Score among the general adult population: Findings from the German Health Interview and Examination Surveys. BMJ Open Diabetes Res Care. 2016;4:280. doi:10.1136/bmjdrc-2016-000280.
30. Scheidt-Nave C, Kamtsiuris P, Gößwald A, Hölling H, Lange M, Busch MA, et al. German health interview and examination survey for adults (DEGS)-design, objectives and implementation of the first data collection wave. BMC Public Health. 2012;12:1–16.
31. Huhn S, Beyer F, Zhang R, Lampe L, Grothe J, Kratzsch J, et al. Effects of resveratrol on memory performance, hippocampus connectivity and microstructure in older adults–A randomized controlled trial. Neuroimage. 2018.
32. Truthmann J, Mensink GBM, Richter A. Relative validation of the KiGGS Food Frequency Questionnaire among adolescents in Germany. Nutr J. 2011;10:133.
33. Medawar E, Zedler M, de Biasi L, Villringer A, Witte AV. Distinct effects of single plant-based vs. animal-based meals on satiety and mood revealed by real-world smartphone-embedded studies. medRxiv. 2021;:2021.10.24.21265240. doi:10.1101/2021.10.24.21265240.

34. Hartmann H, Janssen LK, Rausch F, Horstmann A. The interplay between dopamine genes and diet. OSF. 2022. https://osf.io/vs8cm/. Accessed 24 Feb 2022.
35. Fromm SP, Horstmann A. Psychometric Evaluation of the German Version of the Dietary Fat and Free Sugar-Short Questionnaire. Obes Facts. 2019;12:518–28. doi:10.1159/000501969.
36. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15:155–63. doi:10.1016/j.jcm.2016.02.012.
37. Azaïs-Braesco V, Sluik D, Maillot M, Kok F, Moreno LA. A review of total & added sugar intakes and dietary sources in Europe. Nutr J. 2017;16:1–15.
38. Stephen AM, Champ MM-J, Cloran SJ, Fleith M, van Lieshout L, Mejborn H, et al. Dietary fibre in Europe: current state of knowledge on definitions, sources, recommendations, intakes and relationships to health. Nutr Res Rev. 2017;30:149–90.
39. Ahluwalia N, Dwyer J, Terry A, Moshfegh A, Johnson C. Update on NHANES Dietary Data: Focus on Collection, Release, Analytical Considerations, and Uses to Inform Public Policy. Adv Nutr. 2016;7:121–34.
40. Freedman LS, Midthune D, Arab L, Prentice RL, Subar AF, Willett W, et al. Combining a Food Frequency Questionnaire with 24-Hour Recalls to Increase the Precision of Estimation of Usual Dietary Intakes-Evidence from the Validation Studies Pooling Project. Am J Epidemiol. 2018;187:2227–32.
41. Mitry P, Wawro N, Six-Merker J, Zoller D, Jourdan C, Meisinger C, et al. Usual Dietary Intake Estimation Based on a Combination of Repeated 24-H Food Lists and a Food Frequency Questionnaire in the KORA FF4 Cross-Sectional Study. Front Nutr. 2019;6.
42. Freedman LS, Commins JM, Willett W, Tinker LF, Spiegelman D, Rhodes D, et al. Evaluation of the 24-Hour Recall as a Reference Instrument for Calibrating Other Self-Report Instruments in Nutritional Cohort Studies: Evidence from the Validation Studies Pooling Project. Am J Epidemiol. 2017;186:73–82.
43. Wong JE, Parnell WR, Black KE, Skidmore PML. Reliability and relative validity of a food frequency questionnaire to assess food group intakes in New Zealand adolescents. Nutr J. 2012;11:65. doi:10.1186/1475-2891-11-65.
44. Cheng Z, Shuai P, Qiao Q, Li T. Validity and reliability of a simplified food frequency questionnaire: a cross sectional study among physical health examination adults in southwest region of China. Nutr J. 2020;19:1–10.
45. Sherwani SI, Khan HA, Ekhzaimy A, Masood A, Sakharkar MK. Significance of HbA1c Test in Diagnosis and Prognosis of Diabetic Patients. Biomark Insights. 2016;11:95–104. doi:10.4137/BMI.S38440.
46. Medawar E, Enzenbach C, Roehr S, Villringer A, Riedel-Heller SG, Witte AV. Less Animal-Based Food, Better Weight Status: Associations of the Restriction of Animal-Based Product Intake with Body-Mass-Index, Depressive Symptoms and Personality in the General Population. Nutrients. 2020;12:1492.

47. van de Vijver LPL, van den Bosch LMC, van den Brandt PA, Goldbohm RA. Whole-grain consumption, dietary fibre intake and body mass index in the Netherlands cohort study. Eur J Clin Nutr. 2009;63:31–8. doi:10.1038/sj.ejcn.1602895.
48. Byrd-Williams CE, Strother ML, Kelly LA, Huang TTK. Dietary fiber and associations with adiposity and fasting insulin among college students with plausible dietary reports. Nutrition. 2009;25:896–904. doi:10.1016/j.nut.2009.02.003.
49. Fernstrand AM, Bury D, Garssen J, Verster JC. Dietary intake of fibers: differential effects in men and women on perceived general health and immune functioning. Food Nutr Res. 2017;61:1297053.
50. Fayet-Moore F, Cassettari T, Tuck K, McConnell A, Petocz P. Dietary fibre intake in Australia. Paper I: Associations with demographic, socio-economic, and anthropometric factors. Nutrients. 2018;10. doi:10.3390/nu10050599.
51. Grooms KN, Ommerborn MJ, Pham DQ, Djoussé L, Clark CR. Dietary fiber intake and cardiometabolic risks among US adults, NHANES 1999–2010. Am J Med. 2013;126. doi:10.1016/j.amjmed.2013.07.023.
52. Mumford SL, Schisterman EF, Siega-Riz AM, Gaskins AJ, Wactawski-Wende J, Vanderweele TJ. Effect of dietary fiber intake on lipoprotein cholesterol levels independent of estradiol in healthy premenopausal women. Am J Epidemiol. 2011;173:145–56. doi:10.1093/aje/kwq388.
53. Gaskins AJ, Mumford SL, Zhang C, Wactawski-Wende J, Hovey KM, Whitcomb BW, et al. Effect of daily fiber intake on reproductive function: the BioCycle Study. Am J Clin Nutr. 2009;90:1061–9. doi:10.3945/ajcn.2009.27990.
54. García-Meseguer M, Delicado-Soria A, Serrano-Urrea R. Fiber Patterns in Young Adults Living in Different Environments (USA, Spain, and Tunisia). Anthropometric and Lifestyle Characteristics. Nutrients. 2017;9:1030. doi:10.3390/nu9091030.
55. Dahl WJ, Stewart ML. Position of the Academy of Nutrition and Dietetics: Health Implications of Dietary Fiber. J Acad Nutr Diet. 2015;115:1861–70. doi:10.1016/j.jand.2015.09.003.
56. Wang AYM, Sea MMM, Ng K, Wang M, Chan IHS, Lam CWK, et al. Dietary Fiber Intake, Myocardial Injury, and Major Adverse Cardiovascular Events Among End-Stage Kidney Disease Patients: A Prospective Cohort Study. Kidney Int Reports. 2019;4:814–23.
57. Reynolds AN, Akerman AP, Mann J. Dietary fibre and whole grains in diabetes management: Systematic review and meta-analyses. PLOS Med. 2020;17:e1003053. doi:10.1371/journal.pmed.1003053.
58. Medawar E, Huhn S, Villringer A, Veronica Witte A. The effects of plant-based diets on the body and the brain: a systematic review. Transl Psychiatry. 2019;9:1–17. doi:10.1038/s41398-019-0552-0.
59. Miketinas DC, Bray GA, Beyl RA, Ryan DH, Sacks FM, Champagne CM. Fiber Intake Predicts Weight Loss and Dietary Adherence in Adults Consuming Calorie-Restricted Diets: The POUNDS Lost (Preventing Overweight Using Novel Dietary Strategies) Study. J Nutr. 2019;149:1742–8. doi:10.1093/jn/nxz117.

60. Blechert J, Lender A, Polk S, Busch NA, Ohla K. Food-pics_extended—an image database for experimental research on eating and appetite: additional images, normative ratings and an updated review. Front Psychol. 2019;10.
61. Renner B, Arens-Azevêdo U, Watzl B, Richter M, Virmani K, Linseisen J. DGE-Positionspapier zur nachhaltigeren Ernährung. Ernährungs-Umschau. 2021;68:144–54.
62. Jones JM. Dietary fiber future directions: Integrating new definitions and findings to inform nutrition research and communication. Advances in Nutrition. 2013;4:8–15. doi:10.3945/an.112.002907.

Figures

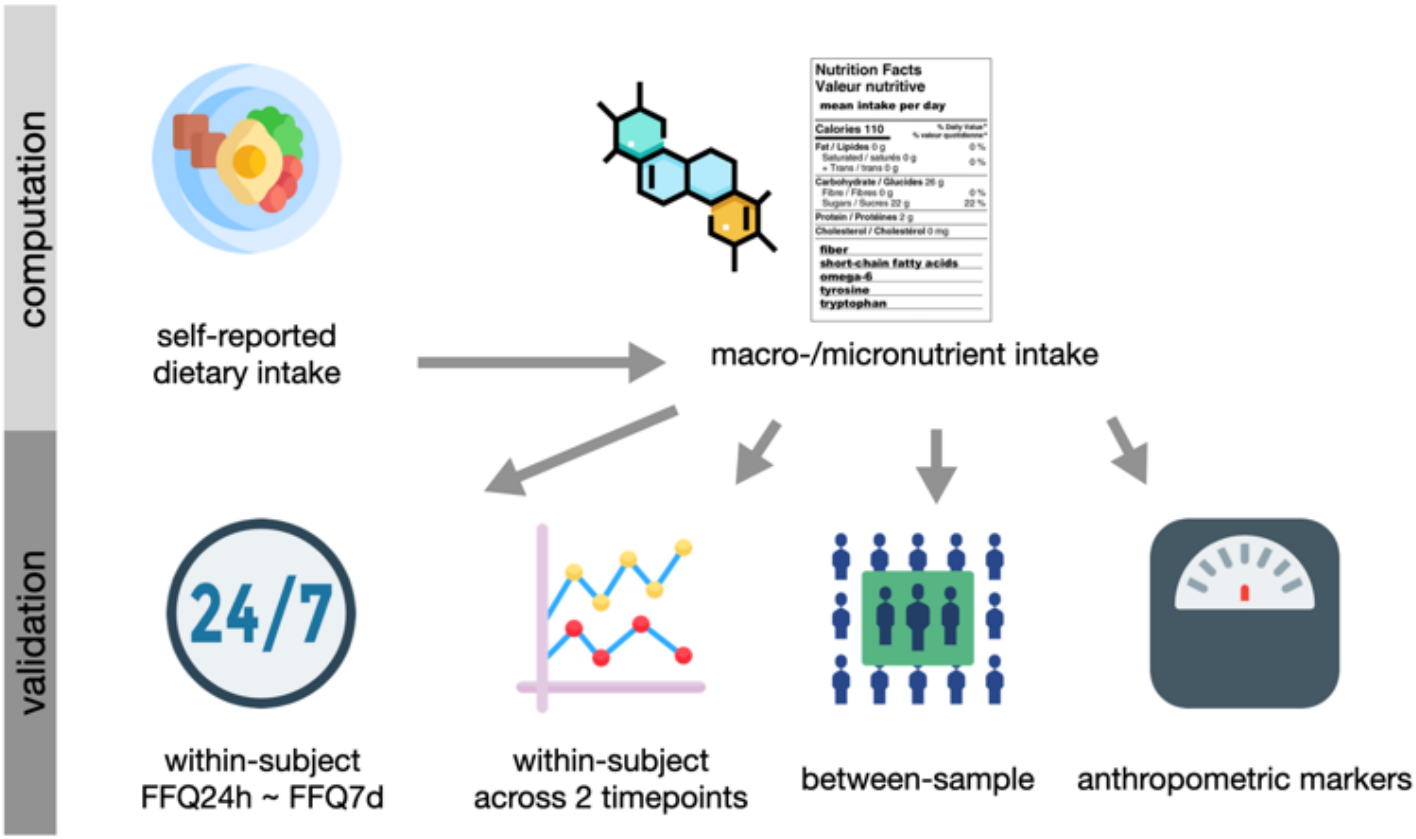


Figure 1

Schematic overview of nutrient intake computation and validation in this study

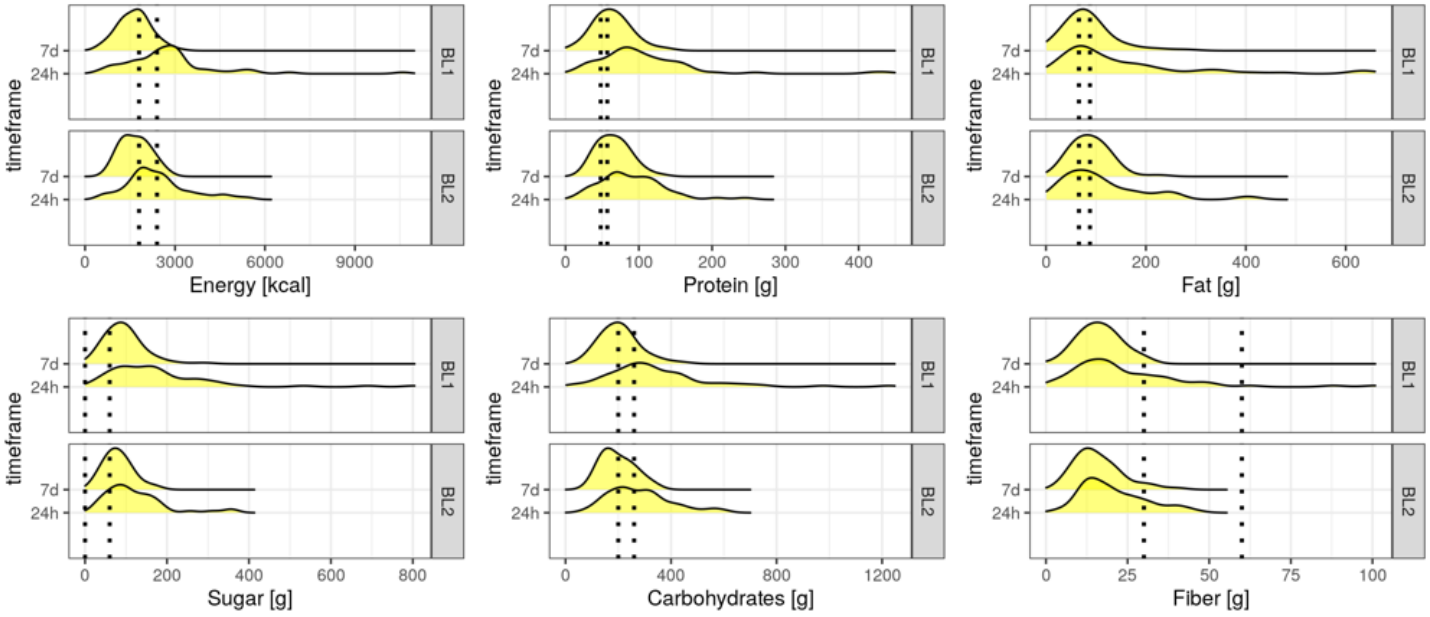


Figure 2

Frequency for computed macronutrients for FFQ 7 days and 24 hours (n=110 datapoints in total from first and second baseline). Dotted lines represent reference values (sources: "German Society for Nutrition" for most nutrients, "German Nutrition Counselling Network (DEBI)" for sugar).

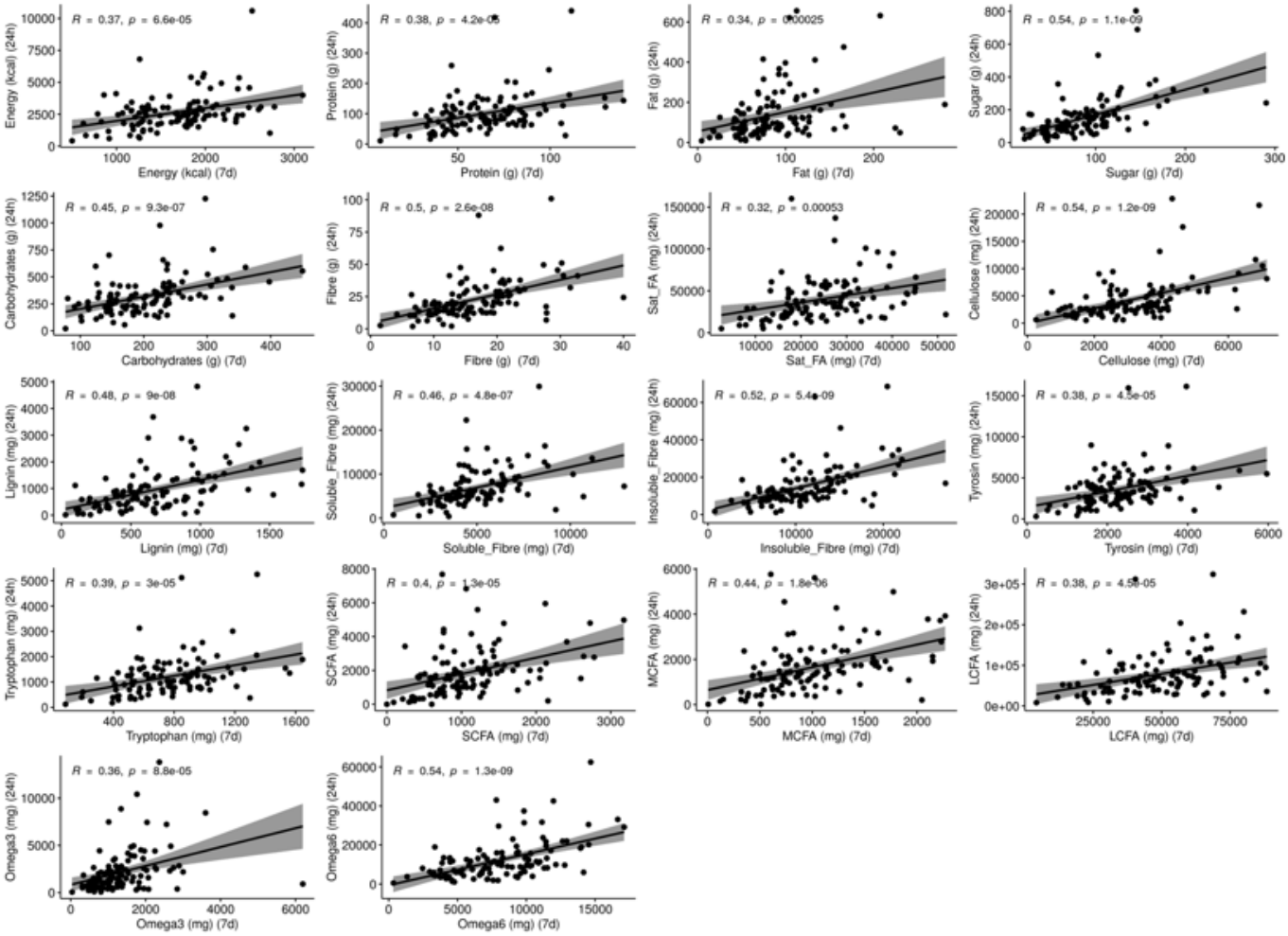


Figure 3

Pearson's correlation between all macro- and micronutrients between FFQ7d and FFQ24h. 95% CI depicted

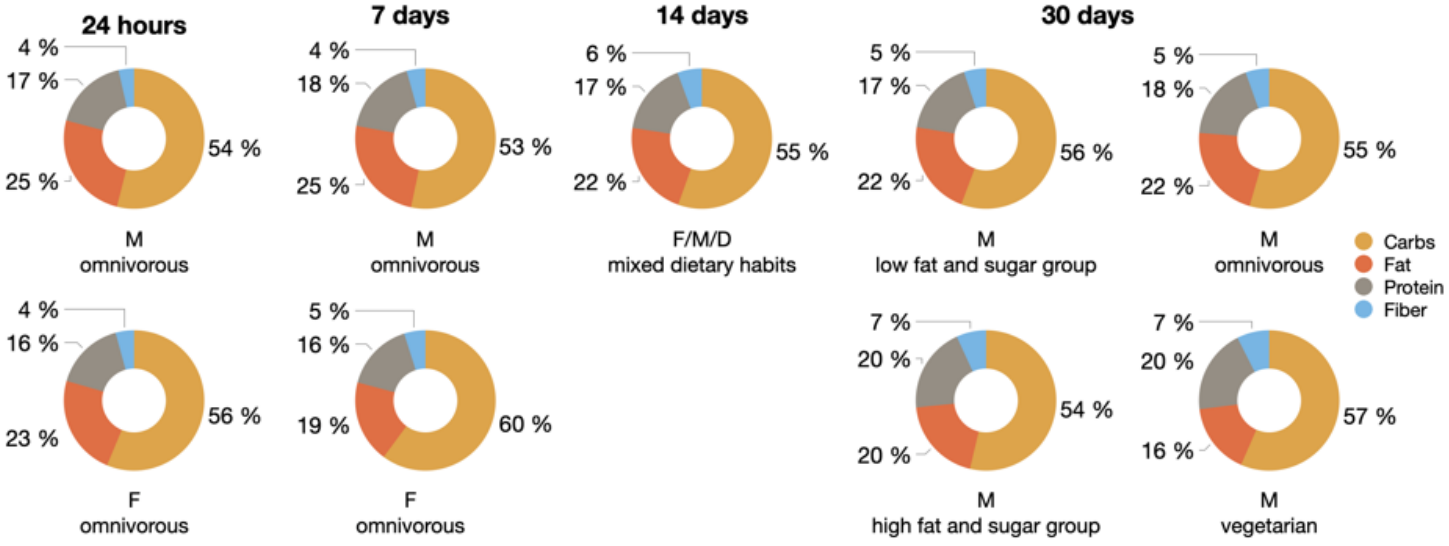


Figure 4

Pie charts for proportions of computed nutrient intakes per time frame of questionnaire per dataset in percent for carbohydrates, fat, protein and fiber intake. 24 hours and 7 days are based on data from the first baseline of the GUT-BRAIN dataset. Calorie intake was not considered here. Abbreviations: F: female, M: male, D: diverse.

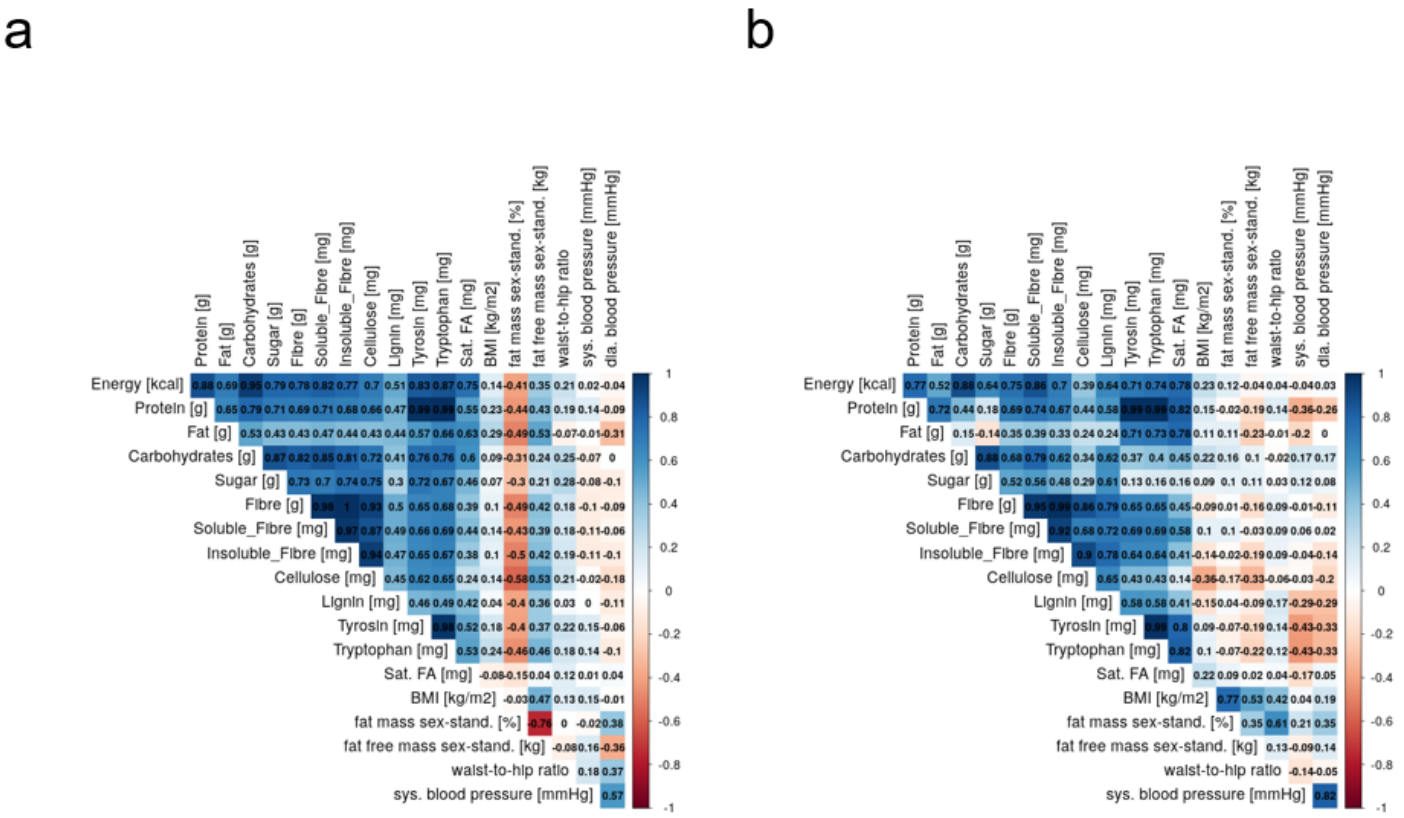


Figure 5

Correlation plot of nutrient-related markers and laboratory markers for BL1 timepoint only for a) males only and b) females only (Spearman's rank correlation coefficient).

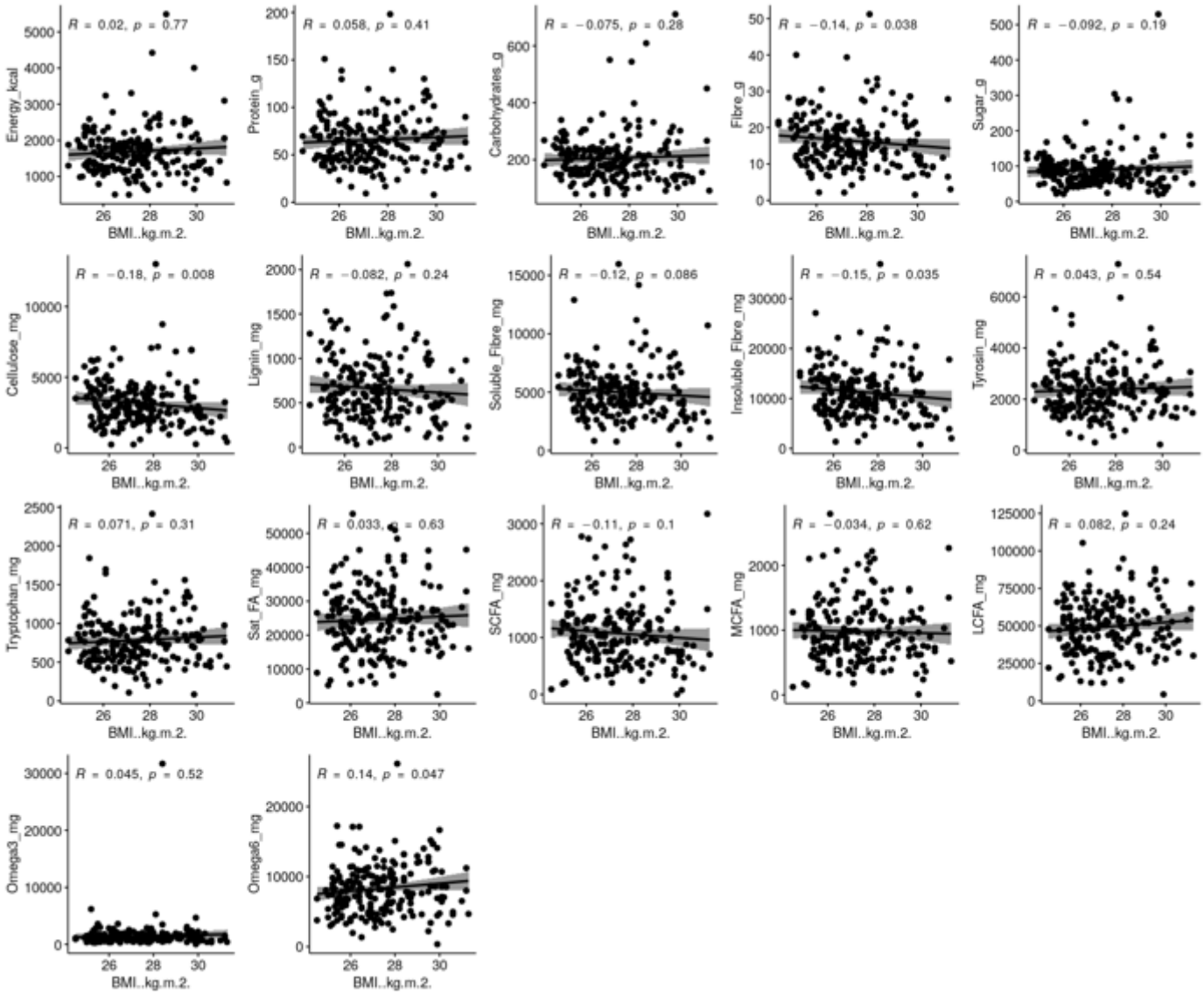
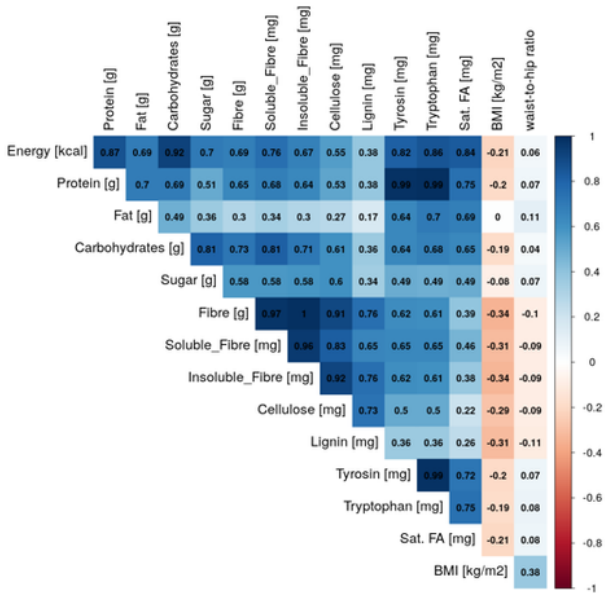


Figure 6

Correlations between weekly nutrient intake and BMI (Spearman's rank correlation coefficient). 95% CI plotted.

a



b

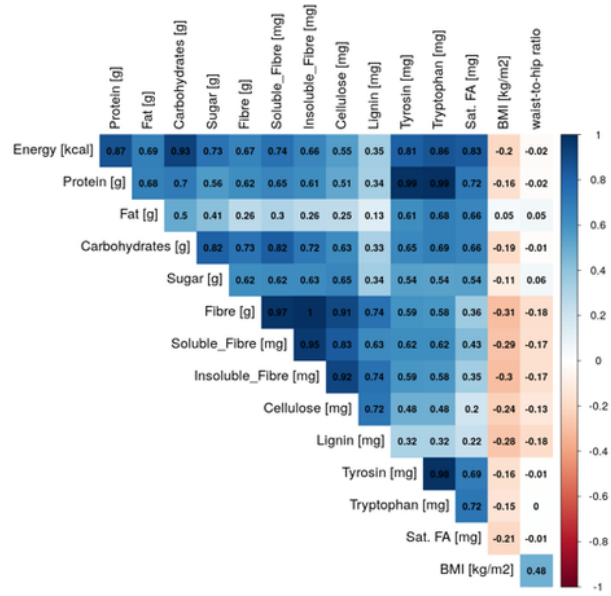


Figure 7

Correlations (Spearman's rank correlation coefficient) between computed nutrient intake and anthropometric markers for a BMI range of 18.6-36.4 kg/m^2 for a) female and male and b) for male only (GUT-BRAIN + GREADT).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformation.docx