Western and carnivorous dietary patterns are associated with greater likelihood of IBD-development in a large prospective population-based cohort

Peters, Vera; Bolte, Laura; Schuttert, Eva; Andreu-Sánchez, Sergio; Dijkstra, Gerard; Weersma, Rinse; Campmans-Kuijpers, Marjo

Published in:
Journal of Crohn's and Colitis

DOI:
10.1093/ecco-jcc/jjab219

Journal of Crohn's and Colitis, 2022, **16**, 931–939
https://doi.org/10.1093/ecco-jcc/jjab219
Advance access publication 1 December 2021
Original Article

Western and Carnivorous Dietary Patterns are Associated with Greater Likelihood of IBD Development in a Large Prospective Population-based Cohort

Vera Peters,[a,b] Laura Bolte,[a,c] Eva [Monique] Schuttert,[a,c] Sergio Andreu-Sánchez,[c] Gerard Dijkstra,[a] Rinse [Karel] Weersma,[a,c,‡] Marjo [Johanna Elisabeth] Campmans-Kuijpers[a,‡]

[a]Department of Gastroenterology and Hepatology, University of Groningen and University Medical Centre Groningen, The Netherlands
[b]Department of Epidemiology, University of Groningen and University Medical Centre Groningen, The Netherlands
[c]Department of Genetics, University of Groningen and University Medical Centre Groningen, The Netherlands

Corresponding author: Marjo [Johanna Elisabeth] Campmans-Kuijpers, University Medical Centre Groningen, Hanzeplein 1, PO-box 30.001, 9700 RB, Department of Gastroenterology, HPC BB41, Groningen, The Netherlands. Tel.: +31 [0]50 361 2620; fax: +31 [0]50 361 9306; email: m.j.e.campmans-kuijpers@umcg.nl
‡Shared last authorship.

Abstract

Objective: Nutrition plays a role in the development of Crohn's disease [CD] and ulcerative colitis [UC]. However, prospective data on nutrition and disease onset are limited. Here, we analysed dietary patterns and scores in relation to inflammatory bowel disease [IBD] development in a prospective population-based cohort.

Methods: We analysed 125 445 participants of whom 224 individuals developed *de novo* UC and 97 CD over a maximum 14-year follow-up period. Participants answered health-related [also prospectively] and dietary questionnaires [FFQ] at baseline. Principal component analysis [PCA] was conducted deriving *a-posteriori* dietary patterns. Hypotheses-based *a-priori* dietary scores were also calculated, including the protein score, Healthy Eating Index, LifeLines Diet Score [LLDS], and alternative Mediterranean Diet Score. Logistic regression models were performed between dietary patterns, scores, and IBD development.

Results: PCA identified five dietary patterns. A pattern characterised by high intake of snacks, prepared meals, non-alcoholic beverages, and sauces along with low vegetables and fruit consumption was associated with higher likelihood of CD development (odds ratio [OR]: 1.16, 95% confidence interval [CI]: 1.03-1.30, $p = 0.013$). A pattern comprising red meat, poultry, and processed meat, was associated with increased likelihood of UC development [OR: 1.11, 95% CI: 1.01-1.20, $p = 0.023$]. A high diet quality score [LLDS] was associated with decreased risk of CD [OR: 0.95, 95% CI: 0.92-0.99, $p = 0.009$].

Conclusions: A Western dietary pattern was associated with a greater likelihood of CD development and a carnivorous pattern with UC development, whereas a relatively high diet quality [LLDS] was protective for CD development. Our study strengthens the importance of evaluating dietary patterns to aid prevention of IBD in the general population.

Key Words: Inflammatory bowel disease [IBD]; dietary patterns; principal component analysis [PCA]; dietary scores; Protein Score; Healthy Eating Index [HEI]; LifeLines Diet Score [LLDS]; Alternate Mediterranean Diet Score [aMED]

1. Introduction

Crohn's disease [CD] and ulcerative colitis [UC], together referred to as inflammatory bowel disease [IBD], are chronic inflammatory disorders of the intestine. It is hypothesized that IBD is triggered and maintained by environmental factors, including diet, in genetically predisposed individuals with gut dysbiosis and an aberrant immune response.[1] The exact interplay between those pathophysiological factors is unknown.[2]

Nutrition, through its interactions with immunity, host barrier function, and the gut microbiota, plays a key role in the pathogenesis of IBD.[3] A Westernised lifestyle has been suggested to contribute to the rising incidence of IBD in developing countries.[4] This is supported by functional studies showing an increase in intestinal inflammation upon administration of saturated fat, cholesterol, or food additives,[5] as well as by retrospective cohort studies showing a correlation between the intake of animal protein and IBD onset.[6] In contrast, a Mediterranean diet, which is widely considered a healthy dietary pattern with anti-inflammatory effects, has been associated with a significantly lower risk of later onset CD.[7]

Whereas nutrients and single food items often are of interest in studies investigating specific diet-disease relationships, it should be recognised that these elements likely act synergistically or antagonistically as part of a large matrix i.e., habitual diet.[8] Therefore, it is believed that dietary patterns have great clinical implications[9] and should be studied in large longitudinal population-based cohorts to assess their role in disease development. Principal component analysis [PCA] is a data-driven dimensionality-reduction method used to identify such *a-posteriori* dietary patterns, explaining most of the habitual intake variety among individuals in a given popu-

Received: July 15, 2021. **Revised:** October 9, 2021. **Accepted:** November 30, 2021

lation. In recent years, this method has become of interest in the nutritional field. Another method to associate overall dietary intake with health outcomes is *a-priori* defined dietary scores, which are based on hypotheses of food items being harmful or beneficial and score the adherence to targeted dietary recommendations.[6,7,10] Here, we focus on four previously published dietary scores: the Protein Score,[11] LifeLines Diet Score [LLDS],[12] Healthy Eating Index [HEI],[13] and alternative Mediterranean diet score [aMED].[14]

In this study, using the LifeLines Cohort Study[15] which prospectively follows 167 729 participants for a minimum of 30 years, we have the unique opportunity to study habitual diet and the development of IBD. We use both *a-posteriori* identified dietary patterns [data-driven method] and *a-priori* dietary scores [target-driven method][16] and link these to the development of IBD. This will generate knowledge of dietary patterns involved in disease development, which can potentially contribute to prevention of these disorders in the future.

2. Methods and Materials

2.1. Cohort description

LifeLines[15] is a multidisciplinary prospective population-based cohort study examining in a unique three-generation design the health and health-related behaviours of 167 729 persons living in the north of The Netherlands. It employs a broad range of investigative procedures in assessing biomedical, sociodemographic, behavioural, physical, and psychological factors contributing to health and disease of the general population. For the present study, dietary information was available for 144 091 participants Figure 1. We excluded participants with age <18 years, implausible Food Frequency Questionnaire [FFQ] data (males <800 or >3934 kcal/day [97.5th percentile], females <500 or >2906 kcal/day [97.5th percentile]),[17] missing data (i.e., missing body mass index [BMI], smoking status, or LLDS), and individuals who already suffered from UC and/or CD at baseline.

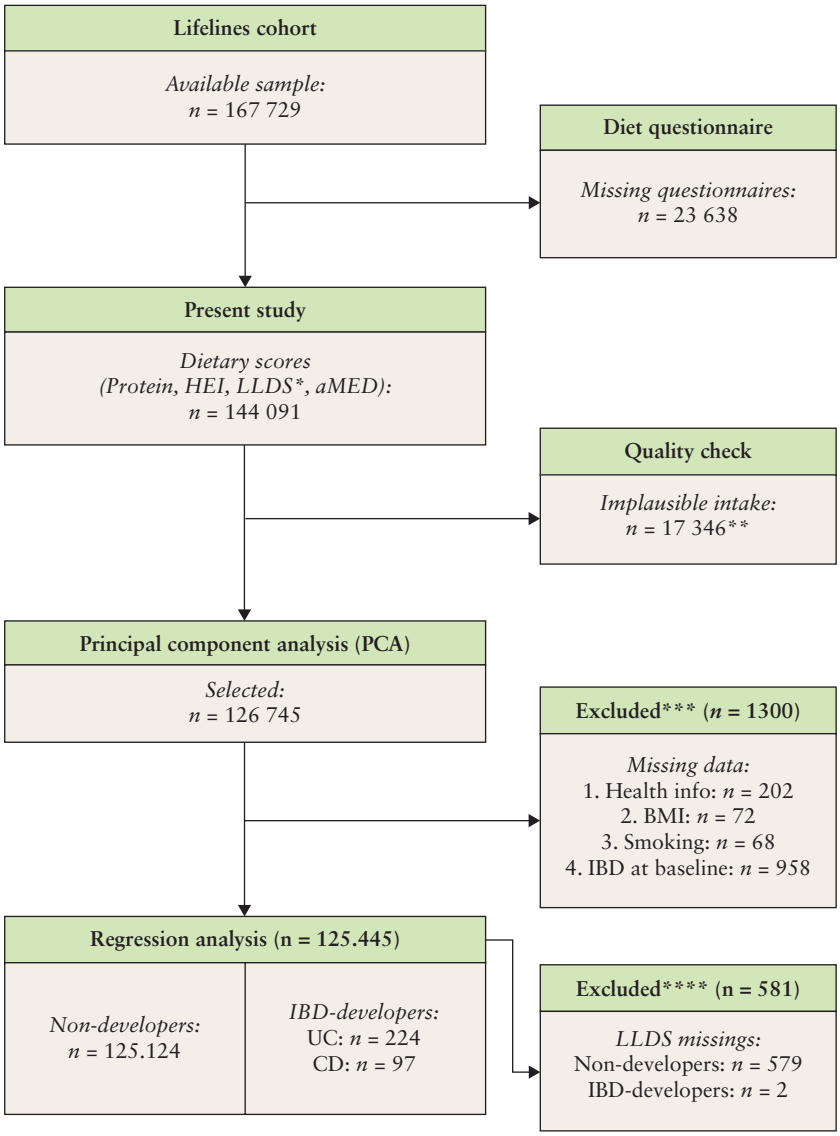

Figure 1. Flowchart of LifeLines participant inclusion. Description: *Only 129 364 LLDS available. **Implausible intake = overall intake for males <800 or >97.5% kcal/day and for females <500 or >97.5% kcal/day. ***Filtering is done sequentially [1 to 4], categories are not mutually exclusive. ****Only excluded when analysing LLDS. *n*, number; HEI, Healthy Eating Index; LLDS, LifeLines Diet Score; aMED, alternative Mediterranean score; PCA, principal component analysis; UC, ulcerative colitis; CD, Crohn's disease.

2.2. Data collection and processing

2.2.1. Disease development data

LifeLines participants reported via multiple questionnaires if they suffer from IBD (baseline assessment 1A [2007–2014], 1B [2011–2015], 1C [2012–2016], 2A [2014–2018], and 3A [2019–2023]) The maximum follow-up time was 14 years. Participants who reported absence of IBD at baseline and registered IBD *de novo* at any follow-up assessment were classified as either 'UC developer' or 'CD developer'. Participants who reported absence of disease at baseline and during every follow-up measurement were classified as 'non-IBD developers'. In case of insufficient or missing data on disease development, participants were regarded as non-IBD developers. In addition, information on potential covariates such as age, sex, BMI, and smoking status [current, former, or never] was retrieved from the LifeLines database.

2.2.2. Dietary data

A semi-quantitative FFQ, which was developed and validated by the division of Human Nutrition of Wageningen University,[18–21] was used to assess habitual dietary intake. The FFQ was administered from 2007 to 2014 [baseline measurement 1A]. Intake over the previous month functioned as a reference period. Intake was reflected in scoring the frequencies of consumption on a four- or seven-item scale along with the usual amount taken. Portion sizes were estimated using natural portions and commonly used household measures. Reported frequencies of consumed food items were linked to the Dutch food composition table [NEVO 2011, RIVM Bilthoven, The Netherlands] to calculate individual mean intake of the reported macronutrients and 110 food items. Food items were grouped into 22 food groups Table S1, available as Supplementary data at *ECCO-JCC* online.

2.2.3. Statistical analyses

All statistical analyses were performed using R [v 3.3.2]. Analyses were corrected for gender, age, BMI, and smoking, and a two-sided *p*-value of <0.05 was considered significant.

2.2.4. Descriptive statistics

Baseline characteristics and dietary intake were presented as mean and standard deviation [SD] for continuous variables and as number and percentages for categorical variables. Between UC developers, CD developers, and non-developers, continuous data were compared using a linear model where the continuous feature acted as a dependent variable and the group as explanatory. An overall *p*-value for the group effect was obtained using a likelihood ratio test between the described model and a null model without group as a covariate. Categorical data were tested by a chi square test [IBD developers vs non-developers].

2.3. Dietary pattern analysis

2.3.1. Principal component analysis [PCA]

PCA, a form of factor analysis, creates sequential linear combinations of food groups to explain the maximal amount of variance in a correlation matrix [i.e., overall diet of individuals]. Single scores are generated for each 'component' as the sum of the products of the strength of the correlation of each food group, with the overall intake reported by the individual. These scores [continuous variables] enable ranking of individuals based on the extent to which they consume foods from groups that are highly weighted in the component.[10]

PCA was conducted with orthogonal [varimax] rotation on 22 standardised food groups [Z-scores] to extract *a-posteriori* identified dietary patterns.[22] Hence, optimal interpretability of the extracted components [dietary patterns] was obtained. Before analysis, suitability of the data was checked using a correlation matrix, Bartlett's Test of Sphericity and the Kaiser–Meyer–Olkin test. Coefficients with absolute values above 0.3 or below -0.3 were considered relevant. Scree plots and interpretability criteria were used to determine the number of patterns to retain. Subsequently for each participant, a factor score [rotated component] per dietary pattern was calculated as the sum of the food group weighted by the factor loadings. Food items may be correlated to several identified dietary patterns and these dietary patterns are not mutually exclusive. Since PCA is sensitive to outliers, an additional *robust* PCA analysis with varimax rotation was performed to check whether the results could be confirmed,[23] using the same assumptions. The R package psych 1.8.12 was used to conduct PCA with orthogonal [varimax] rotation, as for Bartlett's Test of Sphericity and the Kaiser–Meyer–Olkin test.[24] *Robust* PCA was performed using R package pracma v. 2.2.9.[25]

2.3.2. A-priori dietary scores

Four hypothesis-based dietary scores were calculated to allow comparison with previous studies. The Protein Score is based on the hypothesis that a higher overall protein intake with a higher intake of plant-derived protein relative to animal derived protein, is associated with improved health outcomes, including a lower likelihood of developing IBD.[6,11] The LifeLines Diet Score [LLDS] is a population-specific diet quality score that has been based on top-10 most prevalent diseases.[12] Moreover, we calculated internationally used diet quality scores, the Healthy Eating Index [HEI][13] and the [alternative] Mediterranean diet score [aMED].[14] Calculations were conducted according to procedures mentioned in literature.[12,13,26,27] Application and modification of scores are described in the Table S2, available as Supplementary data at *ECCO-JCC* online.

2.4. Logistic regression analysis

To determine whether higher adherence to the identified dietary patterns or dietary scores is associated with IBD development during follow-up, multiple logistic regression analysis was conducted correcting for clinical confounders (gender, age, BMI, and smoking behaviour [as categorical variable: current smoker, former smoker, never smoked]). All dietary patterns extracted from PCA were included in one model; another model included all dietary patterns extracted from *robust* PCA. Each dietary score individually formed a model. Odds ratios [OR] with 95% confidence interval [95% CI], for the association between adherence to the derived dietary patterns or dietary scores and disease development, were calculated. The regression analysis was performed using the glm R function.

2.5. Ethical considerations

LifeLines was approved by the medical ethical committee of the University Medical Centre Groningen. From all individuals, written informed consent was obtained. The study is conducted in accordance with the principles of the Declaration of Helsinki and the UMCG research code. The data underlying this article can be shared on reasonable request; a proposal can be submitted to the

LifeLines Research Office [research@lifelines.nl]. Detailed information on all collected variables within the LifeLines cohort can be found in the catalogue [https://catalogue.lifelines.nl/]. The results are reported according to the STROBE-NUT checklist.[28]

3. Results

3.1. Cohort characteristics

In total 167 729 individuals participated in LifeLines, of whom 126 745 participants were selected for PCA; 125 445 samples were eligible for further analyses through logistic regression analysis Figure 1. The exclusion of participants was discussed in detail in the Methods section. Of these participants, 97 developed CD [0.08%] and 224 developed UC [0.18%]. The prevalence of IBD in our sample was 0.89%, which is comparable to previous reports from Western populations.[4]

Participants had a mean age of 44.8 ± 13.1 years, a BMI of 26.0 ± 4.3 kg/m^2, 58.5% [*n* = 73 568] were females, and 19.2% [*n* = 24 069] were current smokers Table 1. When comparing CD developers with non-developers, no differences were found. UC developers were also compared with non-developers. We found a higher mean age [47.3 ± 13.1 vs 44.8 ± 13.1 years] and a lower percentage of smokers [11.6% vs 19.2%]. When comparing UC developers with CD developers, we reported a higher age [47.3 ± 13.1 vs 43.8 ± 15.0 years] and a lower percentage of smokers [11.6% vs 25.8%].

3.2. Assessment of habitual dietary intake

Table 2 shows the habitual dietary intake of participants. Mean energy intake of all participants was 2017 ± 569 kcal per day. Compared with UC developers and non-developers, CD developers consumed more non-alcoholic beverages [207 ± 213, 210 ± 217 vs 293 ± 301 g/day]. Furthermore, UC developers had a higher intake of vegetables than non-developers [113 ± 59.7 vs 103 ± 57.9 g/day].

3.3. Dietary pattern analysis

The dietary data was found to be likely factorisable [Bartlett's Test: p <0.001, Kaiser–Meyer–Olkin test: 0.69]. Subsequently PCA was performed, identifying five dietary patterns explaining 10.8%, 8.7%, 7.5%, 7.4%, and 7.3%, respectively, [cumulative 41.8%] of total dietary variance Table 3. The first pattern was characterised by high intakes of cooking oils and fats, grain products, potatoes, sugar, cakes, confectionery, condiments and sauces, dairy, and processed meat. The second dietary pattern revealed high intake of snacks, prepared meals, non-alcoholic beverages, condiments and sauces, along with low vegetables and fruit consumption. The third pattern reflected high consumption of red meat, poultry, and processed meat; and the fourth was characterised by high intake of coffee and alcoholic beverages and a low intake of tea. The fifth pattern was characterised by high intake of fish, eggs, nuts, vegetables, legumes, alcoholic beverages, soups, and fruits.

Furthermore, the additional *robust* PCA analysis with varimax rotation Supplementary Table S3, available as Supplementary data at *ECCO-JCC* online identified five patterns. Those *robust* patterns were comparable, although the third and fifth patterns seemed to be reversed to the PCA orthogonal [varimax] rotation analysis Supplementary Figure S1, available as Supplementary data at *ECCO-JCC* online. Since the *robust* PCA confirmed similar patterns, all five patterns were used for regression analysis.

3.4. Logistic regression analysis

Of the five identified dietary patterns, the second pattern Table 4, characterised by high intake of snacks, prepared meals, non-alcoholic beverages, condiments, and sauces along with low vegetables and fruit consumption which is in accordance with a 'Western' pattern, was associated with participants newly reporting CD development [OR: 1.16, 95% CI: 1.03-1.30, p = 0.013]. This association was not confirmed when analysing the second *robust* dietary pattern [OR: 1.20, 95% CI: 0.96-1.50, p = 0.100]

Table 1. Demographic and clinical characteristics of Lifelines participants.

	Complete sample [as used in PCA]	Selection[a] [as used in regression]			
	Complete sample	Non-developers	CD developers	UC developers	*p*-value[b]
	n = 126745	*n* = 125124	*n* = 97	*n* = 224	
Demographic characteristics					
Sex [% female]	73568 [58.5]	73363 [58.5]	63 [64.9]	142 [63.4]	0.060
Age [years]	44.8 ± ᶜ13.1	44.8 ± 13.1	43.8 ± 15.0	47.3 ± 13.1	**0.013***, **
Height [cm]	175 ± 9.36	175 ± 9.36	173 ± 10.1	173 ± 9.05	**0.021****
Weight [kg]	79.7 ± 15.2	79.7 ± 15.2	77.1 ± 13.9	80.1 ± 16.3	0.219
BMI [kg/m^2]	26.0 ± 4.30	26.0 ± 4.29	25.66 ± 4.13	26.6 ± 4.94	0.109
Smoking [%]					
Never smoked	97962 [78.0]	97696 [78.0]	72 [74.2]	194 [86.6]	
Former smoker	3512 [2.8]	3508 [2.8]	0 [0]	4 [1.8]	**0.004***
Current smoker	24120 [19.2]	24069 [19.2]	25 [25.8]	26 [11.6]	

Statistics are performed using a linear regression for continuous variables and chi square test for categorical variables. Values are reported as mean ± standard deviation [SD] or number [%] when appropriate.
BMI, body mass index; CD, Crohn's disease; UC, ulcerative colitis; PCA, principal components analysis..
[a]Participants who did not suffer from inflammatory bowel disease [IBD] at baseline.
[b]Comparison between non-developers, CD- and UC-developers.
Significant *p*-value <0.05 (indicated in **bold**); *CD vs UC; **healthy vs UC.

Table 2. Habitual dietary intake of IBD patients.

	Complete sample [as used in PCA]	Selection[a] [as used in regression]			
	Complete sample	Non-developers	CD developers	UC developers	*p*-value[b]
	n = 126745	*n* = 125124	*n* = 97	*n* = 224	*n* = 125662
Macronutrient intake					
Energy intake [Kcal]	2017 ± 569	2018 ± 590	2052 ± 600	2012 ± 540	0.831
Total protein [g/day]	74.0 ± 19.4	74.0 ± 19.4	73.7 ± 21.1	75.1 ± 19.2	0.706
g/kg	0.95 ± 0.27	0.95 ± 0.276	0.98 ± 0.28	0.96 ± 0.28	0.435
Plant protein [g/day]	30.9 ± 9.9	30.9 ± 9.92	30.3 ± 10.1	31.2 ± 10.1	0.775
g/kg	0.40 ± 0.13	0.40 ± 0.13	0.40 ± 0.13	0.40 ± 0.14	0.863
Animal protein [g/day]	43.2 ± 13.6	43.2 ± 13.6	43.5 ± 15.2	44.0 ± 13.1	0.671
g/kg	0.55 ± 0.18	0.397 ± 0.3	0.40 ± 0.13	0.40 ± 0.14	0.863
Total fat [g/day]	79.8 ± 27.3	79.8 ± 27.3	81.6 ± 26.9	80.3 ± 26.8	0.773
En%	35.3 ± 5.00	35.3 ± 4.98	35.7 ± 4.98	35.6 ± 5.14	0.454
Carbohydrates [g/day]	227 ± 69.3	226 ± 69.3	232 ± 75.6	223 ± 66.3	0.563
En%	44.9 ± 4.66	44.9 ± 5.62	45.0 ± 5.90	44.5 ± 5.68	0.543
Alcohol[c][g/day]	7.17 ± 8.84	7.17 ± 8.84	6.44 ± 8.09	6.69 ± 8.69	0.510
En%[c]	2.48 ± 2.99	2.48 ± 2.98	2.22 ± 2.67	2.28 ± 2.97	0.427
Food group intake [g/day][a]					
Alcoholic beverages	100 ± 146	100 ± 146	79.8 ± 113	90.9 ± 129	0.255
Coffee	417 ± 280	419 ± 280	395 ± 328	421 ± 266	0.713
Condiments and sauces	33.5 ± 22.5	33.5 ± 22.5	34.1 ± 23.2	32.4 ± 19.4	0.761
Cooking oils and fats	23.0 ± 16.3	22.9 ± 16.3	22.8 ± 17.0	23.1 ± 16.0	0.967
Dairy	330 ± 12.2	330 ± 192	329 ± 221	327 ± 191	0.969
Eggs	13.9 ± 14.3	13.9 ± 14.2	13.2 ± 11.3	14.4 ± 15.6	0.769
Fish	12.4 ± 12.8	12.4 ± 12.8	12.6 ± 13.1	12.3 ± 11.9	0.961
Fruits	137 ± 111	137 ± 111	161 ± 115	135 ± 95.8	0.409
Grain products	189 ± 80.6	189 ± 80.6	179 ± 80.1	187 ± 81.4	0.539
Legumes	9.7 ± 15.5	9.68 ± 15.5	9.16 ± 11.5	10.1 ± 15.2	0.871
Non-alcoholic beverages	210 ± 218	210 ± 217	293 ± 301	207 ± 213	**0.0008***, ***
Nuts	12.2 ± 14.3	12.3 ± 14.3	11.6 ± 14.0	12.4 ± 14.7	0.890
Potatoes	90.1 ± 55.3	91.0 ± 55.3	85.7 ± 55.1	91.3 ± 48.5	0.637
Poultry	10.8 ± 8.2	10.8 ± 8.16	10.6 ± 7.45	11.9 ± 8.49	0.129
Prepared meals	30.9 ± 39.5	30.9 ± 39.5	37.1 ± 60.9	27.4 ± 33.4	0.112
Processed meat	29.0 ± 22.0	29.0 ± 22.0	29.5 ± 21.5	28.1 ± 20.6	0.823
Red meat	37.6 ± 19.0	37.6 ± 19.0	36.2 ± 19.6	39.9 ± 19.1	0.145
Snacks	28.8 ± 23.8	28.8 ± 23.8	32.3 ± 53.9	27.9 ± 28.0	0.301
Soups	49.4 ± 52.3	49.4 ± 52.2	48.6 ± 55.4	56.9 ± 65.9	0.096
Sugar, cakes, and confectionery	74.6 ± 45.3	74.7 ± 45.3	74.4 ± 47.8	74.0 ± 43.5	0.973
Tea	245 ± 246	245 ± 246	250 ± 253	244 ± 244	0.974
Vegetables	103 ± 57.9[c]	103 ± 57.9	101 ± 76.5	113 ± 59.7	**0.031****

Values are reported as mean ± standard deviation [SD].
PCA; principal component analysis; En%, macronutrient as percentage of total energy intake [calculated as macronutrient/Kcal * 100];CD, Crohn's disease; UC, ulcerative colitis.
[a]Participants who did not suffer from inflammatory bowel disease [IBD] at baseline.
[b]Comparison of CD vs UC vs non-developers
[c]Crude intake reported, statistics conducted on √-transformed variables.
Significant *p*-value <0.05 (indicated in **bold**); *CD vs UC; **healthy vs UC; ***healthy vs CD.

Supplementary Table S4, available as Supplementary data at *ECCO-JCC* online. The third pattern which can be interpreted as a 'carnivorous' pattern, including high consumption of red meat, poultry, and processed meat, was associated with the risk of UC development [OR: 1.11, 95% CI: 1.01-1.20, p = 0.023], Table 5. This association was confirmed when analysing the reversed third *robust* pattern [OR: 0.84, 95% CI: 0.74-0.95, p = 0.006] Supplementary Table S4. All analyses were corrected for age, gender, BMI, and smoking status.

Regarding *a-priori* dietary scores Table 4, a higher LLDS, reflecting high adherence to dietary guidelines in The Netherlands, was associated with a lower likelihood of newly reporting CD development [OR: 0.95, 95% CI: 0.92-0.99, p = 0.009]. Other dietary patterns were not associated with reporting UC or CD development among participants.

Table 3. Factor loadings of PCA orthogonal [varimax] rotation derived dietary pattern.

	Pattern 1	Pattern 2	Pattern 3	Pattern 4	Pattern 5
Alcoholic beverages	-0.013	0.248	0.080	**0.479**	**0.337**
Coffee	0.167	-0.186	0.009	**0.737**	0.091
Condiments and sauces	**0.455**	**0.333**	0.295	0.113	0.084
Cooking oils and fats	**0.706**	0.013	-0.042	0.076	0.037
Dairy	**0.380**	-0.203	0.016	0.077	-0.004
Eggs	-0.146	0.050	0.188	0.108	**0.489**
Fish	-0.212	-0.096	-0.042	-0.078	**0.609**
Fruits	0.003	**-0.458**	-0.069	-0.255	**0.305**
Grain products	**0.693**	0.071	0.012	-0.004	0.221
Legumes	0.201	-0.070	-0.124	-0.016	**0.413**
Non-alcoholic beverages	0.079	**0.565**	0.092	-0.028	-0.100
Nuts	0.177	0.148	-0.134	-0.002	**0.422**
Potatoes	**0.596**	-0.024	0.244	0.107	0.040
Poultry	-0.111	0.013	**0.715**	-0.123	0.005
Prepared meals	-0.069	**0.584**	-0.004	-0.020	0.180
Processed meat	**0.323**	0.179	**0.451**	0.257	0.087
Red meat	0.158	0.048	**0.790**	0.104	-0.076
Snacks	0.112	**0.728**	0.023	0.071	0.079
Soups	0.096	0.028	0.048	0.109	**0.334**
Sugar, cakes, and confectionery	**0.571**	0.209	-0.072	-0.147	-0.147
Tea	-0.008	-0.149	-0.002	**-0.751**	0.097
Vegetables	0.160	**-0.362**	0.268	-0.217	**0.415**
Explained variance	10.8%	8.7%	7.5%	7.4%	7.3%

Statistics are performed using principal component analysis [PCA]. Factor loadings >0.3 and <-0.3 are indicated in **bold**.

Table 4. Logistic regression analysis on reporting CD development during follow-up.

Model[a,b]	Odds ratio	95% CI	*p*-value
Dietary pattern 1[1]	1.00	0.90-1.11	0.981
Dietary pattern 2[1]	1.16	1.03-1.30	**0.013***
Dietary pattern 3[1]	0.99	0.86-1.13	0.853
Dietary pattern 4[1]	1.01	0.88-1.14	0.921
Dietary pattern 5[1]	0.90	0.77-1.04	0.144
Protein score[2]	0.93	0.86-1.00	0.062
LLDS[3]	0.95	0.92-0.99	**0.009***
HEI[4]	0.99	0.97-1.01	0.371
aMED[5]	0.98	0.86-1.13	0.831

CD, Crohn's disease; CI, confidence interval; LLDS, Lifelines Diet Score; HEI, Healthy Eating Index; aMED, alternative Mediterranean score.
[a]Multiple models are performed [corrected for age, gender, body mass index, and smoking status] and indicated by numbers[1-5].
[b]Dietary pattern extracted from principal component analysis [PCA].
*Significance = *p*-value <0.05 (indicated in **bold**).

Table 5. Logistic regression analysis on reporting UC development during follow-up.

Models[a,b]	Odds ratio	95% CI	*p*-value
Dietary pattern 1[1]	1.00	0.93-1.06	0.941
Dietary pattern 2[1]	1.01	0.92-1.10	0.847
Dietary pattern 3[1]	1.11	1.01-1.20	**0.023***
Dietary pattern 4[1]	1.02	0.94-1.11	0.570
Dietary pattern 5[1]	1.01	0.92-1.12	0.805
Protein score[2]	1.02	0.97-1.07	0.483
LLDS[3]	0.99	0.96-1.01	0.310
HEI[4]	1.01	0.99-1.02	0.421
aMED[5]	1.06	0.97-1.16	0.219

UC, ulcerative colitis; CI, confidence interval; LLDS, Lifelines Diet Score, HEI, Healthy Eating Index; aMED, alternative Mediterranean score.
[a]Multiple models are performed [corrected for age, gender, body mass index, and smoking status] and indicated by numbers[1-5].
[b]Dietary pattern extracted from principal component analysis [PCA].
*Significance = *p*-value <0.05 (indicated in bold).

4. Discussion

In this study, dietary patterns and scores were associated with *de novo* IBD development in a large prospective cohort comprising 125 445 individuals of the general population. Adherence to a 'Western' pattern was associated with increased likelihood of CD development, and a 'carnivorous' pattern with UC development during a maximum follow-up period of 14 years. Furthermore a higher LLDS, reflecting higher relative diet quality, was associated with a lower likelihood of *de novo* CD development. To our knowledge, this is the first study simultaneously investigating the association between both *a-posteriori* dietary patterns and *a-priori* dietary scores, and longitudinal IBD development.

The first dietary pattern was characterised by high intakes of cooking oils and fats, grain products, potatoes, sugar, cakes and confectionery, condiments and sauces, dairy and processed meat, which comports with a 'Traditional [Dutch]' dietary pattern. However vegetables, which are often

consumed together with potatoes, condiments, and meat in the Dutch cuisine,[29] are not reflected in this pattern. Similarly, another Dutch cohort study following 5427 women aged 60–69 for around 8.2 years,[30] did not find a significant association between what they referred to as a traditional Dutch pattern [high intakes of meat, potatoes, vegetables, and alcoholic beverages] and all-cause mortality risk.

The second dietary pattern can be regarded as typical 'Western', consisting of high intake of snacks, prepared meals, non-alcoholic beverages, condiments and sauces, along with low vegetables and fruit consumption. This pattern is frequently discussed in literature.[31–33] According to a recent meta-analysis by Li *et al.*,[34] a dietary pattern can be described as Western if it meets a minimum of two characteristics: high intakes of: [a] refined grains or sugars; [b] red and processed meat; [c] animal protein; [d] animal fats; and [e] high-fat dairy products; [f] a low consumption of fruits and vegetables. The herewith identified 'Western' pattern corresponds with four [a, b, e, and f] of their suggested criteria. In line with our findings, the meta-analysis found an association between Western dietary patterns and risk of CD development (pooled relative risk [RR]: 1.72, 95% CI: 1.01-2.93, $p = 0.045$, $I^2 = 74.8\%$).

The third pattern consists of high consumption of red meat, poultry, and processed meat and will be referred to as the 'carnivorous' dietary pattern and was associated with UC development [OR: 1.11, 95% CI: 1.01-1.22, $p = 0.024$]. This is in line with previous studies, reporting an excessive consumption of red meat and meat products, animal fats, protein, and sugar as risk factors for IBD.[6,35] Recently, Albenberg *et al.*[36] could not establish an association between the amount of red and processed meat consumed and time to symptomatic relapse in a clinical trial, whereas earlier Ge *et al.*[37] demonstrated a greater pooled RR for IBD in a meta-analysis [pooled RR: 1.50, 95% CI: 1.15-1.95, $I^2 = 60.3\%$, $p < 0.001$].

The fourth pattern is characterised by high intake of coffee and alcoholic beverages and a low intake of tea, which we called the 'beverages' pattern. Moderate alcohol consumption is sometimes proposed, although controversial, to be associated with lower all-cause mortality.[38] In a meta-analysis by Nie *et al.*,[39] alcohol consumption was not significantly associated with UC risk whereas coffee consumption showed an inverse association with UC risk, although not significantly. Coffee consumption was previously demonstrated to be preventive for IBD development in an Asian Pacific population.[40]

The fifth pattern is characterised by high intake of fish, eggs, nuts, vegetables, legumes, alcoholic beverages, soups, and fruits. It can be regarded as 'Mediterranean', although the consumption of eggs and soups do not fit. This fifth pattern can also be classified as a 'Healthy [Dutch]' since it includes intake of vegetables, nuts, legumes, fruits, and fish. Such a dietary pattern has been associated with a reduced risk of CD [pooled RR: 0.39, 95% CI: 0.16-0.62, I^2: 67.9%, $p = 0.014$] and UC [pooled RR: 0.61, 95% CI: 0.04-1.18, I^2: 82.8%, $p = 0.003$] in a recent meta-analysis.[16] Surprisingly, we did not find a negative association between our identified dietary pattern and disease development, whereas such an association is often suggested in literature.[7,16,41,42] Traditional Mediterranean dietary habits are changing nowadays and are becoming more Westernised every day.[43] Perhaps our participants consumed a predominantly Westernised Mediterranean diet instead of a Traditional Mediterranean diet, which might explain why we did not find an association in our population.

A-priori determined dietary scores are widely used to measure adherence to current dietary recommendations and associations with health outcomes. Previous research has shown that high intake of animal protein, leading to a lower protein score, was associated with an increased risk of UC.[6] This effect has not been confirmed in our findings regarding the protein score. Nevertheless, animal protein intake is represented in the third dietary pattern by high intake of meat, which pattern was actually associated with UC development.

There was no significant association with IBD risk and the HEI score. Since the HEI is originally composed to suit American data [cups/day] instead of metric data [g/day], we adapted the scoring systems as described in the Table S1. This modification could potentially explain why we did not find an association, whereas previously higher adherence to HEI did show an association with increased diet quality and decreased all-cause mortality.[44]

Conversely the LLDS, reflecting relative diet quality according to the Dutch dietary guidelines,[45] was significantly associated with a decreased risk of developing CD. The aMED score shows similar features to the LLDS, including positive scoring of legumes, nuts, fruits, vegetables, whole grains, and fish. These food groups have been associated with a decreased risk of IBD.[33] Furthermore, dietary fibre intake and long-term high intake of fruit has been associated with a decreased risk for CD.[33,46] A potential mechanism is that dietary fibre interacts with gut microbes and leads to the production of key metabolites such as short-chain fatty acids [SCFAs] which have anti-inflammatory properties.[47,48]

Unexpectedly, no association was found between adherence to a Mediterranean diet, measured by the aMED, and onset of IBD. As aforementioned, no association between our identified 'Mediterranean' dietary pattern and disease development was found either. In contrast, Khalili *et al.*[7] did see that a greater adherence to a Mediterranean diet was associated with significantly lower risk of CD. Due to Westernisation, our Dutch cohort may not be representative enough of the Mediterranean dietary pattern. Besides, fatty acids could not be included in our calculated aMED due to lack of data.

Although the associations between IBD development and hypothesis-based predefined scores (protein score, LLDS [UC only], HEI, and aMED) were not statistically significant, we observed a consistently decreased odds between higher dietary quality and the development of CD [protein score = 0.93, LLDS = 0.95, HEI = 0.99, and aMED = 0.98], Tables 4 and 5.

4.1. Strengths and limitations

A limitation of this study is that overall dietary intake was only assessed at baseline. Consequently, our data cannot conclude on causality but are solely suitable to establish associations between diet and disease development likelihood. Food frequency questionnaires are regarded as a proper and achievable method to assess long-term dietary habits.[49]

Although PCA is a data-driven method, arbitrary decisions need to be made such as how many patterns to retain and how to name or classify a pattern. Moreover some of the dietary scores, except for the LLDS, were developed for other datasets so that adaptations had to be made to fit the Lifelines FFQ data. Furthermore, we feature data of participants self-reporting to have developed IBD over the years. It was not possible to confirm disease status with medical records due to privacy regulations.

Nevertheless, we were able to identify long-term dietary patterns that could be relevant for IBD development and a basis for future intervention studies. Performing a comprehensive dietary pattern analysis in a large prospective population-based cohort, we were able to identify protective dietary patterns as well as potential risk factors for IBD. Target-driven dietary scores have been widely used in the literature. Whereas they are a useful tool to match participant's dietary quality to recommendations, they are based on current [subject to change] knowledge and it is unknown whether these patterns are most advantageous for health.[10] To our knowledge it is the first time that data-driven and target-driven methods were used in parallel to associate dietary patterns with the risk of IBD in such a large cohort.

In conclusion, in this study we have linked long-term dietary patterns to IBD development in 125 445 prospectively followed individuals of the general population. We observed a higher likelihood of developing UC with adherence to a carnivorous dietary pattern and of CD with a Western dietary pattern, whereas following current dietary recommendations for disease prevention [LLDS] was linked to lower development of CD. Our study adds to the importance of evaluating dietary patterns to aid prevention of IBD already at the general population level, and to focus research on wholefood-based strategies and formulated diets for IBD patients.[50] Although these findings need to be confirmed through interventional studies, renouncing a Western or carnivorous dietary pattern has the potential to reduce IBD risk.

Supplementary Data

Supplementary data are available online at *ECCO-JCC* online. The data underlying this article can be shared on reasonable request; a proposal can be submitted to the LifeLines Research Office [research@lifelines.nl]. Detailed information on all collected variables within the LifeLines cohort can be found in the catalogue [https://catalogue.lifelines.nl/].

Funding

LB and RW are supported by a research grant from the Seerave Foundation.

Conflict of Interest

GD reports speakers' fees [outside the submitted work] from Janssen Pharmaceuticals, Takeda, and Pfizer. MC received invited speaking fees [outside the submitted work] from Takeda. RW acted as a consultant for Takeda and received unrestricted research grants from Johnson and Johnson and Takeda Pharmaceuticals.

Acknowledgements

The authors wish to acknowledge the services of the Lifelines Cohort Study, the contributing research centres delivering data to Lifelines, and all the study participants. The Lifelines Biobank initiative has been made possible by subsidy from the Dutch Ministry of Health, Welfare and Sport, the Dutch Ministry of Economic Affairs, the University Medical Centre Groningen [UMCG The Netherlands], University Groningen, and the Northern Provinces of The Netherlands.

Author Contributions

Concept: VP, LB, RW, MC. Data curation: LB, ES, SAS, RW. Formal analysis: VP, LB, ES, SAS, MC. Investigation: VP, LB, ES, SAS, GD, RW, MC. Methodology: VP, LB, ES, SAS, RW, MC. Resources: GD, RW, MC. Supervision: GD, RW, MC. Visualization: VP, ES, SAS. Writing, original draft preparation: VP, ES, SAS, MC. Writing, review, and editing: VP, LB, ES, SAS, GD, RW, MC.

References

1. Ananthakrishnan AN, Bernstein CN, Iliopoulos D, *et al*. Environmental triggers in IBD: a review of progress and evidence. *Nat Rev Gastroenterol Hepatol* 2018:39–49.
2. Ananthakrishnan AN. Epidemiology and risk factors for IBD. *Nat Rev Gastroenterol Hepatol* 2015;12:205–17.
3. Levine A, Sigall Boneh R, Wine E. Evolving role of diet in the pathogenesis and treatment of inflammatory bowel diseases. *Gut* 2018;67:1726–38.
4. Kaplan GG. The global burden of IBD: from 2015 to 2025. *Nat Rev Gastroenterol Hepatol* 2015;12:720–7.
5. Shen W, Gaskins HR, McIntosh MK. Influence of dietary fat on intestinal microbes, inflammation, barrier function and metabolic outcomes. *J Nutr Biochem* 2014;25:270–80.
6. Jantchou P, Morois S, Clavel-Chapelon F, Boutron-Ruault MC, Carbonnel F. Animal protein intake and risk of inflammatory bowel disease: the E3N prospective study. *Am J Gastroenterol* 2010;105:2195–201.
7. Khalili H, Håkansson N, Chan SS, *et al*. Adherence to a Mediterranean diet is associated with a lower risk of later onset Crohn's disease: results from two large prospective cohort studies. *Gut* 2020. doi: 10.1136/gutjnl-2019-319505.
8. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. *Curr Opin Lipidol* 2002;13:3–9.
9. Forbes A, Escher J, Hébuterne X, *et al*. ESPEN guideline: clinical nutrition in inflammatory bowel disease. *Clin Nutr* 2017;36:321–47.
10. Tucker KL. Dietary patterns, approaches, and multicultural perspective. *Appl Physiol Nutr Metab* 2010;35:211–8.
11. Møller G, Sluik D, Ritz C, *et al*. A protein diet score, including plant and animal protein, investigating the association with HbA1c and eGFR: The PREVIEW project. *Nutrients* 2017;9:1–14.
12. Vinke PC, Corpeleijn E, Dekker LH, Jacobs DR, Navis G, Kromhout D. *Development of the Food-Based Lifelines Diet Score [LLDS] and Its Application in 129,369 Lifelines Participants*. Vol. 72. Berlin: Nature Publishing Group: 2018.
13. Krebs-Smith SM, Pannucci TRE, Subar AF, *et al*. Update of the healthy eating index: HEI-2015. *J Acad Nutr Diet* 2018;118:1591–602.
14. Trichopoulou A, Vasilopoulou E. Mediterranean diet and longevity. *Br J Nutr* 2000;84[Suppl 2]:S205–9.
15. Scholtens S, Smidt N, Swertz MA, *et al*. Cohort Profile: LifeLines, a three-generation cohort study and biobank. *Int J Epidemiol* 2015;44:1172–80.
16. Khorshidi M, Djafarian K, Aghayei E, Shab-Bidar S. A posteriori dietary patterns and risk of inflammatory bowel disease: a meta-analysis of observational studies. *Int J Vitam Nutr Res* 2020;90:376–84.
17. Willett WC. *Nutritional Epidemiology*. 3th edn. Oxford, UK: Oxford University Press; 1998.
18. Siebelink E, Geelen A, de Vries JH. Self-reported energy intake by FFQ compared with actual energy intake to maintain body weight in 516 adults. *Br J Nutr* 2011;106:274–81.
19. Sluik D, Geelen A, de Vries JH, *et al*. A national FFQ for the Netherlands [the FFQ-NL 1.0]: validation of a comprehensive FFQ for adults. *Br J Nutr* 2016;116:913–23.
20. Streppel MT, de Vries JH, Meijboom S, *et al*. Relative validity of the food frequency questionnaire used to assess dietary intake in the Leiden Longevity Study. *Nutr J* 2013;12:75.

21. Brouwer-Brolsma EM, Perenboom C, Sluik D, *et al*. Development and external validation of the 'flower-FFQ': a food frequency questionnaire designed for the Lifelines Cohort Study. *Public Health Nutr* 2021. doi: 10.1017/S1368980021002111.
22. Peters V, Spooren CEGM, Pierik MJ, *et al*. Dietary intake pattern is associated with occurrence of flares in IBD patients. *J Crohns Colitis* 2021;**15**:1305–1315.
23. Hubert M, Rousseeuw PJ, Vanden Branden K. ROBPCA: a new approach to robust principal component analysis. *Technometrics* 2005;47:64–79.
24. Revelle W. *Psych: Procedures for Psychological, Psychometric, and Personality Research*. Evanston, IL: Northwestern University; 2019.
25. Borchers HW. *Pracma: Practical Numerical Math Functions. R package version 1.3*. 2015. https://CRAN.R-project.org/package=pracma.
26. Halton TL, Willett WC, Liu S, *et al*. Low-carbohydrate-diet score and the risk of coronary heart disease in women. *N Engl J Med* 2006;355:1991–2002.
27. Fung TT, McCullough ML, Newby PK, *et al*. Diet-quality scores and plasma concentrations of markers of inflammation and endothelial dysfunction. *Am J Clin Nutr* 2005;82:163–73.
28. Lachat C, Hawwash D, Ocké MC, *et al*. Strengthening the Reporting of Observational Studies in Epidemiology-Nutritional Epidemiology [STROBE-nut]: an extension of the STROBE statement. *PLoS Med* 2016;13:e1002036.
29. Buisman ME, Jonkman J. Dietary trends from 1950 to 2010: a Dutch cookbook analysis. *J Nutr Sci* 2019. doi: 10.1017/JNS.2019.3.
30. Waijers PMCM, Ocké MC, Van Rossum CTM, *et al*. Dietary patterns and survival in older Dutch women. *Am J Clin Nutr* 2006;83:1170–6.
31. Rashvand S, Behrooz M, Samsamikor M, Jacobson K, Hekmatdoost A. Dietary patterns and risk of ulcerative colitis: a case-control study. *J Hum Nutr Diet* 2018;31:408–12.
32. Maconi G, Ardizzone S, Cucino C, Bezzio C, Russo AG, Bianchi Porro G. Pre-illness changes in dietary habits and diet as a risk factor for inflammatory bowel disease: a case-control study. *World J Gastroenterol* 2010;16:4297–304.
33. D'Souza S, Levy E, Mack D, *et al*. Dietary patterns and risk for Crohn's disease in children. *Inflamm Bowel Dis* 2008;14:367–73.
34. Li T, Qiu Y, Yang HS, *et al*. Systematic review and meta-analysis: association of a pre-illness Western dietary pattern with the risk of developing inflammatory bowel disease. *J Dig Dis* 2020;21:362–71.
35. Pieczyńska J, Prescha A, Zabłocka-Słowińska K, *et al*. Occurrence of dietary risk factors in inflammatory bowel disease: influence on the nutritional status of patients in clinical remission. *Adv Clin Exp Med* 2019;28:587–92.
36. Albenberg L, Brensinger CM, Wu Q, *et al*. A diet low in red and processed meat does not reduce rate of Crohn's disease flares. *Gastroenterology* 2019;157:128–36.e5.
37. Ge J, Han TJ, Liu J, *et al*. Meat intake and risk of inflammatory bowel disease: a meta-analysis. *Turk J Gastroenterol* 2015;26:492–7.
38. Costanzo S, de Gaetano G, Di Castelnuovo A, Djoussé L, Poli A, van Velden DP. Moderate alcohol consumption and lower total mortality risk: justified doubts or established facts? *Nutr Metab Cardiovasc Dis* 2019;29:1003–8.
39. Nie JY, Zhao Q. Beverage consumption and risk of ulcerative colitis. *Med* 2017. doi: 10.1097/MD.0000000000009070.
40. Ng SC, Tang W, Leong RW, *et al*.; Asia-Pacific Crohn's and Colitis Epidemiology Study ACCESS Group. Environmental risk factors in inflammatory bowel disease: a population-based case-control study in Asia-Pacific. *Gut* 2015;64:1063–71.
41. Godny L, Reshef L, Pfeffer-Gik T, *et al*. Adherence to the Mediterranean diet is associated with decreased fecal calprotectin in patients with ulcerative colitis after pouch surgery. *Eur J Nutr* 2020;59:3183–90.
42. Papada E, Amerikanou C, Forbes A, Kaliora AC. Adherence to Mediterranean diet in Crohn's disease. *Eur J Nutr* 2020;59:1115–21.
43. D'Alessandro A, De Pergola G. The Mediterranean Diet: its definition and evaluation of a priori dietary indexes in primary cardiovascular prevention. *Int J Food Sci Nutr* 2018;69:647–59.
44. Onvani S, Haghighatdoost F, Surkan PJ, Larijani B, Azadbakht L. Adherence to the Healthy Eating Index and Alternative Healthy Eating Index dietary patterns and mortality from all causes, cardiovascular disease and cancer: a meta-analysis of observational studies. *J Hum Nutr Diet* 2017;30:216–26.
45. Brink E, van Rossum C, Postma-Smeets A, *et al*. Development of healthy and sustainable food-based dietary guidelines for the Netherlands. *Public Health Nutr* 2019;22:2419–35.
46. Ananthakrishnan AN, Khalili H, Konijeti GG, *et al*. A prospective study of long-term intake of dietary fiber and risk of Crohn's disease and ulcerative colitis. *Gastroenterology* 2013;145:970–7.
47. Venegas DP, Fuente MKD la, Landskron G, *et al*. Short chain fatty acids [SCFAs]-mediated gut epithelial and immune regulation and its relevance for inflammatory bowel diseases. *Front Immunol* 2019;10:277.
48. Bolte LA, Vich Vila A, Imhann F, *et al*. Long-term dietary patterns are associated with pro-inflammatory and anti-inflammatory features of the gut microbiome. *Gut* 2021;70:1287–98.
49. Willett W. Reproducibility and validity of food-frequency questionnaires. In: Willett W, editor. *Nutritional Epidemiology*. 2nd edn. New York, NY: Oxford University Press; 1998: 321–46.
50. Campmans-Kuijpers MJE, Dijkstra G. Food and food groups in inflammatory bowel disease [IBD]: the design of the Groningen Anti-Inflammatory Diet [GRAID]. *Nutrients* 2021. doi: 10.3390/nu13041067.