Contribution to the understanding of how principal component analysis–derived dietary patterns emerge from habitual data on food consumption

Carolina Schwedhelm,[1,2] Khalid Iqbal,[1,2] Sven Knüppel,[1] Lukas Schwingshackl,[1,2] and Heiner Boeing[1,2]

[1]Department of Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke (DIfE), Nuthetal, Germany and [2]NutriAct—Competence Cluster Nutrition Research Berlin-Potsdam, Germany

ABSTRACT

Background: Principal component analysis (PCA) is a widely used exploratory method in epidemiology to derive dietary patterns from habitual diet. Such dietary patterns seem to originate from intakes on multiple days and eating occasions. Therefore, analyzing food intake of study populations with different levels of food consumption can provide additional insights as to how habitual dietary patterns are formed.

Objective: We analyzed the food intake data of German adults in terms of the relations among food groups from three 24-h dietary recalls (24hDRs) on the habitual, single-day, and main-meal levels, and investigated the contribution of each level to the formation of PCA-derived habitual dietary patterns.

Design: Three 24hDRs were collected in 2010–2012 from 816 adults for an European Prospective Investigation into Cancer and Nutrition (EPIC)–Potsdam subcohort study. We identified PCA-derived habitual dietary patterns and compared cross-sectional food consumption data in terms of correlation (Spearman), consistency (intraclass correlation coefficient), and frequency of consumption across all days and main meals. Contribution to the formation of the dietary patterns was obtained through Spearman correlation of the dietary pattern scores.

Results: Among the meals, breakfast appeared to be the most consistent eating occasion within individuals. Dinner showed the strongest correlations with "Prudent" (Spearman correlation = 0.60), "Western" (Spearman correlation = 0.59), and "Traditional" (Spearman correlation = 0.60) dietary patterns identified on the habitual level, and lunch showed the strongest correlations with the "Cereals and legumes" (Spearman correlation = 0.60) habitual dietary pattern.

Conclusions: Higher meal consistency was related to lower contributions to the formation of PCA-derived habitual dietary patterns. Absolute amounts of food consumption did not strongly conform to the habitual dietary patterns by meals, suggesting that these patterns are formed by complex combinations of variable food consumption across meals. Dinner showed the highest contribution to the formation of habitual dietary patterns. This study provided information about how PCA-derived dietary patterns are formed and how they could be influenced. *Am J Clin Nutr* 2018;107:227–235.

Keywords: EPIC-Potsdam study, principal component analysis, dietary pattern analysis, consumption pattern, meal pattern, consistency of consumption

INTRODUCTION

The consumption of foods is highly intercorrelated and it is therefore difficult to study and interpret the effects of single foods on health outcomes. In order to get new insights, nutritional epidemiologists have started increasingly to focus on dietary patterns rather than on single foods or nutrients (1).

Two approaches, i.e., a priori (hypothesis-driven) and a posteriori (data-driven), are primarily used to derive dietary patterns. The a posteriori approach, also called exploratory analysis, is a preferred way to identify population-specific dietary patterns. Principal component analysis (PCA) is the most widely used method for such an analysis. This method is a variable-reducing procedure based on correlation or covariance matrices of the original variables, creating linear combinations (components, factors, or patterns) (2).

Foods are consumed on single eating occasions as meals or snacks. Meals are regular, specific eating occasions in which large amounts of foods are consumed, whereas snacks comprise smaller intakes between meals (3). Cumulatively, eating occasions (meals and snacks) form the overall diet, also called the

Supported by the German Federal Ministry of Education and Research (BMBF FKZ: 01ER0808 and 01EA1408A).

Supplemental Tables 1 and 2 and Supplemental Figures 1–6 are available from the "Supplemental data" link in the online posting of the article and from the same link in the online table of contents at https://academic.oup.com/ajcn/.

Address correspondence to CS (e-mail: carolina.schwedhelm@dife.de).

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition; FFQ, food-frequency questionnaire; ICC, intraclass correlation coefficient; PCA, principal component analysis; 24hDR, 24-hour dietary recall.

Received February 10, 2017. Accepted for publication November 14, 2017.

First published online February 26, 2018; doi: https://doi.org/10.1093/ajcn/nqx027.

TABLE 1
Selected baseline sociodemographic, lifestyle, and dietary characteristics of the studied population sample[1]

Characteristics	Men	Women	Total
n (%)	411 (50.5)	403 (49.5)	814 (100)
Age, y	66.4 ± 8.0	64.5 ± 8.7	65.5 ± 8.4
BMI, kg/m^2	27.7 ± 3.9	27.4 ± 4.8	27.5 ± 4.4
Physical activity level (TEE:REE ratio),[2] *n* (%)			
Extremely inactive (<1.4)	72 (20.6)	64 (19.1)	136 (19.9)
Sedentary (1.4 to <1.7)	168 (48.1)	195 (58.0)	363 (53.0)
Moderately active (1.7 to <2.0)	98 (28.1)	61 (18.1)	159 (23.2)
Vigorously active (2.0 to <2.4)	10 (2.9)	15 (4.5)	25 (3.6)
Extremely active (≥2.4)	1 (0.3)	1 (0.3)	2 (0.3)
Education, *n* (%)			
No vocational training/current vocational training	124 (30.2)	143 (35.5)	267 (32.8)
Technical college	63 (15.3)	124 (30.8)	187 (23.0)
University	224 (54.5)	136 (33.7)	360 (44.2)
Smoking status, *n* (%)			
Never smoker	132 (32.1)	245 (60.8)	377 (46.3)
Former smoker	235 (57.2)	118 (29.3)	353 (43.4)
Smoker	44 (10.7)	40 (9.9)	84 (10.3)
Participants consuming ≥1 meal,[3] *n* (%)			
Breakfast	411 (100)	403 (100)	814 (100)
Lunch	408 (99.3)	400 (99.3)	808 (99.3)
Afternoon snack	406 (98.8)	398 (98.8)	804 (98.8)
Dinner	411 (100)	403 (100)	814 (100)
Participants consuming meals on all days,[4] *n* (%)			
Breakfast	403 (98.0)	393 (97.5)	796 (97.8)
Lunch	323 (78.6)	327 (81.1)	650 (79.9)
Afternoon snack	278 (67.6)	285 (70.7)	563 (69.2)
Dinner	379 (92.2)	356 (88.3)	735 (90.3)
Energy intake, kcal/d	2341 ± 600	1770 ± 422	2058 ± 595
Energy intake, kcal/meal			
Breakfast	521 ± 214	380 ± 153	451 ± 199
Lunch	585 ± 249	471 ± 177	528 ± 224
Afternoon snack	292 ± 208	232 ± 167	263 ± 191
Dinner	609 ± 230	438 ± 175	524 ± 222

[1] Values are means ± SDs unless otherwise indicated. REE, resting energy expenditure; TEE, total energy expenditure.
[2] $n = 685$.
[3] Number of participants consuming the meal type ≥1 time.
[4] Number of participants consuming the meal type on all (available) recalled days.

habitual diet, and therefore influence the formation of a posteriori dietary patterns (1, 4). PCA is nonetheless frequently applied to data describing the habitual diets of individuals. By ignoring information on eating occasions, e.g., using data from food-frequency questionnaires (FFQs) as is often the case, characteristics of dietary patterns related to food intake on the level of eating occasions are lost. Such information is important as the way specific foods are consumed, e.g., simultaneously or separately during the different meals, might have an impact on health outcomes (5). Furthermore, it is not clear which combinations of foods are the driving forces forming a PCA-derived dietary pattern and why correlations among food groups exist on the habitual level.

By analyzing food consumption data from multiple 24-h dietary recalls (24hDRs), in which food intake can be studied by eating occasion, on the single-day level, and on the habitual level, we can investigate how the intake of foods is related in these different levels and how PCA-derived habitual dietary patterns relate to this structure of food intakes.

Therefore, the objectives of this study were as follows: *1*) to investigate relations among the different food groups on the habitual, single-day, and main-meal levels in terms of correlations, consistency, and frequency of consumption; *2*) to identify PCA-derived habitual dietary patterns; and *3*) to investigate the contribution of single days and different meals to the formation of these patterns. From this, we expect to obtain additional insights as to how habitual dietary patterns are formed, and how they could be influenced.

METHODS

Study sample

For this study, dietary data were obtained from a validation and calibration substudy within the European Prospective Investigation into Cancer and Nutrition (EPIC)–Potsdam cohort (initial recruitment date for EPIC-Potsdam: August 1994). The validation substudy was conducted between 2010 and 2012; recruitment started in August 2010. Details of the validation substudy are available elsewhere (6). In brief, participants aged 35–64 y at baseline from the EPIC-Potsdam cohort who were still actively participating in follow-up interviews were eligible for recruitment to the substudy. A total of 1447 invitations were sent

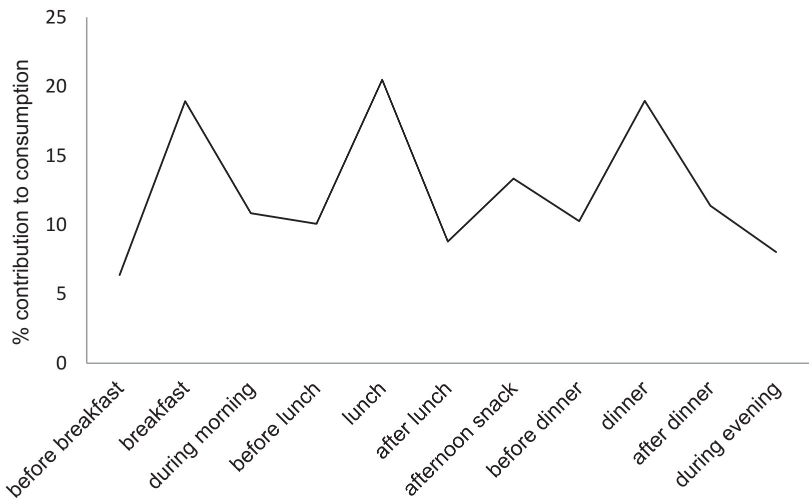

FIGURE 1 Mean contribution (% amount in grams) of eating occasions to food consumption over the day ($n = 814$).

out, drawn from a random age- and sex-stratified sample of the eligible study participants. Of these, 816 men and women finally provided data for this substudy. Ethical approval of the study was attained from the Ethics Committee of the Medical Association of the State of Brandenburg. All participants gave written informed consent.

Dietary assessment

Every participant provided three 24hDRs (mean ± SD = 2.99 ± 0.14) using EPIC-Soft (7) within a period of 4–24 mo (mean ± SD = 7 ± 1.27 mo). The first recall took place in person during the participants' first visit to the study center. The other two 24hDRs were obtained on random days including weekends over the telephone. All 24hDRs were performed by trained interviewers. Food consumption was recorded in 11 eating occasions throughout every day (see **Supplemental Table 1**) and was recorded for each food in grams per eating occasion. A total of 2431 such 24hDRs were collected.

Foods were collapsed into 39 food groups, shown in the text following, as has been done previously (see **Supplemental Table 2**) (8, 9) and averaged over the days and meals to derive the habitual and meal intakes, respectively. For studying the food intake on the meal level, we selected 4 main eating occasions (breakfast, lunch, afternoon snack, dinner), based on the participant-identified labels and peaks in contribution to the whole day's food intake (in grams). All eating occasions were included for the analysis of single days and habitual food intake.

Statistical methods

We performed correlation analyses using Spearman correlation coefficients to identify correlation arrangements on all levels of consumption. To further explore the observed relations, we calculated the intraclass correlation coefficient (ICC) for each food group as a measure of consistency of consumption over days and meals. Three participants with only one 24hDR were excluded from this analysis. The range for the ICC is from 0 (no agreement in food consumption over the days/meals) to 1 (perfect agreement in food consumption over the days/meals). We also calculated frequency of consumption in percentage of times that the foods were consumed (out of all respective eating occasions).

Dietary intake variables had a nonnormal distribution with extreme values (outliers); therefore, we used the Spearman correlation to calculate a posteriori PCA-derived dietary patterns on the habitual level. Especially in the presence of outliers, it is important to respect the assumption of normality of traditional PCA based on Pearson correlation coefficients (10, 11). Furthermore, basing our PCA on Spearman correlations ensured comparability with the correlation analyses described above. We calculated factor loadings after a varimax rotation to obtain uncorrelated components that are more easily interpretable (12). To obtain clearer patterns and avoid noise, we applied an often-used threshold ($\geq|0.3|$) (13, 14) for considering factor loadings as important contributors to a pattern and for labeling these patterns according to the food groups with the highest loadings. Four factors were retained based on scree plot analysis.

We calculated individual pattern scores for each of the retained components, taking into account all food groups (including those with factor loadings $<|0.3|$). Similarly, we obtained individual pattern scores for the different days and meals using the respective intakes and the obtained PCA-standardized scoring coefficients. We investigated the contribution of single days and different meals to the PCA-derived habitual dietary patterns by correlating pattern scores for days and meals with pattern scores on the habitual level. All statistical analyses were performed in SAS (version 9.4, Enterprise Guide 6.1, SAS Institute Inc.).

RESULTS

Two participants were excluded: one who was too young (<35 y old at EPIC-Potsdam baseline) and one with dementia. The final study sample consisted of 814 men and women from the validation and calibration substudy within the EPIC-Potsdam cohort (see **Supplemental Figure 1**). Participants were between 47 and 81 y of age (mean ± SD = 65.5 ± 8.4 y of age) at the time of their first visit. All participants (n = 814) consumed ≥1 breakfast and 1 dinner. Six participants did not consume lunch and 10 participants did not consume an afternoon snack on any of the recalled days (**Table 1**). **Figure 1** shows the percentage of contribution of each eating

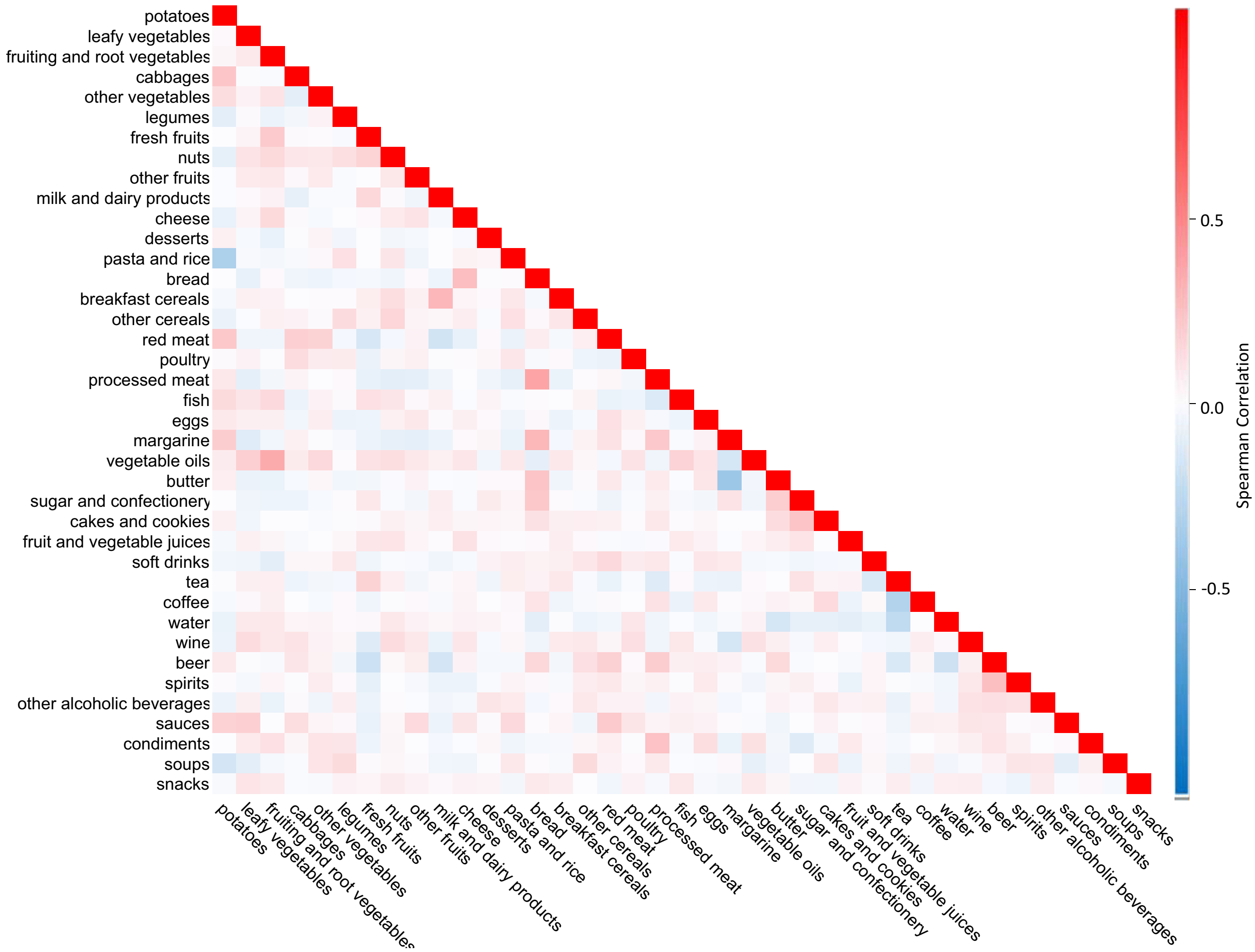

FIGURE 2 Heat map showing the Spearman correlation matrix for habitual food intake in grams ($n = 814$) by food group. The color corresponds to the strength of correlations (red: positive correlation; white: no correlation; blue: negative correlation).

occasion to the total amount of food consumption over the day (grams per eating occasion). We observed clear peaks at breakfast, lunch, afternoon, and dinner. The highest contribution to the daily amount in grams was seen for lunch (20.5%), followed by dinner (19.0%), breakfast (18.9%), and afternoon snack (13.3%).

Correlations among food groups on the habitual and meal food consumption levels

The Spearman correlation matrix of the habitual food intake (in grams) of the 814 participants is shown as a heat map on Figure 2. Some food groups that are commonly eaten together showed strong positive correlations, such as "bread" with "margarine," "butter," and "cheese." On the other hand, food groups that are commonly substitutes for each other were strongly negatively correlated, such as "potatoes" with "pasta and rice," as well as "tea" with "coffee."

We further correlated food intakes (in grams) within strata of meals (**Supplemental Figures 2–5**). In general, we saw stronger correlations than on the habitual level. Across all meals, we observed strong positive correlations with "bread" and "cheese." Some meal-specific strong positive correlations were observed between "breakfast cereals" and "milk and dairy products" for breakfast (Supplemental Figure 2), and between "cakes and cookies" and "coffee" for the afternoon snack (Supplemental Figure 4). Some meal-specific strong negative correlations were observed between "margarine" and "butter," and "tea" and "coffee" for breakfast. Out of the 4 meals, dinner showed the weakest correlations among food groups (Supplemental Figure 5), though these were similar to the correlations for habitual food intake.

Consistency and frequency of consumption

Next, we analyzed whether correlation results related to consistency or stability of intake across days or meals, and to frequency of intake. Table 2 shows the ICCs in food consumption across single days and across meals. Consistency of consumption was highest across the days and across breakfast. The highest consistency of consumption across days was seen for "margarine" (ICC = 0.62), "coffee" (ICC = 0.58), "tea" (ICC = 0.55), "water" (ICC = 0.54), and "butter" (ICC = 0.53). Across breakfast, the highest consistency was observed for "tea" (ICC = 0.69), "milk and dairy products" (ICC = 0.63), "coffee" (ICC = 0.61), "margarine" (ICC = 0.60), "butter" (ICC = 0.59), "breakfast cereals" (ICC = 0.54), "sugar and confectionery" (ICC = 0.53), and

TABLE 2
Intraindividual consistency (as ICC) of consumption across days and meals (n = 811)[1]

Food group	Day	Breakfast	Lunch	Afternoon snack	Dinner
Potatoes	0.17	0.00	0.22	0.00	0.10
Leafy vegetables	0.03	0.00	0.04	0.00	0.08
Fruiting and root vegetables	0.12	0.36	0.05	0.08	0.12
Cabbages	0.04	0.00	0.03	0.00	0.05
Other vegetables	0.00	0.03	0.03	0.01	0.02
Legumes	0.29	0.47	0.02	0.33	0.00
Fresh fruits	0.33	**0.50**[2]	0.10	0.04	0.20
Nuts	0.19	0.36	0.02	0.00	0.07
Other fruits	0.01	0.00	0.04	0.00	0.00
Milk and dairy products	0.45	**0.63**	0.14	0.08	0.30
Cheese	0.22	0.41	0.06	0.00	0.14
		0.40			
Desserts	0.09	0.00	0.09	0.00	0.14
Pasta, rice	0.04	0.00	0.04	0.00	0.01
Bread	0.45	0.44	0.19	0.09	0.35
Breakfast cereals	0.45	**0.54**	0.19	0.00	0.34
Other cereals	0.09	0.46	0.00	0.00	0.01
Red meat	0.11	0.10	0.07	0.02	0.05
Poultry	0.00	0.00	0.00	0.00	0.01
Processed meat	0.22	0.44	0.07	0.08	0.17
Fish	0.08	0.31	0.01	0.00	0.09
Eggs	0.11	0.17	0.05	0.00	0.01
Margarine	**0.62**	**0.60**	0.16	0.12	0.45
Vegetable oils	0.11	0.20	0.10	0.01	0.13
Butter	**0.53**	**0.59**	0.13	0.03	0.30
Sugar and confectionery	0.38	**0.53**	0.02	0.07	0.07
Cakes and cookies	0.18	0.16	0.00	0.18	0.02
Fruit and vegetable juices	0.38	0.33	0.20	0.10	0.21
Soft drinks	0.35	0.01	0.05	0.07	0.22
Tea	**0.55**	**0.69**	0.23	0.29	0.41
Coffee	**0.58**	**0.61**	0.20	0.36	0.13
Water	**0.54**	0.24	0.23	0.13	0.21
Wine	0.32	0.00	0.21	0.00	0.20
Beer	0.48	—	0.19	0.11	0.28
Spirits	0.12	—	0.00	0.04	0.00
Other alcoholic beverages	0.00	—	0.00	0.00	0.00
Sauces	0.02	0.22	0.05	0.00	0.00
Condiments	0.18	0.37	0.01	0.03	0.06
Soups	0.09	0.41	0.05	0.00	0.08
Snacks	0.04	0.15	0.00	0.00	0.00

[1] n = 811 participants with at least two 24hDRs; across all available observations. ICC, intraclass correlation coefficient.
[2] ICC ≥0.50 are shown in bold.

"fresh fruits" (ICC = 0.50). The other 3 meals showed very low consistency of consumption.

We were able relate some of the strongest correlations between food intakes on the habitual level (Figure 2) to consistency across days (Table 2). For instance, "breakfast cereals" and "milk and dairy products" correlated positively, and both food groups showed the same and relatively high consistency (ICC = 0.45). "Margarine" and "butter," as well as "coffee" and "tea," showed a strong negative correlation and high consistency as well (ICC = 0.62, 0.53, 0.58, and 0.55, respectively). Similar links could be found on the meal level in the case of breakfast; the strong positive correlation between the intake (in grams) of "breakfast cereals" and "milk and dairy products" is reflected by the high consistency (ICC = 0.54 and 0.63, respectively), and the strong negative correlations between "margarine" and "butter" and between "tea" and "coffee" is related to high consistency (ICC = 0.60, 0.59, 0.69, and 0.61, respectively). However, other strong correlations involving "fruiting and root vegetables" and "other vegetables" did not relate to high consistency of consumption across breakfasts.

Table 3 shows the food consumption frequencies as the percentage of times that the foods were consumed. On single days, frequencies of consumption were higher due to the cumulative effect of all eating occasions in a day; the most frequently consumed foods in a whole day were "bread" (98.1%), followed by "water" (92.5%) and "coffee" (92.1%). The most frequently consumed foods at breakfast were "bread" (88.6%) and "coffee" (72.8%). Lunch showed in general a low frequency of consumption, suggesting a high variability. During afternoon snacks, "coffee" and "cakes and cookies" were the most frequently consumed foods (63.3% and 51.6%, respectively). During dinners, "bread" (72.0%), "processed meat" (53.7%), and "fruiting and root vegetables" (52.6%) were consumed the most frequently.

Not all food groups with strong correlations (see Figure 2) were consumed frequently (see Table 3). Across days, only "bread" and "processed meat" correlated strongly and were also frequently consumed (98.1% and 78.5%, respectively). For breakfast, frequencies of consumption did not further relate to the correlations. At lunch, "potatoes" (in grams) were the most frequently consumed food (49.1%) and showed strong correlations with other food groups on Supplemental Figure 3. For afternoon snacks, the strong positive correlation between "coffee" and "cakes and cookies" (Supplemental Figure 4) related to a high frequency of consumption (63.3% and 51.6%, respectively). Finally, we could relate dinner food intake correlations (Supplemental Figure 5) to food consumption frequency values for "bread" with "cheese" and with "processed meat" (consumed at 72.0%, 47.4%, and 53.7% of dinners, respectively).

PCA-derived habitual dietary intake patterns

Four habitual dietary patterns explaining 20.92% of the variance in food intake were retained based on the scree plot (**Supplemental Figure 6**). **Table 4** shows the factor loadings for the PCA-habitual dietary patterns as well as the average habitual food intakes in grams per day for orientation.

The pattern labelled "Prudent" was characterized by high intake of "leafy vegetables," "fruiting and root vegetables," "fresh fruits," "nuts," "fish," "vegetable oils," and "wine," and by low intake of "margarine" and explained 6.13% of the total variance. The pattern labelled "Western" was characterized by high intake of "potatoes," "cabbages," "red meat," "beer," "sauces," and "condiments," and by low intake of "fresh fruits," "milk and dairy products," and "tea"; it explained 5.49% of the total variance. The pattern labelled "Traditional" was characterized by a high intake of "bread," "processed meat," "butter," "sugar and confectionery," and "cakes and cookies" and a low intake of "water." This dietary pattern explained an additional 4.74% of the total variance. Finally, the habitual dietary pattern labelled "Cereals and legumes" was characterized by a high consumption of "legumes," "pasta and rice," "other cereals," "other alcoholic beverages," and "soups" and by a low intake of "potatoes." This last pattern explained an additional 4.56% of the total variance.

TABLE 3

Frequency of consumption of 39 food groups across days and meals: percentage of days and meals in which foods were consumed ($n = 814$)[1]

Food group	Days ($n = 2431$)	Breakfasts ($n = 2411$)	Lunches ($n = 2236$)	Afternoon snacks ($n = 2119$)	Dinners ($n = 2346$)
Potatoes	54.3	0.0	49.1	1.2	9.9
Leafy vegetables	17.7	0.7	8.6	0.3	9.3
Fruiting and root vegetables	72.5	11.3	35.4	2.4	52.6
Cabbages	20.6	0.0	17.0	0.4	5.0
Other vegetables	60.8	2.6	45.3	1.5	27.7
Legumes	5.5	1.0	3.3	0.4	1.5
Fresh fruits	81.7	28.6	35.2	12.4	25.2
Nuts	13.4	4.4	1.7	0.9	1.7
Other fruits	7.6	0.7	3.1	0.4	2.9
Milk and dairy products	86.8	68.5	29.7	47.3	19.5
Cheese	73.7	40.8	10.2	2.0	47.4
Desserts	14.2	0.0	7.3	3.5	1.5
Pasta, rice	16.9	0.8	13.2	0.6	4.1
Bread	98.1	88.6	23.2	8.6	72.0
Breakfast cereals	7.5	6.1	0.6	0.2	0.4
Other cereals	21.2	4.4	11.4	0.9	3.9
Red meat	34.8	1.5	26.5	1.2	10.5
Poultry	13.4	0.5	8.0	0.3	5.6
Processed meat	78.5	33.8	30.6	3.6	53.7
Fish	22.4	4.4	8.4	0.6	12.2
Eggs	30.5	16.5	9.7	0.7	6.1
Margarine	55.5	32.0	24.1	2.7	32.7
Vegetable oils	41.1	2.5	27.0	0.9	18.7
Butter	69.2	46.6	28.2	3.9	36.5
Sugar and confectionery	85.3	66.9	17.3	20.1	17.1
Cakes and cookies	56.4	2.0	3.6	51.6	1.7
Fruit and vegetable juices	40.8	12.3	13.2	5.2	12.2
Soft drinks	15.0	0.3	6.4	1.8	5.1
Tea	57.6	24.6	8.9	11.9	29.6
Coffee	92.1	72.8	9.7	63.3	1.8
Water	92.5	16.6	45.9	26.0	36.6
Wine	21.9	0.3	3.3	2.1	5.8
Beer	26.0	0.0	4.0	1.6	12.7
Spirits	3.7	0.0	0.1	0.4	0.4
Other alcoholic beverages	5.8	0.0	0.7	0.9	0.9
Sauces	46.3	3.9	29.4	1.7	20.6
Condiments	41.1	11.5	17.7	4.8	19.1
Soups	27.9	1.0	22.0	1.0	7.3
Snacks	2.2	0.6	0.5	0.2	1.1

[1]In percentages, over a period of 3 observations. If <3 recalls were available, the total of the available observations counted as 100%; days and meals were treated as independent observations. Descriptive results.

The Spearman correlation coefficients for adherence to the 4 habitual dietary patterns (pattern scores) are shown in the upper part of **Table 5**. The "Prudent" habitual dietary pattern scores correlated inversely with the "Western" and with the "Cereals and legumes" habitual dietary pattern scores (Spearman correlation = –0.65 and –0.61, respectively). The "Western" habitual dietary pattern scores correlated positively with those of the "Cereals and legumes" habitual dietary pattern (Spearman correlation = 0.61) (Table 5). The "Traditional" habitual dietary pattern scores were slightly inversely correlated with those of the "Western" and the "Cereals and legumes" patterns (Spearman correlation = –0.20 and –0.16, respectively) and slightly positively correlated with the "Prudent" pattern score (Spearman correlation = 0.15).

Pattern scores for the PCA-derived patterns on the meal level showed different correlations for each meal type with the habitual level, suggesting that every meal type contributed to a different extent to the formation of the PCA-derived habitual dietary patterns. Pattern scores for the "Prudent" habitual dietary pattern correlated strongest for dinner, followed by lunch and breakfast, and afternoon snacks last (Spearman correlation = 0.60, 0.53, 0.53, and 0.34, respectively). Pattern scores for the "Western" habitual dietary pattern also correlated strongest for dinner, followed by breakfast, lunch, and afternoon snacks (Spearman correlation = 0.59, 0.51, 0.42, and 0.39, respectively). Correlations for the "Traditional" dietary pattern scores were also strongest for dinner, followed by lunch, then afternoon snacks, and breakfast last (Spearman correlation = 0.60, 0.58, 0.44, and 0.33, respectively). Finally, pattern scores for the "Cereals and legumes" habitual dietary pattern correlated strongest for lunch, followed by dinner, breakfast, and afternoon snacks (Spearman

TABLE 4
Average habitual food intake and factor loadings for the 4 PCA-derived habitual dietary patterns for all ($n = 814$) participants[1]

Food groups	Average habitual intake (g/d)	Factor loadings for dietary patterns			
		Prudent	Western	Traditional	Cereals and legumes
Potatoes	81.7	0.07	**0.35**[2]	0.09	**–0.61**
Leafy vegetables	11.6	**0.41**	0.13	–0.19	–0.02
Fruiting and root vegetables	103	**0.55**	0.06	0.01	–0.13
Cabbages	22.5	0.06	**0.35**	–0.09	–0.11
Other vegetables	32.9	0.19	0.26	–0.05	0.01
Legumes	6.64	0.00	0.02	–0.05	**0.42**
Fresh fruits	231	**0.37**	**–0.37**	0.08	–0.12
Nuts	3.95	**0.45**	0.03	0.00	0.27
Other fruits	10.2	0.29	0.21	–0.01	0.04
Milk and dairy products	167	0.20	**–0.35**	0.03	–0.04
Cheese	37.4	0.28	0.05	0.22	0.04
Desserts	17.6	–0.05	0.02	0.02	0.02
Pasta and rice	23.1	0.07	–0.12	0.01	**0.54**
Bread	113	–0.08	0.11	**0.68**	–0.08
Breakfast cereals	3.40	0.28	–0.14	0.04	0.14
Other cereals	5.30	0.12	0.08	0.16	**0.38**
Red meat	39.5	–0.12	**0.54**	0.07	–0.10
Poultry	14.8	0.09	0.19	–0.15	0.15
Processed meat	60.8	–0.25	0.28	**0.39**	–0.09
Fish	24.1	**0.37**	0.00	0.03	–0.08
Eggs and egg products	18.7	0.11	0.29	0.07	–0.06
Margarine	13.2	**–0.31**	0.16	0.17	–0.29
Vegetable oils	5.06	**0.58**	0.13	–0.11	–0.03
Butter	17.6	0.06	0.04	**0.47**	0.10
Sugar and confectionery	38.0	0.00	–0.19	**0.54**	0.01
Cakes and cookies	59.2	0.08	0.02	**0.43**	0.11
Fruit and vegetable juices	94.5	0.25	–0.08	0.22	0.00
Soft drinks	48.1	–0.17	0.21	0.01	0.20
Tea	355	0.23	**–0.38**	0.22	–0.12
Coffee	447	–0.06	0.22	0.09	0.12
Water	740	0.08	0.04	**–0.44**	0.01
Wine	57.3	**0.30**	0.26	–0.06	0.28
Beer	173	–0.07	**0.51**	0.28	0.12
Spirits	1.59	–0.08	0.26	0.12	0.22
Other alcoholic beverages	4.99	–0.01	0.15	0.05	**0.37**
Sauces	24.2	0.16	**0.40**	–0.04	–0.05
Condiments	2.79	0.12	**0.31**	–0.04	0.06
Soups	51.8	–0.17	0.04	0.08	**0.42**
Snacks	1.60	0.25	–0.02	0.02	0.02
Total variance explained, %	20.92 (all factors)	6.13	5.49	4.74	4.56

[1]Habitual dietary patterns were PCA-derived using Spearman correlation matrix. PCA, principal component analysis.
[2]Factor loadings with an absolute value $\geq$0.30 are shown in bold.

correlation = 0.60, 0.53, 0.36, and 0.26, respectively). Correlations on the single-day level to the formation of the habitual dietary patterns showed little variation across days and patterns (results not shown).

DISCUSSION

In this study, we investigated the relation between PCA-derived dietary patterns in a German adult population, using habitual intakes with different levels of intake, i.e., daily and meal level. We observed a higher contribution of the meals to the formation of the PCA-derived habitual dietary patterns when consistency of consumption was low and the intake of foods was substantial in quantity for the respective meal. Consequently, correlation analyses of the obtained habitual dietary pattern scores across the investigated levels of food consumption suggest that dinner, followed by lunch, played an important role in the formation of these habitual dietary patterns.

The participants in our study sample showed very regular breakfast consumption habits: 97.8% ate breakfast on all available 24hDRs and 0% skipped breakfast on all days. Regular breakfast consumption has been associated with higher dietary quality (15, 16). In line with our findings, eating patterns have been found to be more consistent in the morning and less consistent in meals that are away from home, in social situations, and containing alcohol (17), which are more likely to happen in

TABLE 5
Spearman correlations of habitual dietary pattern scores on the habitual and meal levels ($n = 814$)[1]

Dietary pattern scores (habitual and meal levels)	Habitual dietary pattern scores			
	Prudent	Western	Traditional	Cereals and legumes
Habitual diet				
Prudent	1.00[2]			
Western	–0.65	1.00		
Traditional	0.15	–0.20	1.00	
Cereals and legumes	–0.61	0.61	–0.16	1.00
Breakfast	0.53	0.51	0.33	0.36
Lunch	0.53	0.42	0.58	0.60
Afternoon snack	0.34	0.39	0.44	0.26
Dinner	0.60	0.59	0.60	0.53

[1]Habitual level refers to the average daily food consumption; meal level refers to the meal-specific average food consumption. Habitual dietary patterns were PCA-derived using the Spearman correlation matrix. PCA, principal component analysis.

[2]All values on the table had probability <0.0001.

the evening. Furthermore, in a study analyzing the same population, breakfast was found to be associated with cardiometabolic risk factors (18). This observed association is potentially reinforced by the consistency of this meal. Therefore, the breakfast meal might offer an opportunity for an effective long-term adherence to dietary recommendations, potentially resulting in a consistent impact in the overall habitual food consumption and health status of individuals.

The strong correlations observed in meals between different food groups usually persisted in the habitual level; however, correlations between food groups in the habitual level were not representative of those seen in meals. Additionally, we could not explain some strong correlations concerning foods that are not frequently consumed. This suggests that foods might be related in a different way, e.g., concomitant consumption. Therefore, further analyses involving food networks could give us some information about a possible dependency of consumption among food groups and the extent to which this can influence the observed correlations. A recent study explored dietary networks in the German EPIC-Potsdam population to describe consumption patterns, i.e., how foods are consumed in relation to each other (19). The major dietary networks consisted of red and processed meat, cooked vegetables, sauces, and potatoes, among other food groups. Since this population is very similar to the one used in our study, we assume the networks would show similarity with our results. Although the dietary networks constructed by Iqbal et al. (19) were based on dietary data from FFQs (and therefore nonmeal specific), these networks share some characteristics with our results, such as positive correlations between red meat and sauces, fruits and vegetables, and a negative relation between margarine and butter. Comparing these dietary networks to our PCA-derived habitual dietary patterns, a few resemblances were noticeable with all patterns.

The correlations that we observed for the dietary pattern scores on the meal level suggest that dinner food intake had a higher impact on the habitual dietary pattern formation for the patterns “Prudent,” “Western,” and “Traditional,” but that lunch had a greater impact on the formation of the pattern “Cereals and legumes.” The PCA method seeks to group food consumption in a way that explains the most possible variance. Due to the nature of this method, the observed meal-specific contribution to the PCA-derived habitual dietary patterns tended to be higher for large meals with low consistency of consumption and high interindividual variation. We chose PCA as the basis of our analysis of dietary patterns because it is the most frequently used exploratory method (20–22). Similar analyses can be carried out analyzing other exploratory methods to derive dietary patterns such as cluster analysis, reduced rank regression, and dietary intake networks.

A limitation specific to our study is the number of 24hDRs available per person. Studies assessing the validity of three 24hDRs to estimate energy intake have reached mixed conclusions (23–25). Especially within a homogeneous population, intraindividual variation of dietary factors can be even larger than the interindividual variation (26–28). In our results, day-to-day variation in diet within the same individual becomes evident, underlining the importance of using repeated measurements for increased reliability. Due to the methodological limitation that FFQs do not provide meal-specific information (3), 1-wk food records or a higher number of repeated 24hDRs are examples of tools that could improve the accuracy of the estimated intraindividual variation (25, 29, 30). Finally, a methodological challenge in our study is the nature of the data and the subsequent high proportion of zeros; dealing with meal-specific food consumption of a wide number of food groups results in many zero values, since only some food groups are consumed per person per meal. As a consequence, our data are highly positively skewed. All analyses performed in this study were appropriate for nonparametric data with the exception of the calculation of the ICC (31, 32). However, to avoid further compromising interpretability, we chose not to transform our data. As an exploratory analysis, with the purpose of comparing values across different observations but not involving corrections or testing, this method provides a better estimate of within-person variability relative to the between-person variability of food consumption.

In conclusion, we could show that habitual patterns to some extent originate at the meal level, which could lead to a better understanding of how PCA-derived dietary patterns arise. Meal pattern analyses with the description of correlations and variance components in general might help to generate more effective dietary advice in the longer term.

We thank the Human Study Centre (HSC) of the German Institute of Human Nutrition Potsdam-Rehbrücke, namely the trustee and the examination unit for the collection, the data hub for the processing, the participants for the provision of the data, and the head of the HSC, Manuela Bergmann, for the contribution to the study design and leading the underlying process of data generation. We thank Ellen Kohlsdorf for data handling and technical assistance.

The authors’ responsibilities were as follows—CS: analyzed the data; CS and KI: wrote the paper; KI and SK: contributed with essential methodological advice; CS: had primary responsibility for the final content; and all authors: contributed to the conception and design of the research, and read and approved the final manuscript. None of the authors reported a conflict of interest related to the study.

REFERENCES

1. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. Curr Opin Lipidol 2002;13:3–9.

2. Gorst-Rasmussen A, Dahm CC, Dethlefsen C, Scheike T, Overvad K. Exploring dietary patterns by using the treelet transform. Am J Epidemiol 2011;173:1097–104.
3. Leech RM, Worsley A, Timperio A, McNaughton SA. Understanding meal patterns: definitions, methodology and impact on nutrient intake and diet quality. Nutr Res Rev 2015;28:1–21.
4. Myhre JB, Loken EB, Wandel M, Andersen LF. Meal types as sources for intakes of fruits, vegetables, fish and whole grains among Norwegian adults. Public Health Nutr 2015;18:2011–21.
5. Woolhead C, Gibney MJ, Walsh MC, Brennan L, Gibney ER. A generic coding approach for the examination of meal patterns. Am J Clin Nutr 2015;102:316–23.
6. Neamat-Allah J, Wald D, Hüsing A, Teucher B, Wendt A, Delorme S, Dinkel J, Vigl M, Bergmann MM, Feller S, et al. Validation of anthropometric indices of adiposity against whole-body magnetic resonance imaging—a study within the German European Prospective Investigation into Cancer and Nutrition (EPIC) Cohorts. PLoS One 2014;9:e91586.
7. Voss S, Charrondiere UR, Slimani N, Kroke A, Riboli E, Wahrendorf J, Boeing H. EPIC-SOFT ein europaisches Computerprogramm fur 24-Stunden-Erinnerungsprotokolle [German] (EPIC-SOFT a European computer program for 24-hour dietary protocols). Z Ernahrungswiss 1998;37:227–33.
8. Schulz M, Hoffmann K, Weikert C, Nothlings U, Schulze MB, Boeing H. Identification of a dietary pattern characterized by high-fat food choices associated with increased risk of breast cancer: the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam Study. Br J Nutr 2008;100:942–6.
9. Haubrock J, Nothlings U, Volatier JL, Dekkers A, Ocke M, Harttig U, Illner AK, Knuppel S, Andersen LF, Boeing H. Estimating usual food intake distributions by using the multiple source method in the EPIC-Potsdam Calibration Study. J Nutr 2011;141:91–20.
10. Jolliffe I. Principal component analysis. 2nd ed. Wiley Online Library; 2002.
11. O'Rourke N, Hatcher L. A step-by-step approach to using SAS for factor analysis and structural equation modeling, SAS Institute; 2013.
12. Norm O'Rourke, Larry Hatcher, Edward J. Stephanski. Principal component analysis. A step-by-step approach to using SAS for univariate & multivariate statistics. 2nd ed. New York: Wiley-Interscience; 2005.
13. Castello A, Lope V, Vioque J, Santamarina C, Pedraz-Pingarron C, Abad S, Ederra M, Salas-Trejo D, Vidal C, Sanchez-Contador C, et al. Reproducibility of data-driven dietary patterns in two groups of adult Spanish women from different studies. Br J Nutr 2016;116:734–42.
14. Markussen MS, Veierod MB, Ursin G, Andersen LF. The effect of under-reporting of energy intake on dietary patterns and on the associations between dietary patterns and self-reported chronic disease in women aged 50–69 years. Br J Nutr 2016;116:547–58.
15. Laska MN, Hearst MO, Lust K, Lytle LA, Story M. How we eat what we eat: identifying meal routines and practices most strongly associated with healthy and unhealthy dietary factors among young adults. Public Health Nutr 2015;18):2135–45.
16. Azadbakht L, Haghighatdoost F, Feizi A, Esmaillzadeh A. Breakfast eating pattern and its association with dietary quality indices and anthropometric measurements in young women in Isfahan. Nutrition 2013;29:420–5.
17. Vainik U, Dube L, Lu J, Fellows LK. Personality and situation predictors of consistent eating patterns. PLoS One 2015;10: e0144134.
18. Iqbal K, Schwingshackl L, Gottschald M, Knuppel S, Stelmach-Mardas M, Aleksandrova K, Boeing H. Breakfast quality and cardiometabolic risk profiles in an upper middle-aged german population. Eur J Clin Nutr 2017. doi: 10.1038/ejcn.2017.116.
19. Iqbal K, Buijsse B, Wirth J, Schulze MB, Floegel A, Boeing H. Gaussian graphical models identify networks of dietary intake in a German adult population. J Nutr 2016;146:646–52.
20. Alhazmi A, Stojanovski E, McEvoy M, Garg ML. The association between dietary patterns and type 2 diabetes: a systematic review and meta-analysis of cohort studies. J Hum Nutr Diet 2014;27:251–60.
21. Carvalho CA, Fonseca PC, Nobre LN, Priore SE, Franceschini Sdo C. [Methods of a posteriori identification of food patterns in Brazilian children: a systematic review]. Cien Saude Colet 2016;21:143–54.
22. Yusof AS, Isa ZM, Shah SA. Dietary patterns and risk of colorectal cancer: a systematic review of cohort studies (2000-2011). Asian Pac J Cancer Prev 2012;13:4713–7.
23. Ma Y, Olendzki BC, Pagoto SL, Hurley TG, Magner RP, Ockene IS, Schneider KL, Merriam PA, Hebert JR. Number of 24-hour diet recalls needed to estimate energy intake. Ann Epidemiol 2009;19:553–9.
24. Lins IL, Bueno NB, Grotti Clemente AP, Pfrimer K, Sawaya AL, de Menezes Toledo Florencio TM. Energy intake in socially vulnerable women living in Brazil: assessment of the accuracy of two methods of dietary intake recording using doubly labeled water. J Acad Nutr Diet 2016;116:1560–7.
25. Carroll RJ, Midthune D, Subar AF, Shumakovich M, Freedman LS, Thompson FE, Kipnis V. Taking advantage of the strengths of 2 different dietary assessment instruments to improve intake estimates for nutritional epidemiology. Am J Epidemiol 2012;175:340–7.
26. Beaton GH, Milner J, McGuire V, Feather TE, Little JA. Source of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. Carbohydrate sources, vitamins, and minerals. Am J Clin Nutr 1983;37:986–95.
27. Sempos CT, Johnson NE, Smith EL, Gilligan C. Effects of intraindividual and interindividual variation in repeated dietary records. Am J Epidemiol 1985;121:120–30.
28. Liu K, Stamler J, Dyer A, McKeever J, McKeever P. Statistical methods to assess and minimize the role of intra-individual variability in obscuring the relationship between dietary lipids and serum cholesterol. J Chron Dis 1978;31:399–418.
29. McAvay G, Rodin J. Interindividual and intraindividual variation in repeated measures of 24-hour dietary recall in the elderly. Appetite 1988;11:97–110.
30. Thompson FE, Subar AF. Dietary assessment methodology. In: AM Coulston,CJ BousheyMG Ferruzzi, editors. Nutrition in the prevention and treatment of disease. Oxford: Academic Press; 2013. p. 5–46.
31. Rosner B. Fundamentals of Biostatistics. 5th edn. Pacific Grove, CA: Brooks/Cole; 2000.
32. Looney SW, Hagan JL. Analysis of Biomarker Data: A Practical Guide. Hoboken, NJ: John Wiley; 2015.