



The reliability and relative validity of predefined dietary patterns were higher than that of exploratory dietary patterns in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam population

Franziska Jannasch^{1,2,3*}, Daniela Nickel^{1,2,3} and Matthias B. Schulze^{1,2,3,4}

¹Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal 14558, Germany

²German Center for Diabetes Research (DZD), München-Neuberberg, Germany

³NutriAct – Competence Cluster Nutrition Research Berlin-Potsdam, Nuthetal, Germany

⁴Institute of Nutrition Science, University of Potsdam, Nuthetal, Germany

(Submitted 4 May 2020 – Final revision received 24 August 2020 – Accepted 29 August 2020 – First published online 8 September 2020)

Abstract

The aim of this study was to assess the ability of the FFQ to describe reliable and valid dietary pattern (DP) scores. In a total of 134 participants of the European Prospective Investigation into Cancer and Nutrition-Potsdam study aged 35–67 years, the FFQ was applied twice (baseline and after 1 year) to assess its reliability. Between November 1995 and March 1997, twelve 24-h dietary recalls (24HDR) as reference instrument were applied to assess the validity of the FFQ. Exploratory DP were derived by principal component analyses. Investigated predefined DP were the Alternative Healthy Eating Index (AHEI) and two Mediterranean diet indices. From dietary data of each FFQ, two exploratory DP were retained, but differed in highly loading food groups, resulting in moderate correlations (r 0.45–0.58). The predefined indices showed higher correlations between the FFQ (r (AHEI) 0.62, r (Mediterranean Diet Pyramid Index (MedPyr)) 0.62 and r (traditional Mediterranean Diet Score (tMDS)) 0.51). From 24HDR dietary data, one exploratory DP retained differed in composition to the first FFQ-based DP, but showed similarities to the second DP, reflected by a good correlation (r 0.70). The predefined DP correlated moderately (r 0.40–0.60). To conclude, long-term analyses on exploratory DP should be interpreted with caution, due to only moderate reliability. The validity differed extensively for the two exploratory DP. The investigated predefined DP showed a better reliability and a moderate validity, comparable to other studies. Within the two Mediterranean diet indices, the MedPyr performed better than the tMDs in this middle-aged, semi-urban German study population.

Key words: Dietary patterns; Reliability; Validity

One unquestioned health input comes from a large variety of nutrients that are metabolised in the human body every day. However, nutrient intake results from diet and dietary behaviour. Nutritional epidemiology focused for a long time on the health relevance of the nutrient intake or the role of single food items. However, during the last two decades, the investigation of dietary patterns (DP) arose as a promising alternative to investigations on the dietary influence of nutrients or single foods on health outcomes⁽¹⁾.

DP can generally be investigated using (1) predefined dietary indices for a specific pattern or (2) by the use of exploratory methods like principal component analysis (PCA). The latter

results in data-driven DP, thus, highly dependent on the underlying study-specific data structure of food groups. Despite the growing number of studies on exploratory DP^(2,3), data on the validity and reliability of exploratory DP are sparse. In a systematic literature review, we were able to identify only seven studies investigating validity and reliability of exploratory DP in five countries worldwide. The included DP were mainly derived by PCA or factor analysis, except for one Australian study which used reduced rank regression⁽⁴⁾. In this review, a moderate validity for all DP was concluded. This was reflected by correlations ranging between 0.36 and 0.75 and a comparable number and composition of DP between the FFQ as study instru-

Abbreviations: 24HDR, 24-h dietary recall; AHEI, Alternative Healthy Eating Index; DP, dietary pattern; DS, diet screener; EPIC, European Prospective Investigation into Cancer and Nutrition; MEDAS, Mediterranean Diet Adherence Screener; mHDR, mean of twelve 24HDR; PCA, principal component analysis; tMDS, traditional Mediterranean Diet Score.

* **Corresponding author:** Dr Franziska Jannasch, email Franziska.Jannasch@dife.de



ment and dietary records (covering 3–7 d) or 24-h dietary recalls (24HDR) as reference instruments.

Predefined DP based on indices usually depend on food groups or single food items, which were selected due to universally accepted knowledge about their role for specific health outcomes. Numerous such dietary indices have been investigated^(2,3,5). However, an index well investigated across many study populations is the Alternative Healthy Eating Index (AHEI)⁽²⁾. Alternatively of a universal index, predefined DP can also be used to reflect supposedly healthy regional dietary habits. Such an example is the Mediterranean diet which has been investigated in numerous studies, showing associations with lower chronic disease risk across many study populations (including non-Mediterranean populations)⁽⁶⁾. However, a variety of indices for the Mediterranean diet exist. Most of them differ in the way estimates of dietary intake are handled in the score calculation. They either use the population-specific distribution ('traditional Mediterranean Diet Score' (tMDS)⁽⁷⁾) or cut-offs of absolute intake (Mediterranean Diet Pyramid Index (MedPyr)⁽⁸⁾. Despite the variety of studies on associations between these predefined indices and risk for many chronic diseases, systematic investigations on their validity and reliability are sparse^(9,10).

The European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam study is one of two cohorts in Germany within the European-wide multi-centre EPIC study. The originally developed dietary assessment instrument was a semi-quantitative FFQ with 158 items and eighty-seven coloured photographs. Prior to baseline, the relative validity and reproducibility of the food group intake were assessed⁽¹¹⁾. The relative validity, assessed by the use of the mean of twelve 24HDR (mHDR) as reference instrument, was moderate for most food groups (r 0.40–0.60), but extremely low for legumes (r 0.14) and extremely high for alcoholic beverages (r 0.90). The reproducibility, assessed by the repeated application of the FFQ within a 6-month interval, was good for half of the food groups (correlation > 0.70) and moderate for the other half ($r < 0.70$). The reproducibility was lowest for bread (r 0.49) and highest for alcoholic beverages (r 0.89)^(11,12). Validation studies of nutrient intake in the EPIC-Potsdam cohort and EPIC-Heidelberg cohorts^(13,14) indicated a comparable relative validity of nutrient intake measured by the FFQ⁽¹²⁾.

The use of DP in analyses of the association between diet and chronic disease risk often resulted in largely differing estimates between study populations⁽²⁾. Mostly, dietary data were assessed with only one application of the FFQ in the studies, which is a cheap and easy method. However, we also need to ensure not only on food group and nutrient level that the provided dietary data are not the result of randomness. Furthermore, the rather low validity and reliability of some particular food groups⁽¹¹⁾ should not affect the ability of the FFQ as study instrument in terms of describing valid and reliable DP. Only if it can be assured that the derived DP are stable constructs, associations with chronic disease risk can be interpreted with confidence.

Therefore, the aim of this study is to investigate the ability of the FFQ to describe validly exploratory DP derived by PCA and predefined dietary indices based on the comparison with DP derived by dietary data from repeated 24HDR. Another aim is

the investigation of the ability of the FFQ to reliably describe these DP through the repeated application of the FFQ within 1 year.

Methods

Subjects

The EPIC-Potsdam study started with the recruitment of participants between August 1994 and September 1998. The study participants were randomly selected from population registries in Potsdam and the surrounding area, who met the age criteria for men (40–64 years) and for women (35–64 years). Additionally, interested people, who called voluntarily to participate in the study, were also included, if they met the age criteria. On average, the participation rate was 22.7 % from all people, who were contacted via mail, resulting in 27 548 participants at baseline⁽¹³⁾. Within the recruitment procedure, during the time period from November 1995 to March 1997, a total of 160 participants of the EPIC-Potsdam study were asked after the baseline examinations whether they participate in a sub-study study with repeated dietary assessments and blood draws. All participants invited agreed to take part and were divided into three age- and sex-specific strata according to the EPIC protocol^(14,15). This sub-study was conducted to assess the validity and reliability of the baseline dietary assessment instrument, a 146-item FFQ by comparing it, among others, with twelve 24HDR. Those participants, where information on at least ten (from a maximum of twelve) 24HDR was available (n 134), were included in the present analysis. All study procedures were approved by the Ethical Committee of the State of Brandenburg and participants agreed by signing the informed consent form. The age range of men (n 75) and women (n 59) was 40–67 years and 35–66 years, respectively.

Sociodemographic, lifestyle and dietary assessment

Information on the highest achieved educational level (no vocational training/vocational training, technical college and university) and physical activity (sports and biking in hours per week) was inquired with self-administered questionnaires⁽¹⁶⁾. The first dietary assessment was the FFQ which participants had to fill in within the regular recruitment procedure where the participants were asked in advance to the visit of the study centre to fill in a self-administered, scanner-readable questionnaire⁽¹⁴⁾. The FFQ collects semi-quantitatively for each food item information on the usual portion size and the average frequency of intake of 146 food items during the past 12 months. Questions about fat content, types of fats and sauces used for food preparation and the seasonality of specific food items were included additionally. Portion size for each item was estimated via photographs of three different portion sizes or with standard portion sizes like, for example, a cup (=150 ml). If several food items belonging to the food groups were inquired, there is potential to overestimate their consumption. Hence, questions on the usual consumption of the main food groups like bread, meat, fruits and vegetables per week were added at the end of the questionnaire to be



able to correct for potential overestimation. The questionnaire was scanned immediately upon arrival of the participant at the study centre. Completeness and plausibility were verified with a software program, and missing data were completed or corrected at the end of the examinations together with the participant⁽¹⁴⁾. The same FFQ was administered again after 12 months. Over the 1-year study phase, twelve computer-assisted and interviewer-based 24HDR (three in each season) were conducted on all weekdays and weekend days. The assessment of foods and recipes in this software programme was standardised and structured by meals and eating occasions during the day. To ease the estimation of the portion size of consumed foods, coloured photographs of different portion sizes were provided⁽¹⁷⁾.

Assessment of anthropometry

All anthropometric measures were obtained by trained personnel in standardised procedures during the visit in the study centre once at the beginning of the validation study and for a second time after 1 year. Body weight was measured using a digital scale to the nearest 100 g and body height to the nearest mm. BMI was calculated as body weight in kg divided by the squared body height in metres for both time points⁽¹⁴⁾. Both anthropometric measures were obtained to investigate their stability between the application of both FFQ and its potential influence on the stability of dietary intake.

Statistical analysis

All statistical analyses were performed with the software packages Statistical Analysis System (SAS) Enterprise Guide 7.1 with SAS version 9.4 (SAS Institute Inc.).

Dietary pattern construction

Selected DP, based on the dietary data of the FFQ, were constructed. For this purpose, the 146 single food items were assigned to thirty-nine respective food groups. Exploratory DP were derived by applying PCA on these food groups. First, principal components were retained by using the eigenvalue > 1 criterion (in PCA each food group has a variance of 1) to reduce the dimensionality. Second, the scree plot visualises a possible 'scree' in the top-to-ground slope and principal components were identified above the scree as explaining the majority of variance. Third, the interpretability criterion (defined as ≥ 3 food groups with absolute factor loadings ≥ 0.4 , as previously described⁽¹⁸⁾) will retain only those principal components, which indeed reflect the complexity of a DP.

Two different conceptual constructs were considered for the analysis on predefined dietary indices. The AHEI is an example for an index to be claimed as healthful, as published by Kröger *et al.* within the InterAct consortium including our recent study population⁽¹⁹⁾ (online Supplementary Table S1). The Mediterranean diet as a regional diet was investigated as tMDS⁽⁷⁾, where scoring criteria were slightly modified according to the non-Mediterranean population⁽²⁰⁾ (online Supplementary Table S2). Second, the MedPyr was derived based on predefined criteria from recommendations on fifteen foods and food groups categorised to be consumed either in high, moderate or low

quantities in accordance to the algorithm developed by Tong *et al.*^(8,20) (online Supplementary Table S3).

Data analysis

Characteristics of the participants at baseline and after 1 year were described as medians with interquartile ranges and as percentages for categorical variables. To compare the exploratory DP between the two FFQ and the 24HDR, the explained variance in food groups by the DP and the contributing food groups with their respective factor loadings was shown. To potentially explain the deviant structure of the exploratory DP, the single thirty-nine food groups (described in detail in online Supplementary Table S4) were also compared between the different instruments. The median intake of the food groups as measured by the FFQ at baseline (FFQ_b), by the FFQ after 1 year (FFQ₁) and the mean of twelve 24HDR (mHDR) were calculated. The median difference, the mean absolute deviation from the median differences and the Spearman correlation coefficients between both FFQ (reliability) and between FFQ₁ and mHDR (validity) were calculated for all food groups in accordance to a previous analysis by Bohlscheid *et al.*⁽¹¹⁾. Deattenuated correlation coefficients were provided to correct for the intra-individual variation between the twelve 24HDR with the following formula:

$$r_{\text{deatt}} = r_0 \times \sqrt{1 + \gamma/k},$$

where r_{deatt} is the corrected correlation coefficient, r_0 is the uncorrected correlation coefficient, γ is the ratio of the estimated within-person to between-person variation in the 24HDR and k is the number of 24HDR⁽²¹⁾. For the calculation of γ , we used the SAS macro provided by Lu *et al.*⁽²²⁾.

To investigate the potential influence of energy adjustment on the DP structure and on the reliability and validity of these DP, an additional PCA was performed with food groups, adjusted for total energy intake by the residual method⁽²³⁾.

For the three determined predefined dietary indices, mean difference and its standard deviation were calculated for both FFQ and mHDR. Spearman rank correlation coefficients were calculated, which were again also provided as deattenuated correlation coefficients. The three indices were divided into quintiles, and it was investigated how many of the participants stayed in the same quintile were in the adjacent or opposite quintile, when comparing the two FFQ and mHDR. Cohen's weighted κ gives an additional parameter, how well the indices are in accordance in these comparisons.

To investigate if the DP, derived by the study instrument, systematically differ from the reference instrument and if the variance is dependent on the level of the DP, we applied Bland–Altman plots for the exploratory DP scores and the three predefined indices.

Results

Median age of the participants in this validation study was 59.2 years for men and 55.5 years for women. In men and women, the body weight at baseline remained almost the same

Table 1. Sex-specific characteristics of the included study population (*n* 134) at baseline and after the validation study period (1 year) (Median values and interquartile ranges (IQR); numbers and percentages)

	Male (<i>n</i> 75)		Female (<i>n</i> 59)	
	Median	IQR	Median	IQR
Age at baseline (years)	59.24	51.17, 63.62	55.53	47.59, 59.26
BMI at baseline (kg/m ²)	26.67	24.42, 28.99	25.56	22.60, 27.50
BMI after 1 year (kg/m ²)	26.44	24.44, 28.92	25.82	22.75, 27.68
Weight at baseline (kg)	81.80	73.70, 89.30	65.80	59.90, 71.90
Weight after 1 year (kg)	81.60	73.50, 89.10	66.20	60.00, 71.60
Physical activity (h/week)*	1.50	0.00, 3.00	1.50	0.00, 3.50
	<i>n</i>	%	<i>n</i>	%
Education				
No vocational/vocational training	15	20.00	21	35.59
Technical college	16	21.33	21	35.59
University	44	58.67	17	28.81

* Physical activity is the sum of information on sports (h/week) and biking (h/week).

at the end of the study period (men 81.8 *v.* 81.6 kg; women 65.8 *v.* 66.2 kg). Accordingly, the median BMI was also stable over the time of 1 year. The highest percentage of men had a university degree, while for women it was the minority. Both men and women had a rather low physical activity with a median of 1.5 h per week (Table 1).

Reliability of exploratory dietary patterns

The application of PCA on the dietary data from FFQ_b resulted in two DP scores, explaining in total 19.4 % of variance in the thirty-nine food groups (details of food items belonging to the food groups in online Supplementary Table S4). The first DP score was characterised by high factor loadings of different vegetables, fruits, cereals, vegetable oils and the food group miscellaneous. The second DP score had high contributions of potatoes, red and processed meat, offals, butter, beer and sauces. Again, after 1 year, two DP scores were derived from FFQ₁ data which explained similar variance (18.8 %) as the two pattern scores at baseline. Vegetables, potatoes, legumes, poultry, fish, vegetable oils and soups were highly loading on the first pattern score. The second pattern score was predominantly characterised by high negative loadings of leafy and root vegetables, fruits, milk and dairy products and cereals, while potatoes, red meat, vegetable oils, beer, spirits and sauces were positively loading (Table 2). Comparing PCA results between FFQ_b and FFQ₁, the first derived DP score differed in high loading food groups and was only comparable in contributions of vegetables and vegetable oils. This was reflected in a rather moderate correlation (*r* 0.45). The second DP showed a more comparable structure between FFQ_b and FFQ₁ characterised by high loadings of potatoes, red meat, beer and sauces. The correlation was higher than for the first DP, although still moderate (*r* 0.58) (Table 3).

Validity of the exploratory dietary patterns

For mHDR dietary data, the PCA resulted in one DP score, which explained 10.5 % variance in the food groups. This DP score was characterised by high loadings of potatoes, other vegetables (e.g. mushrooms, garlic, onions and mixed salads), bread, red and processed meat, margarine and other fats, beer and spirits and a high negative loading of milk and dairy products (Table 2). The comparison of the first DP score from FFQ₁ with the DP score derived by mHDR showed only few comparable food groups like other vegetables, fish and soups and the correlation between both patterns was rather low (*r* 0.30) (Table 3). The second DP score derived by FFQ₁ showed more consistent high loading food groups with the DP derived by mHDR, specifically for potatoes, milk and dairy products, red meat, processed meat, other fats, coffee, beer and spirits. This was reflected by a higher correlation (*r* 0.70) (Table 3).

The Bland–Altman plots for the two derived DP scores did not indicate a systematic deviation for the difference between the two dietary assessment instruments with higher DP score points (online Supplementary Figs. S1 and S2).

Sensitivity analyses on exploratory dietary patterns

When we repeated the PCA using energy-adjusted food groups, a similar number of DP scores were retained. However, the energy-adjusted DP scores differ extensively in their pattern structure compared with the unadjusted DP scores (online Supplementary Table S5). The reliability of the energy-adjusted DP was lower than for the unadjusted DP scores (*r* −0.20; *r* 0.12). The validity for the first DP score by FFQ₁ against the DP score by mHDR was moderate (*r* 0.56) and therefore higher than for the unadjusted DP scores. However, the good relative validity of the second DP score disappeared after energy adjustment (*r* 0.01) (Table 3).

The investigation of the reproducibility of the thirty-nine food groups in the FFQ showed for the majority of food groups no statistically significant median differences (online Supplementary Table S6). The median difference from fourteen out of thirty-nine food groups was significantly different from zero, for example, for potatoes, sugar and cake and cookies. The mean deviation of median differences was very high in fruits (238 g) and sugar (173 g) in comparison with their mean intake of both FFQ, which indicates a high variability among individual differences. Correlations for the majority of food groups were moderate (*r* 0.50–0.70). For almost all beverages in particular, correlations were fairly good (coffee: *r* 0.73) to good (beer: *r* 0.93). The lowest correlations were observed for cheese (*r* 0.24) and vegetable oils (*r* 0.22), although the significance level was not achieved (online Supplementary Table S7).

The investigation of the validity of thirty-nine food groups resulted in a median difference of seventeen food groups, significantly different from zero, with a range from 0 to 94 g. Of these seventeen food groups, five had negative differences, suggesting that the FFQ₁ estimates a significant higher intake of nuts, bread, butter, fruit and vegetable juices and wine than the mHDR. The highest median differences were observed for fruits (51.8 g) and coffee (93.6 g). The mean deviations of median differences were



Table 2. Dietary patterns of FFQ at baseline (FFQ_b), FFQ after 1 year (FFQ₁) and mean of twelve 24-h dietary recalls (mHDR) and the respective factor loadings* for the thirty-nine food groups

	FFQ _b		FFQ ₁		mHDR
	DP score 1	DP score 2	DP score 1	DP score 2	DP score 1
Explained variance	10.42 %	8.96 %	10.15 %	8.69 %	10.21 %
Potatoes	–	59	45	46	50
Leafy vegetables	64	–	49	–45	–26
Fruiting vegetables	59	–	27	–37	–
Root vegetables	65	–	52	–43	–
Cabbages	37	–	53	–	–
Other vegetables	54	36	71	–	48
Legumes	–	37	47	–	24
Fruits	49	–	22	–45	–21
Nuts	–	–	–	–25	–27
Other fruits	31	37	26	–	–
Milk and dairy products	36	–	–	–40	–47
Cheese	–	20	–	–	22
Pasta, rice	37	–	39	–	–
Bread	–	34	–	–	42
Other cereals	45	–	–	–52	–25
Red meat	–	70	58	49	59
Poultry	28	–	30	–	26
Processed meat	24	42	22	31	64
Offal	–	56	–	29	–
Fish	–	–	41	–	39
Eggs	–	26	–	–	–
Vegetable oils	66	–21	49	49	–
Margarines	–	20	–	–	46
Butter	–	40	–	24	–
Other fats	–	–	–	32	51
Sugar	–	24	–	32	–
Cakes, cookies	–	–	–	–	–
Fruit and vegetable juices	32	–	–	–	–
Soft drinks	–	31	–	–	–
Coffee	–	–	–	34	34
Tea	–	–	–	–	–
Other non-alcoholic drinks	20	–29	–	–32	–29
Wine	–	–26	–	–	–
Beer	–24	40	24	50	66
Spirits	–23	33	–	42	51
Other alcoholic beverages	–	–	–	–	–
Sauces	–	66	36	54	31
Soups	37	39	65	–	30
Miscellaneous	52	–	27	–	22

* Factor loadings < |20| are not shown.

Table 3. Dietary pattern (DP) scores derived by FFQ at baseline (FFQ_b), FFQ after 1 year (FFQ₁) and mean of twelve 24-h dietary recalls (mHDR) (Spearman rank correlation coefficients and 95 % confidence intervals)

	DP score 1		DP score 2	
	<i>r</i>	95 % CI	<i>r</i>	95 % CI
Unadjusted				
FFQ _b v. FFQ ₁	0.45	0.31, 0.58	0.58	0.46, 0.68
FFQ ₁ v. mHDR	0.30	0.15, 0.45	0.70	0.59, 0.77*
Energy adjusted				
FFQ _b v. FFQ ₁	–0.20	–0.36, –0.03	0.12	–0.05, 0.28
FFQ ₁ v. mHDR	0.56	0.43, 0.66	0.01	–0.16, 0.18*

* DP score 2 (FFQ₁) v. DP score 1 (mHDR).

high for milk and dairy products (103 g), bread (99 g) and beer (95 g) (online Supplementary Table S6). The correlations were not as high as in the reproducibility analysis. Still, potatoes, fruits, milk and dairy products, other cereals, processed meat, margarine, butter and many beverages showed moderate to good

correlations ($r_{\text{deatt}} > 0.60$). The lowest correlations were observed for vegetable oils ($r_{\text{deatt}} -0.01$), but also for 'other fruits' (e.g. fruit salad and olives) ($r_{\text{deatt}} -0.04$) (online Supplementary Table S7).

Reliability of predefined indices

The mean AHEI derived with FFQ_b data was 25.2 (SD 5.8) points of a maximum achieved 43 points, with women achieving more points than men. The theoretical maximum of seventy points for this score was not achieved in this population. The mean difference to the FFQ₁ for all participants was –2.9 points and was slightly higher in women than in men. The correlation in all participants was moderate and higher in men ($r 0.62$) than in women ($r 0.59$) (Table 4).

The mean of tMDS was slightly lower for the FFQ₁ compared with the FFQ_b, but the mean difference was less than one scoring point out of a maximum of fourteen points (from eighteen possible points) in all participants. In men, the mean difference was



Table 4. Alternative Healthy Eating Index (AHEI), traditional Mediterranean Diet Score (tMDS) and Mediterranean Diet Pyramid Index (MedPyr) between FFQ at baseline (FFQ_b), FFQ after 1 year (FFQ₁) and the mean of twelve 24-h dietary recalls (mHDR) (*n* 134) (Mean values, mean differences, standard deviations and correlations)

	FFQ _b		FFQ ₁		mHDR		FFQ _b v. FFQ ₁		mHDR v. FFQ ₁		FFQ _b v. FFQ ₁		mHDR v. FFQ ₁	
	Mean	SD	Mean	SD	Mean	SD	Mean difference	SD	Mean difference	SD	<i>r</i>		<i>r</i>	<i>r</i> _{deatt}
AHEI (Max 70P)														
All	25.20	5.77	28.05	6.09	26.73	4.98	−2.85	5.16	−1.32	5.16	0.62		0.58	0.60
Men	24.16	5.68	26.84	6.11	25.77	5.32	−2.68	5.18	−1.07	5.63	0.62		0.52	0.54
Women	26.53	5.66	29.59	5.76	27.95	4.25	−3.07	5.17	−1.64	4.53	0.59		0.63	0.64
tMDS (Max 18P)														
All	9.09	2.55	8.96	2.81	8.43	2.50	0.13	2.67	−0.53	2.86	0.51		0.38	0.40
Men	9.21	2.40	9.01	2.52	8.17	2.53	0.20	2.54	−0.84	2.95	0.47		0.32	0.34
Women	8.93	2.74	8.90	3.16	8.76	2.44	0.03	2.85	−0.14	2.71	0.54		0.56	0.60
MedPyr (Max 15P)														
All	6.59	1.28	6.56	1.27	6.25	1.10	0.03	1.11	−0.31	1.27	0.62		0.43	0.45
Men	6.57	1.22	6.51	1.24	6.13	1.12	0.06	1.10	−0.37	1.31	0.60		0.39	0.41
Women	6.61	1.36	6.62	1.32	6.40	1.07	−0.01	1.14	−0.22	1.23	0.64		0.49	0.52

r, Spearman rank correlation; *r*_{deatt}, correlation coefficient corrected for intra-individual variation between the 24HDR; tMDS, traditional Mediterranean Diet Score; MedPyr, Mediterranean Diet Pyramid Index.

Table 5. Agreement to the quintiles of the Alternative Healthy Eating Index (AHEI), traditional Mediterranean Diet Score (tMDS) and Mediterranean Diet Pyramid Index (MedPyr) between FFQ at baseline (FFQ_b), FFQ after 1 year (FFQ₁) and mean of twelve 24-h dietary recalls (mHDR) (*κ* Values; 95 % confidence intervals)

	Changes in dietary quality indices quintiles								κ^*	95 % CI	
	Lower adjacent quintile	%	No change	%	Higher adjacent quintile	%	Opposite quintiles	%			
AHEI											
FFQ _b v. FFQ ₁	30	22.4	43	32.1	25	18.7	2	1.5	0.35	0.24, 0.46	
FFQ ₁ v. mHDR	30	22.4	47	35.1	24	17.9	1	0.8	0.37	0.25, 0.48	
tMDS											
FFQ _b v. FFQ ₁	23	17.2	41	30.6	30	22.4	3	2.2	0.32	0.21, 0.43	
FFQ ₁ v. mHDR	24	17.9	42	31.4	31	23.1	3	2.2	0.32	0.21, 0.43	
MedPyr											
FFQ _b v. FFQ ₁	33	24.6	49	36.6	28	20.9	2	1.5	0.46	0.36, 0.55	
FFQ ₁ v. mHDR	17	12.7	45	33.6	31	23.1	4	3.0	0.30	0.18, 0.41	

* *κ* = Cohen's weighted *κ* 0.21–0.40 'sufficient accordance', 0.41–0.60 'moderate accordance'.

higher than in women, which was also reflected by a lower correlation in men (*r* 0.47) than in women (*r* 0.54) (Table 4).

The mean of MedPyr derived at baseline and after 1 year was similar (6.6 points) out of an achieved maximum of 9.93 points. Overall, a theoretical maximum of fifteen points was not achieved in this study population. The mean difference for the MedPyr score between FFQ_b and FFQ₁ was very low and lower in women than in men. The correlation was moderate (*r* 0.62) and higher in women than in men (*r* 0.64 *v.* 0.60) (Table 4).

We furthermore evaluated the agreement to the quintiles for the comparison of the two FFQ. It revealed a higher proportion of study participants in the 'no change' group of the MedPyr (36.6 %) compared with the tMDS (30.6 %) and AHEI (32.1 %) (Table 5). This was also confirmed by the Cohen's weighted *κ*, which indicates a moderate accordance for MedPyr (*κ* 0.46, 95 % CI 0.36, 0.55). Still, the majority of participants were grouped in the adjacent quintiles, with a higher proportion of participants being in the lower adjacent quintile for the AHEI and MedPyr after 1 year. Contrarily, according to the tMDS, more participants were assigned to the higher adjacent quintile. A severe

misclassification in the opposite quintiles was only the case in a minor proportion (1.5–2.2 %) of the participants for all three indices.

Validity of predefined indices

The mean difference of the AHEI for the comparison of the FFQ₁ with mHDR was negative and highest in women (−1.6 P) and the correlation was moderate, indicating a reasonable relative validity of the FFQ (Table 4). The Bland–Altman plot showed that the FFQ₁ overestimated the adherence to AHEI with higher score points (>35 points) in comparison with the reference instrument mHDR (Fig. 1).

The relative validity of the tMDS, evaluated by the mean difference between the FFQ₁ and the mHDR, which ranged from −0.84 for men and −0.14 for women, indicated sex-specific differences. This was also reflected in a higher correlation in women (*r*_{deatt} 0.60) than in men (*r*_{deatt} 0.34) (Table 4).

Similar to tMDS, a slightly higher score was found for the MedPyr derived by FFQ₁ than the mHDR and the correlation indicated a rather low relative validity (*r*_{deatt} 0.45). Again,

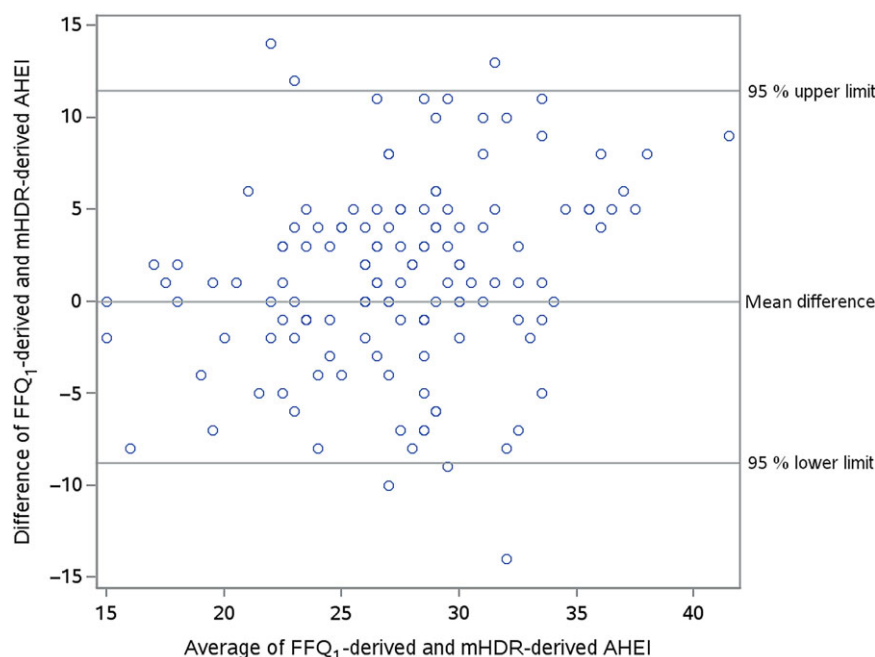


Fig. 1. Bland–Altman plot of the AHEI FFQ₁ v. mHDR. AHEI, Alternative Healthy Eating Index; FFQ₁, FFQ applied after 1 year; mHDR, mean of twelve 24-h dietary recalls.

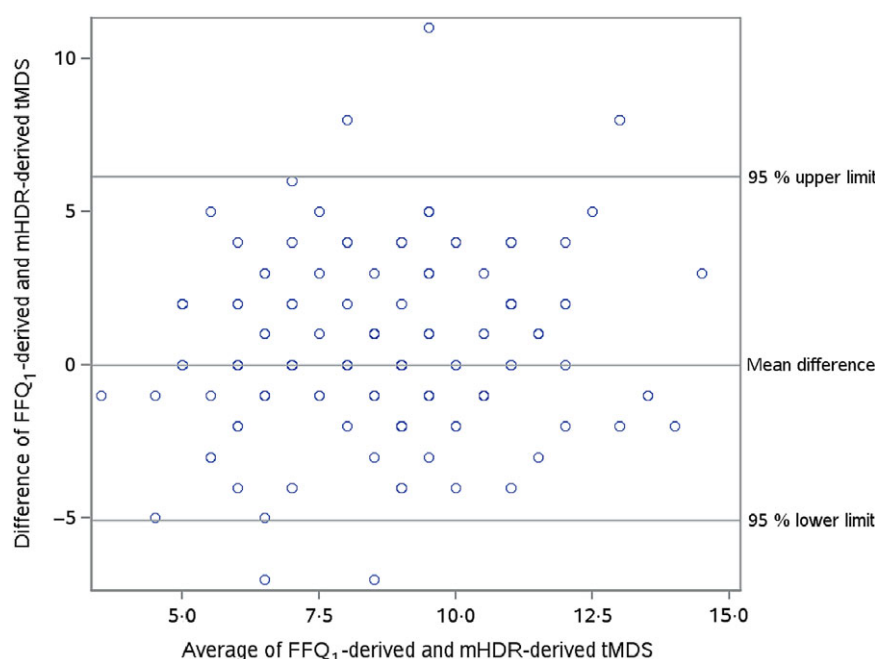


Fig. 2. Bland–Altman plot of the tMDS FFQ₁ v. mHDR. tMDS, traditional Mediterranean Diet Score; FFQ₁, FFQ applied after 1 year; mHDR, mean of twelve 24-h dietary recalls.

correlations in women were slightly higher than in men for the validity (r_{deatt} 0.52 *v.* 0.41) (Table 4).

For both Mediterranean diet scores, the Bland–Altman plots did not suggest a systematic deviation of the difference between the FFQ₁ and mHDR in dependence of index values (Figs. 2 and 3).

According to the agreement to the quintiles, most of the participants were classified into adjacent quintiles, but a considerable proportion were classified into the same quintile of

the AHEI (35.1 %), followed by the MedPyr (33.6 %) and the tMDS (31.4 %) (Table 5). For both indices describing the Mediterranean diet, a higher proportion of participants were classified into the higher adjacent quintile by mHDR data compared with FFQ data. The opposite was observed for the AHEI. For all three indices, a very small proportion of participants were classified into the opposite quintiles, ranging from 0.8 to 3.0 %. The κ as a measure of accordance was overall lower than in the

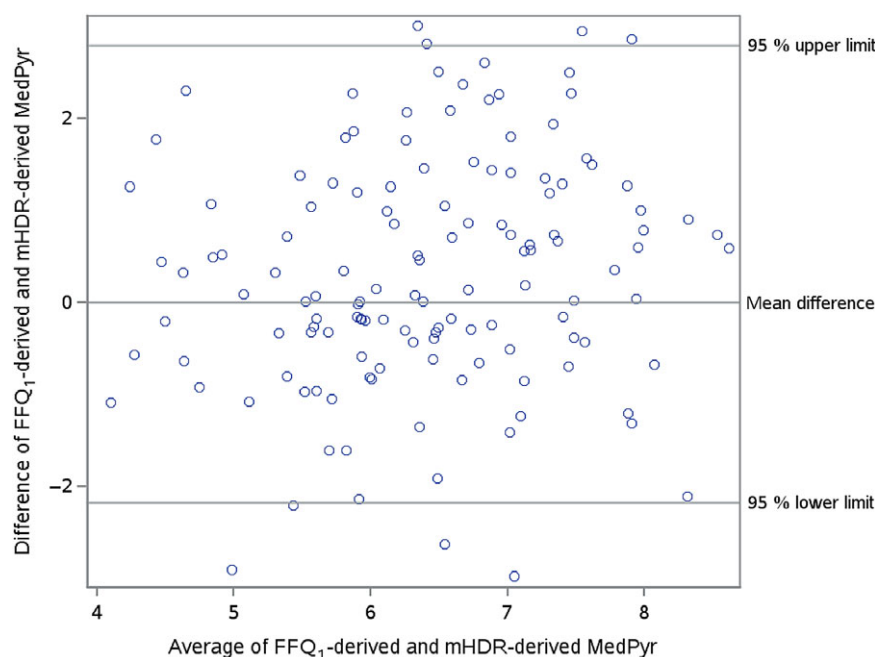


Fig. 3. Bland–Altman plot of the MedPyr FFQ₁ v. mHDR. MedPyr, Mediterranean Diet Pyramid Index; FFQ₁, FFQ applied after 1 year; mHDR, mean of twelve 24-h dietary recalls.

reliability investigation of the three indices and highest for the AHEI (κ 0.37, 95 % CI 0.25, 0.48).

Discussion

The aim of the current analysis was to investigate how valid and reliable the semi-quantitative FFQ as dietary intake assessment instrument can assess DP in the EPIC-Potsdam study. The reliability was assessed by the repeated application of the same FFQ within the study period of 1 year. The relative validity was examined by the comparison of DP scores from an FFQ with DP scores from the mHDR as reference instrument. In terms of reliability, the exploratory DP scores differed considerably according to their structure of contributing food groups and resulted in a moderate correlation, although the second DP score seemed to be more comparable than the first derived DP score. The relative validity of the first DP score by FFQ₁ was very low, but the second DP score showed a good relative validity. The predefined dietary indices AHEI, tMDS and MedPyr showed rather small differences between the two FFQ and resulted in moderate to good reliability. The relative validity of the three indices was reflected by lower, yet still moderate correlations than in the reliability analysis.

Reliability of exploratory dietary patterns

A former analysis on exploratory DP in 1184 randomly selected participants of EPIC-Potsdam revealed two DP, which were similar in structure of high loading food groups compared with the two patterns derived with the FFQ_b in this analysis⁽¹⁸⁾. Hence, although a limited sample of 134 participants was used in the recent analysis, the structure of the two baseline DP seems to be generalisable for this study population. Compared with

existing studies on exploratory DP, the reliability of the FFQ (r 0.45–0.58) to assess DP scores was considerably weaker. A potential influence of weight changes during the study period on the stability of the DP scores could be ruled out, because weight remains stable in both men and women. Correlation coefficients in most previous studies ranged between 0.63 and 0.87^(24–27), pointing towards a moderate to good reliability. However, Nanri *et al.* showed a comparable reliability of DP scores (r 0.55–0.56) to our results⁽²⁸⁾. A possible explanation for the weaker reliability than in other studies could be the deviant structure of contributing food groups in the two DP scores between the FFQ. Furthermore, certain food groups like vegetables and vegetable oils, which are mainly contributing to the first DP score in the FFQ, had a rather low reliability (cabbages: r 0.37, root vegetables: r 0.40, vegetable oils: r 0.22) in comparison with other food groups. Those food groups, contributing to the second DP score in both FFQ, were characterised by a moderate to good reliability (potatoes: r 0.67, red meat: r 0.61, beer: r 0.93, sauces: r 0.62), which was also reflected by a higher reliability of the second DP score compared with the first DP score.

Validity of exploratory dietary patterns

In other studies, the validity of exploratory DP was assessed with diverse reference instruments, using either 24HDR^(24,25) or dietary records covering different numbers of days^(26–30). Studies using the same reference instrument as in our analysis showed a wide range of correlation coefficients (r 0.30–0.70)^(24,25), which cover our results (first DP score: r 0.30; second DP score: r 0.70). While the validity of the two DP in the Health Professional Follow-Up Study⁽²⁶⁾ was moderate to good (r 0.52–0.74), Nanri *et al.*⁽²⁸⁾ reported rather low to moderate correlations (r 0.32–0.49 in men; r 0.36–0.63 in women). Ambrosini



et al.⁽²⁹⁾ concluded that adjusting the correlations for total energy intake resulted in a stronger validity. However, in our analysis, the comparability of the DP score structure did not improve, when the PCA was applied to energy-adjusted food groups and correlations between FFQ-derived and mHDR-derived DP scores were weaker than for the unadjusted ones. A possible explanation for the weak correlations of the exploratory DP, in comparison with other studies, could be the rather low explained variance in the food groups. While the two DP derived with data from the FFQ₁ explained in total 18.8 % of variance, DP from other studies explained 22–84 %^(25,27,29,30), resulting partly in moderate to good reliability and validity⁽²⁵⁾. However, explained variance is highly dependent on the number of DP scores to be retained and also on the number of food groups, included in the analysis. Hence, a comparison with other studies is impeded by these influences. Another limitation of the PCA, possibly explaining the rather low validity, is the retainment of deviant numbers of DP scores. This was the case in the comparison with the FFQ₁-derived DP scores with the mHDR-derived DP score in our analysis.

Reliability of predefined indices

Data on the reliability of dietary assessment instruments measuring predefined indices are sparse, which impede the evaluation of our results in comparison with other studies. The comparison of indices describing Mediterranean diet in other studies was hampered by different assessment instruments like a short diet screener (Mediterranean Diet Adherence Screener (MEDAS)) instead of an FFQ. The reliability was measured with the MEDAS applied 1 month apart and resulted in a moderately good correlation (r 0.69, P < 0.001) and a fair agreement (κ 0.38)⁽³¹⁾. Weaker correlations in our analysis could be due to the twelve times longer interval between both FFQ applications. Furthermore, it is not surprising that a short screener developed to assess the Mediterranean diet performs better than a FFQ with a longer list of inquired food items. Another study investigated the AHEI-2010 and alternative Mediterranean diet in a multi-ethnic Asian population to assess the reliability of a short thirty-seven-item diet screener (DS)⁽⁹⁾. The application of two DS within a 4-month interval resulted in a better reliability (AHEI: intra-class correlation (ICC) = 0.69; alternative Mediterranean diet: ICC = 0.71) than it was observed in our analysis. Again, this could be due to the shorter interval between the applications of the DS. However, comparability between a DS and FFQ remains constrained.

Since the tMDS was dependent on the distribution of food groups in the study population by scoring their respective tertiles, the derived scores by the FFQ_b and FFQ₁ were not the same constructs. To address this limitation, a confirmatory approach of using tMDS tertiles from FFQ_b on the dietary data of FFQ₁ was done in a sensitivity analysis (online Supplementary Table S8). In this case, the mean of tMDS in the FFQ₁ was higher than the mean of tMDS by FFQ_b, which was also reflected by negative mean differences. In the original analysis, the mean difference was higher in men, while in the confirmatory analysis, the difference is higher in women. The correlation coefficient did not improve for men, but was slightly improved for women

(r 0.54 *v.* r 0.59). Overall, the results were only marginally improved in comparison with the original analysis. Hence, the ability of the FFQ to reliably describe the tMDS was not affected.

Validity of predefined indices

The validity of the tMDS was rather low (r_{deatt} 0.40) in this population. In a Spanish validation study in 107 participants, FFQ-derived Mediterranean diet scores were compared with 24HDR. The correlations for a modified Mediterranean diet score resulted in similarly low values (r 0.33)⁽¹⁰⁾. This score was comparable with the tMDS in our analysis. Another investigated score in this study, the Mediterranean-like diet score, resulted in a slightly better correlation (r 0.42)⁽¹⁰⁾. It was partly comparable in its components to the MedPyr, which also resulted in a better correlation (r_{deatt} 0.45) in our study.

Another study investigated the validity of a fourteen-item English version of the MEDAS via the comparison with 3-d food records⁽³¹⁾. The derived Mediterranean diet score correlated slightly better (r 0.50) than the tMDS (r_{deatt} 0.40) and MedPyr (r_{deatt} 0.45) in the current analysis. However, compared with our FFQ, the MEDAS is a short assessment instrument with the intention to exclusively measure components of the Mediterranean diet.

In a multi-ethnic Asian population, the mean DS was compared with a 163-item FFQ as a reference instrument, resulting in moderate correlations for the indices (AHEI: r 0.51; alternative Mediterranean diet: r 0.50) and moderate agreement (κ 0.48–0.58)⁽⁹⁾. However, a comparison with our study results remains limited, since the study instrument and the reference instrument in our analysis differ from the instruments used in this study.

We would expect that several strengths and limitations in our study contributed to the observed results. While the 24HDR can serve as a promising reference instrument, because multiple applications of short-term measurements over a year could account for seasonally consumed food groups, it also has some limitations. So, for example, based on the repeated collection of one day's dietary intake, a certain food can or cannot be consumed due to the individual consumption frequency. Hence, the proportion of participants reporting non-consumption could be higher than for FFQ data. In a former analysis by Hoffmann *et al.*⁽³²⁾, it was also observed that although mean intake estimates of twelve 24HDR were comparable to more sophisticated methods (e.g. Buck or S-Nusser method), percentiles, standard deviation and skewness differed between the methods, implying that individual consumption estimates can differ. To address these limitations, a more sophisticated method like the National Cancer Institute (NCI) method would result in a different distribution and improved estimates of usual intake⁽³³⁾. In the NCI method, the information from an FFQ can be used to enhance pseudo-individual estimates of usual dietary intake. The FFQ information can be used as an adjustment variable. The days with and without consumption are included in the estimate, that is, not only the amount consumed is considered⁽³⁴⁾. Hence, we would expect in our analysis higher correlations between the food groups, possibly resulting in a different pattern structure of exploratory DP and in a better relative validity. This could also apply to the relative validity of the predefined DP.



Nevertheless, we decided to use the twelve 24HDR as reference instrument due to two reasons: on the one hand to ensure the comparability to other study results and on the other hand to use a reference instrument being independent of the study instrument. However, for future analyses on usual intake estimates, it is reasonable to choose the NCI method over simple averaging the repeated 24HDR due to the advantage of combining different sources of information and considering that in most observational studies, the number of available recalls is restricted to 2–4 recalls rather than twelve recalls as in our validation study setting.

Conclusion

To conclude, the application of PCA as an exploratory method to dietary data from the FFQ yielded two DP scores with moderate reliability over a 12-month period in this study population. Therefore, when changes of dietary habits will be investigated by using exploratory DP, the distinction will be impeded, if deviations in DP are either the result of real changes in diet over time or if they are the result of measurement error in this study instrument. Since this is highly dependent on the reliability of those food groups contributing most to the DP scores, we would recommend to always investigate the study instrument with regard to the reliability of single food groups in addition to the reliability of exploratory DP before analyses of long-term changes in diet will be undertaken. In contrast, the three examined predefined dietary indices showed moderate to good reliability. In terms of validity, FFQ-based exploratory DP differed in numbers and composition to those based on repeated 24HDR, but the second DP score showed a good validity. Reasonable validity was also observed for the three predefined dietary indices. Comparing the two investigated indices reflecting the Mediterranean diet, the MedPyr had a higher reliability and validity than the tMDS in this middle-aged semi-urban German study population.

Acknowledgements

The authors would like to thank the participants of the EPIC-Potsdam study, who agreed to take part in this validation study. Furthermore, the authors would like to thank E. Kohlsdorf to provide the data for this analysis and M. Bergmann as the head of the Human Study Center for her support with additional information on the data collection and the revision of the manuscript draft.

This work was supported by a grant from the German Federal Ministry of Education and Research (BMBF) to the German Center for Diabetes Research (DZD) and the State of Brandenburg. Furthermore, this work was supported by the NutriAct – Competence Cluster Nutrition Research Berlin-Potsdam funded by the German Federal Ministry of Education and Research (FKZ: 01EA1806A). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

F. J. analysed the data, and planned, drafted and revised the paper. D. N. contributed to the technical review and critically revised and edited the paper. M. B. S. contributed to conceptualising and planning, and drafted and revised the paper.

There are no conflicts of interest.

Supplementary material

For supplementary materials referred to in this article, please visit <https://doi.org/10.1017/S0007114520003517>

References

1. Hu FB (2002) Dietary pattern analysis: a new direction in nutritional epidemiology. *Curr Opin Lipidol* **13**, 3–9.
2. Jannasch F, Kroger J & Schulze MB (2017) Dietary patterns and type 2 diabetes: a systematic literature review and meta-analysis of prospective studies. *J Nutr* **147**, 1174–1182.
3. Mozaffarian D (2016) Dietary and policy priorities for cardiovascular disease, diabetes, and obesity: a comprehensive review. *Circulation* **133**, 187–225.
4. Jannasch F, Riordan F, Andersen LF, *et al.* (2018) Exploratory dietary patterns: a systematic review of methods applied in pan-European studies and of validation studies. *Br J Nutr* **120**, 601–611.
5. Milajerdi A, Namazi N, Larijani B, *et al.* (2018) The association of dietary quality indices and cancer mortality: a systematic review and meta-analysis of cohort studies. *Nutr Cancer* **70**, 1091–1105.
6. Galbete C, Schwingshackl L, Schwedhelm C, *et al.* (2018) Evaluating Mediterranean diet and risk of chronic disease in cohort studies: an umbrella review of meta-analyses. *Eur J Epidemiol* **33**, 909–931.
7. Trichopoulou A, Costacou T, Bamia C, *et al.* (2003) Adherence to a Mediterranean diet and survival in a Greek population. *N Engl J Med* **348**, 2599–2608.
8. Tong TY, Wareham NJ, Khaw KT, *et al.* (2016) Prospective association of the Mediterranean diet with cardiovascular disease incidence and mortality and its population impact in a non-Mediterranean population: the EPIC-Norfolk study. *BMC Med* **14**, 135.
9. Whitton C, Ho JCY, Rebello SA, *et al.* (2018) Relative validity and reproducibility of dietary quality scores from a short diet screener in a multi-ethnic Asian population. *Public Health Nutr* **21**, 2735–2743.
10. Benitez-Arciniega AA, Mendez MA, Baena-Diez JM, *et al.* (2011) Concurrent and construct validity of Mediterranean diet scores as assessed by an FFQ. *Public Health Nutr* **14**, 2015–2021.
11. Bohlscheid-Thomas S, Hoting I, Boeing H, *et al.* (1997) Reproducibility and relative validity of food group intake in a food frequency questionnaire developed for the German part of the EPIC project. European Prospective Investigation into Cancer and Nutrition. *Int J Epidemiol* **26**, Suppl. 1, S59–S70.
12. Bohlscheid-Thomas S, Hoting I, Boeing H, *et al.* (1997) Reproducibility and relative validity of energy and macronutrient intake of a food frequency questionnaire developed for the German part of the EPIC project. European Prospective Investigation into Cancer and Nutrition. *Int J Epidemiol* **26**, Suppl. 1, S71–S81.
13. Boeing H, Korfmann A & Bergmann MM (1999) Recruitment procedures of EPIC-Germany. European Investigation into Cancer and Nutrition. *Ann Nutr Metab* **43**, 205–215.
14. Kroke A, Klipstein-Grobusch K, Voss S, *et al.* (1999) Validation of a self-administered food-frequency questionnaire administered in the European Prospective Investigation into Cancer and Nutrition (EPIC) Study: comparison of energy, protein, and macronutrient intakes estimated with the doubly labeled water, urinary nitrogen, and repeated 24-h dietary recall methods. *Am J Clin Nutr* **70**, 439–447.



15. Riboli E (1992) Nutrition and cancer: background and rationale of the European Prospective Investigation into Cancer and Nutrition (EPIC). *Ann Oncol* **3**, 783–791.
16. Kroke A, Bergmann MM, Lotze G, *et al.* (1999) Measures of quality control in the German component of the EPIC study. European Prospective Investigation into Cancer and Nutrition. *Ann Nutr Metab* **43**, 216–224.
17. Voss S, Charondiere UR, Slimani N, *et al.* (1998) EPIC-SOFT a European computer program for 24-hour dietary protocols. *Z Ernahrungswiss* **37**, 227–233.
18. Jannasch F, Kroger J, Agnoli C, *et al.* (2019) Generalizability of a diabetes-associated country-specific exploratory dietary pattern is feasible across European populations. *J Nutr* **149**, 1047–1055.
19. InterAct C (2014) Adherence to predefined dietary patterns and incident type 2 diabetes in European populations: EPIC-InterAct Study. *Diabetologia* **57**, 321–333.
20. Galbete C, Kroger J, Jannasch F, *et al.* (2018) Nordic diet, Mediterranean diet, and the risk of chronic diseases: the EPIC-Potsdam study. *BMC Med* **16**, 99.
21. Rosner B & Willett WC (1988) Interval estimates for correlation coefficients corrected for within-person variation: implications for study design and hypothesis testing. *Am J Epidemiol* **127**, 377–386.
22. Lu L & Shara N (2007) Reliability analysis: calculate and compare intra-class correlation coefficients (ICC) in SAS2007. <https://lexjansen.com/nesug/nesug07/sa/sa13.pdf> (accessed February 2020).
23. Willett WC, Howe GR & Kushi LH (1997) Adjustment for total energy intake in epidemiologic studies. *Am J Clin Nutr* **65**, 1220S–1228S.
24. Asghari G, Rezazadeh A, Hosseini-Esfahani F, *et al.* (2012) Reliability, comparative validity and stability of dietary patterns derived from an FFQ in the Tehran Lipid and Glucose Study. *Br J Nutr* **108**, 1109–1117.
25. Hong X, Ye Q, Wang Z, *et al.* (2016) Reproducibility and validity of dietary patterns identified using factor analysis among Chinese populations. *Br J Nutr* **116**, 842–852.
26. Hu FB, Rimm E, Smith-Warner SA, *et al.* (1999) Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. *Am J Clin Nutr* **69**, 243–249.
27. Khani BR, Ye W, Terry P, *et al.* (2004) Reproducibility and validity of major dietary patterns among Swedish women assessed with a food-frequency questionnaire. *J Nutr* **134**, 1541–1545.
28. Nanri A, Shimazu T, Ishihara J, *et al.* (2012) Reproducibility and validity of dietary patterns assessed by a food frequency questionnaire used in the 5-year follow-up survey of the Japan Public Health Center-Based Prospective Study. *J Epidemiol* **22**, 205–215.
29. Ambrosini GL, O'Sullivan TA, de Klerk NH, *et al.* (2011) Relative validity of adolescent dietary patterns: a comparison of a FFQ and 3 d food record. *Br J Nutr* **105**, 625–633.
30. Okubo H, Murakami K, Sasaki S, *et al.* (2010) Relative validity of dietary patterns derived from a self-administered diet history questionnaire using factor analysis among Japanese adults. *Public Health Nutr* **13**, 1080–1089.
31. Papadaki A, Johnson L, Toumpakari Z, *et al.* (2018) Validation of the English version of the 14-item Mediterranean diet adherence screener of the PREDIMED study, in people at high cardiovascular risk in the UK. *Nutrients* **10**, 138.
32. Hoffmann K, Boeing H, Dufour A, *et al.* (2002) Estimating the distribution of usual dietary intake by short-term measurements. *Eur J Clin Nutr* **56**, Suppl. 2, S53–S62.
33. Tooze JA, Midthune D, Dodd KW, *et al.* (2006) A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. *J Am Diet Assoc* **106**, 1575–1587.
34. Kipnis V, Midthune D, Buckman DW, *et al.* (2009) Modeling data with excess zeros and measurement error: application to evaluating relationships between episodically consumed foods and health outcomes. *Biometrics* **65**, 1003–1010.