Supplementary Appendix 1 - PURE Study Participant Selection Methodology as Excerpted from Teo et al (26).

Selection of Countries
The choice and number of countries selected in PURE reflects a balance between involving a large number of communities in countries at different economic levels, with substantial heterogeneity in social and economic circumstances and policies, and the feasibility of centers to successfully achieve long-term follow-up. Thus, PURE included sites in which investigators are committed to collecting good-quality data for a low-budget study over the planned 10-year follow-up period and did not aim for a strict proportionate sampling of the entire world.

Selection of Communities
Within each country, urban and rural communities were selected based on broad guidelines. A common definition for "community" that is applicable globally is difficult to establish. In PURE, a community was defined as a group of people who have common characteristics and reside in a defined geographic area. A city or large town was not usually considered to be a single community, rather communities from low-, middle-, and high-income areas were selected from sections of the city and the community area defined according to a geographical measure (e.g., a set of contiguous postal code areas or a group of streets or a village). The primary sampling unit for rural areas in many countries was the village. The reason for inclusion of both urban and rural communities is that for many countries, urban and rural 9 environments exhibit distinct characteristics in social and physical environment, and hence, by sampling both, we ensured considerable variation in societal factors across PURE communities. The number of communities selected in each country varied, with the aim to recruit communities with substantial heterogeneity in social and economic circumstances balanced against the capacity of local investigators to maintain follow-up. In some countries (e.g., India, China, Canada, and Colombia), communities from several states/provinces were included to capture regional diversity, in policy, socioeconomic status, culture, and physical environment. In other countries (e.g., Iran, Poland, Sweden, and Zimbabwe), fewer communities were selected.

Selections of Households and Individuals
Within each community, sampling was designed to achieve a broadly representative sample of that community of adults aged between 35 and 70 years. The choice of sampling frame within each center was based on both "representativeness" and feasibility of long-term follow-up, following broad study guidelines. Once a community was identified, where possible, common and standardized approaches were applied to the enumeration of households, identification of individuals, recruitment procedures, and data collection. The method of approaching households differed between regions. For example, in rural areas of India and China, a community announcement was made to the village through contact of a community leader, followed by in-person door-to-door visits of all households. In contrast in Canada, initial contact was by mail followed by telephone inviting members of the households to a central clinic. Households were eligible if at least 1 member of the household was between the 10 ages of 35 and 70 years and the household members intended to continue living in their current home for a further 4 years. For each approach, at least 3 attempts at contact were made. All individuals within these households between 35 and 70 years providing written informed consent were enrolled. When an eligible household or eligible individual in a household refused to participate, demographics and self-reported data about CVD risk factors, education, and history of CVD, cancers and deaths in the households within the two previous years were recorded. To ensure standardization and high data quality, we used a comprehensive operations manual, training workshops, DVDs, regular communication with study personnel and standardized report forms. We entered all data in a customized database programmed with range and consistency

checks which was transmitted electronically to the Population Health Research Institute in Hamilton (Ontario, Canada) where further quality checks were implemented.

Supplementary Appendix 2 – Validation exercise for 20% of IBD participants

Of 467 cases of IBD, we aimed to validate 20% of cases (93 participants). Medical records for these 93 participants (where available in English) were requested from participating centres selected at random, upon which centres were able to provide medical records (such as consultation, imaging, and endoscopy reports) for 88 participants. Manual review of these records was able to confirm these 83 cases included 26 diagnoses of Crohn's disease and 57 diagnoses of ulcerative colitis. 83/93 (89.2%) cases could be verified. For the 5 cases that records were not provided, this was due to lack of access to medical records at that site.

Supplementary Table 1 – PURE food frequency questionnaire validation studies

Country	Validated	Reference dietary method	Reference
Argentina	Yes	Multiple dietary recalls	Dehghan et al. PLoS One· 2012;7(5):e37958
Brazil	Yes	Multiple dietary recalls	Under preparation
Canada	Yes	Multiple dietary recalls	Kelemen L et al. J Am Diet Assoc. 2003 103(9):1178-84
Chile	Yes	Multiple dietary recalls	Dehghan et al. Public Health Nutr. 2013;16 (10):1782-8.
China	Yes	Multiple dietary recalls	Zhao WH et al. Biomedical and environmental sciences. 2010; 23(suppl.), 1-38.
Colombia	Yes	Multiple dietary recalls	Dehghan et al. J Nutr Educ Behav. 2012;44(6):609-13.
Iran	Yes	Multiple dietary recalls	Under preparation
India	Yes	Multiple dietary recalls	Iqbal R et al. Public Health Nutr. 2009; 12(1):12-18 Bharati A et al. Asia Pac J Clin Nutr 2008; 14(1):178-185. Mahajan R et al. The National Medical Journal of India vol. 26, no. 5, 2013
Malaysia	Yes	Multiple dietary recalls	Book chapter
Palestine	Yes	Multiple dietary recalls	Under preparation
Poland	Yes	Multiple dietary recalls	Dehghan et al. J Hum Nutr Diet. 2012; 25(3):225-32
Sweden	Yes		Khani B et al. J Nutr. 2004, 134:1541-1545
South Africa	Yes	Multiple dietary recalls	MacIntyre UE et al. Public health nutr. 2000; 4(1), 63-71
Turkey	Yes	Multiple dietary recalls	Gunes eat al. J Pak Med Assoc. 2015; 65(7):756-63.
UAE	Yes	Multiple dietary recalls	Dehghan et al. Nutr J. 2005;4:18
Kuwait/UAE	Yes	Multiple dietary recalls	Dehghan et al. Saudi Med J 2009; Vol30(1)
Zimbabwe	No	FFQ development	Development of FFQ Merchant et al. Nutr J. 2005;4:37

Supplementary Table 2 - Association Between Total Ultra-processed Food Intake (grams/day) and Development of IBD

IBD			
	<50g/day	50 to <100g/day	≥100g/day	P trend
No. of participants	77,318	12,820	24,928	
No. of events (%)	218 (0.28)	66 (0.51)	180 (0.72)	
Unadjusted HR (95% CI)	1 (reference)	1.99 (1.51 to 2.62)	2.75 (2.25 to 3.35)	<0.001
Min. adjusted HR (95% CI)	1 (reference)	1.42 (1.07 to 1.90)	1.59 (1.26 to 2.01)	0.004
Fully adjusted HR (95% CI)	1 (reference)	1.33 (0.88 to 2.01)	1.73 (1.23 to 2.45)	0.007
Crohn's disease			
No. of participants	77,318	12,820	24,928	
No. of events (%)	34 (0.04)	12 (0.09)	41 (0.16)	
Unadjusted HR (95% CI)	1 (reference)	2.31 (1.20 to 4.47)	4.00 (2.54 to 6.30)	<0.001
Min. adjusted HR (95% CI)	1 (reference)	1.22 (0.61 to 2.45)	1.50 (0.87 to 2.59)	0.34
Fully adjusted HR (95% CI)	1 (reference)	2.42 (0.91 to 6.42)	2.83 (1.17 to 6.83)	0.07
Ulcerative colitis			
No. of participants	77,318	12,820	24,928	
No. of events (%)	184 (0.24)	54 (0.42)	139 (0.56)	
Unadjusted HR (95% CI)	1 (reference)	1.93 (1.43 to 2.62)	2.51 (2.02 to 3.13)	<0.001
Min. adjusted HR (95% CI)	1 (reference)	1.47 (1.07-2.02)	1.60 (1.24 to 2.08)	0.001
Fully adjusted HR (95% CI)	1 (reference)	1.16 (0.73 to 1.87)	1.57 (1.07 to 2.30)	0.05

Heterogeneity of results from Crohn's disease and ulcerative colitis - Chi² p-value=0.595; I² = 0%.
Minimal adjustments are for age, gender, and geographic region. Fully adjusted model includes age, gender, geographic region, education, alcohol intake, smoking status, BMI, total energy intake and location.

Supplementary Table 3 – Association Between Total Ultra-processed Food Intake (servings/day) and Development of IBD in Participants < Age 50

IBD			
	<1 serving/day	1 to <5 servings/day	≥ 5 servings/day	P trend
No. of participants	38,034	11,964	5,108	
No. of events (%)	110 (0.29)	49 (0.41)	34 (0.67)	
Unadjusted HR (95% CI)	1 (reference)	1.54 (1.10 to 2.16)	2.36 (1.61 to 3.47)	<0.001
Min. adjusted HR (95% CI)	1 (reference)	1.26 (0.89 to 1.80)	1.47 (0.95 to 2.29)	0.18
Fully adjusted HR (95% CI)	1 (reference)	1.27 (0.68 to 2.37)	2.02 (1.03 to 3.97)	0.10

Supplementary Table 4 - Association Between Total Ultra-processed Food Intake (servings/day) and Development of IBD in Participants ≥ Age 50

IBD				
	<1 serving/day	1 to <5 servings/day	≥ 5 servings/day	P trend
No. of participants	38,381	13,489	6,634	
No. of events (%)	89 (0.23)	85 (0.63)	61 (0.92)	
Unadjusted HR (95% CI)	1 (reference)	2.94 (2.19 to 3.96)	4.06 (2.93 to 5.62)	<0.001
Min. adjusted HR (95% CI)	1 (reference)	1.48 (1.06 to 2.08)	1.39 (0.94 to 2.06)	0.07
Fully adjusted HR (95% CI)	1 (reference)	1.90 (1.24 to 2.94)	1.76 (1.06 to 2.91)	0.01

Supplementary Table 5 - Association Between Total Processed Meat Intake (servings/day) and Development of IBD

IBD				
	<1 serving/week	1 serving/week to <1 serving/day	≥1 serving/day	P trend
No. of participants	33,827	50,274	3,045	
No. of events (%)	43 (0.13)	263 (0.52)	24 (0.79)	
Unadjusted HR (95% CI)	1 (reference)	4.20 (3.05 to 5.80)	6.30 (3.82 to 10.38)	<0.001
Min. adjusted HR (95% CI)	1 (reference)	2.39 (1.69 to 3.37)	2.36 (1.39 to 4.01)	<0.001
Fully adjusted HR (95% CI)	1 (reference)	1.92 (1.24 to 2.98)	2.07 (1.14 to 3.76)	0.01
Crohn's disease				
No. of participants	33,827	50,274	3,045	
No. of events (%)	2 (0.01)	56 (0.11)	4 (0.13)	
Unadjusted HR (95% CI)	1 (reference)	19.23 (4.69 to 78.79)	22.67 (4.15 to 123.75)	
Min. adjusted HR (95% CI)	1 (reference)	6.13 (1.43 to 26.36)	3.80 (0.66 to 22.04)	
Fully adjusted HR (95% CI)	1 (reference)	3.99 (0.83 to 19.22)	2.50 (0.39 to 16.03)	
Ulcerative colitis				
No. of participants	33,827	50,274	3,045	
No. of events (%)	41 (0.12)	207 (0.41)	20 (0.66)	
Unadjusted HR (95% CI)	1 (reference)	3.47 (2.48 to 4.85)	5.50 (3.22 to 9.39)	
Min. adjusted HR (95% CI)	1 (reference)	2.19 (1.53 to 3.13)	2.43 (1.38 to 4.29)	
Fully adjusted HR (95% CI)	1 (reference)	1.77 (1.11 to 2.80)	2.19 (1.16 to 4.16)	

Heterogeneity of results from Crohn's disease and ulcerative colitis - Chi² p-value=0.929; I² = 0%.
Minimal adjustments are for age, gender, and geographic region.   Fully adjusted model includes age, gender, geographic region, education, alcohol intake, smoking status, BMI, total energy intake and location.

Supplementary Table 6- Association Between Total Soft Drink Intake (servings/day) and Development of IBD

IBD				
	<0.5 serving/week	0.5 to <3 servings/week	≥3 servings/week	P trend
No. of participants	83,067	17,814	12,106	
No. of events (%)	261 (0.31)	95 (0.53)	66 (0.55)	
Unadjusted HR (95% CI)	1 (reference)	1.80 (1.42 to 2.27)	1.91 (1.46 to 2.51)	<0.001
Min. adjusted HR (95% CI)	1 (reference)	1.28 (1.00 to 1.63)	1.30 (0.99 to 1.72)	0.05
Fully adjusted HR (95% CI)	1 (reference)	1.50 (1.10 to 2.03)	1.94 (1.42 to 2.66)	0.001
Crohn's disease				
No. of participants	83,067	17,814	12,106	
No. of events (%)	49 (0.06)	24 (0.13)	17 (0.14)	
Unadjusted HR (95% CI)	1 (reference)	2.41 (1.48 to 3.93)	2.62 (1.51 to 4.54)	
Min. adjusted HR (95% CI)	1 (reference)	1.51 (0.91 to 2.49)	1.54 (0.87 to 2.72)	
Fully adjusted HR (95% CI)	1 (reference)	2.07 (1.12 to 3.83)	2.36 (1.23 to 4.55)	
Ulcerative colitis				
No. of participants	83,067	17,814	12,106	
No. of events (%)	212 (0.26)	71 (0.40)	49 (0.40)	
Unadjusted HR (95% CI)	1 (reference)	1.66 (1.26 to 2.17)	1.75 (1.28 to 2.39)	
Min. adjusted HR (95% CI)	1 (reference)	1.22 (0.92 to 1.61)	1.24 (0.90 to 1.71)	
Fully adjusted HR (95% CI)	1 (reference)	1.35 (0.95 to 1.93)	1.84 (1.29 to 2.64)	

Heterogeneity of results from Crohn's disease and ulcerative colitis - Chi² p-value=0.722; I² = 0%.
Minimal adjustments are for age, gender, and geographic region.   Fully adjusted model includes age, gender, geographic region, education, alcohol intake, smoking status, BMI, total energy intake and location.

Supplementary Table 7- Association Between Sweets Intake (grams/day) and Development of IBD

	0 g/day	<100g/day	≥100g/day	P trend
IBD				
No. of participants	20,812	68,360	23,109	
No. of events (%)	60 (0.29)	283 (0.41)	123 (0.53)	
Unadjusted HR (95% CI)	1 (reference)	1.43 (1.08 to 1.89)	1.91 (1.40 to 2.60)	0.002
Min. adjusted HR (95% CI)	1 (reference)	1.00 (0.75 to 1.34)	1.22 (0.88 to 1.68)	0.18
Fully adjusted HR (95% CI)	1 (reference)	1.86 (1.08 to 3.23)	2.58 (1.44 to 4.62)	0.003
Crohn's disease				
No. of participants	20,812	68,360	23,109	
No. of events (%)	16 (0.08)	51 (0.07)	23 (0.10)	
Unadjusted HR (95% CI)	1 (reference)	0.97 (0.55 to 1.70)	1.34 (0.71 to 2.53)	
Min. adjusted HR (95% CI)	1 (reference)	0.47 (0.26 to 0.87)	0.54 (0.27 to 1.08)	
Fully adjusted HR (95% CI)	1 (reference)	0.74 (0.24 to 2.30)	1.17 (0.35 to 3.84)	
Ulcerative colitis				
No. of participants	20,812	68,360	23,109	
No. of events (%)	44 (0.21)	232 (0.34)	100 (0.43)	
Unadjusted HR (95% CI)	1 (reference)	1.60 (1.16 to 2.20)	2.11 (1.48 to 3.01)	
Min. adjusted HR (95% CI)	1 (reference)	1.20 (0.86 to 1.67)	1.48 (1.02 to 2.13)	
Fully adjusted HR (95% CI)	1 (reference)	2.31 (1.22 to 4.36)	3.08 (1.57 to 6.05)	

Heterogeneity of results from Crohn's disease and ulcerative colitis - Chi² p-value=0.165; I² = 41%.

Minimal adjustments are for age, gender, and geographic region. Fully adjusted model includes age, gender, geographic region, education, alcohol intake, smoking status, BMI, total energy intake and location.

Supplementary Table 8- Association Between Salty Food and Snacks Intake (grams/day) and Development of IBD

IBD				
	<50g/day	50 to <100g/day	≥100g/day	P trend
No. of participants	91,756	15,611	5,849	
No. of events (%)	320 (0.35)	85 (0.54)	59 (1.01)	
Unadjusted HR (95% CI)	1 (reference)	1.63 (1.28 to 2.07)	3.14 (2.38 to 4.15)	<0.001
Min. adjusted HR (95% CI)	1 (reference)	1.11 (0.87 to 1.43)	1.73 (1.28 to 2.32)	0.002
Fully adjusted HR (95% CI)	1 (reference)	1.22 (0.89 to 1.66)	2.06 (1.41 to 3.00)	0.009
Crohn's disease				
No. of participants	91,756	15,611	5,849	
No. of events (%)	60 (0.07)	14 (0.09)	13 (0.22)	
Unadjusted HR (95% CI)	1 (reference)	1.42 (0.80 to 2.55)	3.70 (2.03 to 6.74)	
Min. adjusted HR (95% CI)	1 (reference)	0.72 (0.40 to 1.32)	1.32 (0.70 to 2.49)	
Fully adjusted HR (95% CI)	1 (reference)	0.97 (0.50 to 1.89)	1.73 (0.80 to 3.73)	
Ulcerative colitis				
No. of participants	91,756	15,611	5,849	
No. of events (%)	260 (0.28)	71 (0.45)	46 (0.79)	
Unadjusted HR (95% CI)	1 (reference)	1.67 (1.29 to 2.18)	3.02 (2.20 to 4.13)	
Min. adjusted HR (95% CI)	1 (reference)	1.23 (0.94 to 1.62)	1.84 (1.32 to 2.58)	
Fully adjusted HR (95% CI)	1 (reference)	1.30 (0.91 to 1.85)	2.16 (1.40 to 3.33)	

Heterogeneity of results from Crohn's disease and ulcerative colitis - Chi² p-value=0.245; I² = 27.8%.
Minimal adjustments are for age, gender, and geographic region. Fully adjusted model includes age, gender, geographic region, education, alcohol intake, smoking status, BMI, total energy intake and location.

Supplementary Table 9 - Association of Urinary Sodium (grams/day) and Development of IBD (sensitivity analyses)

	<2.5g/day	2.5 to <3.5g/day	3.5 to <5g/day	≥5g/day	P trend
No. of participants	8,014	9,262	17,618	12,485	
No. of events (%)	42 (0.52)	72 (0.78)	151 (0.86)	100 (0.80)	
Unadjusted HR (95% CI)	1 (reference)	1.59 (1.09 to 2.32)	1.73 (1.23 to 2.43)	1.60 (1.11 to 2.29)	0.02
Min. adjusted HR (95% CI)	1 (reference)	1.82 (1.24 to 2.67)	2.06 (1.46 to 2.92)	2.06 (1.42 to 2.99)	0.004
Fully adjusted HR (95% CI)	1 (reference)	1.18 (0.75 to 1.84)	1.29 (0.86 to 1.93)	0.99 (0.62 to 1.59)	0.45
Excluding China:					
	<2.5g/day	2.5 to <3.5g/day	≥3.5g/day	P trend	
No. of participants	8,014	9,262	30,103		
No. of events (%)	42 (0.52)	72 (0.78)	251 (0.83)		
Unadjusted HR (95% CI)	1 (reference)	1.59 (1.09 to 2.32)	1.67 (1.21 to 2.32)	0.009	
Min. adjusted HR (95% CI)	1 (reference)	1.82 (1.24 to 2.67)	2.06 (1.48 to 2.88)	<0.001	
Fully adjusted HR (95% CI)	1 (reference)	1.17 (0.75 to 1.82)	1.19 (0.81 to 1.75)	0.68	

Minimal adjustments are for age, gender, and geographic region.   Fully adjusted model includes age, gender, geographic region, education, alcohol intake, smoking status, BMI, total energy intake and location.

Supplementary Table 10 - Association of White Meat Intake and Development of IBD

	<1 serving/week	1 to <3 servings/week	≥3 servings/week	P trend
No. of participants	60,837	26,518	25,320	
No. of events (%)	191 (0.31)	152 (0.57)	87 (0.34)	
Unadjusted HR (95% CI)	1 (reference)	1.88 (1.52 to 2.32)	1.17 (0.91 to 1.51)	<0.001
Min. adjusted HR (95% CI)	1 (reference)	1.16 (0.92 to 1.45)	1.06 (0.82 to 1.36)	0.46
Fully adjusted HR (95% CI)	1 (reference)	1.31 (0.99 to 1.74)	1.38 (0.97 to 1.95)	0.10

Minimal adjustments are for age, gender, and geographic region.   Fully adjusted model includes age, gender, geographic region, education, alcohol intake, smoking status, BMI, total energy intake and location.

Supplementary Table 11 - Association of Red Meat (Unprocessed) Intake and Development of IBD

	<3 servings/week	3 to <7 servings/week	≥7 servings/week	P trend
No. of participants	50,694	30,897	27,009	
No. of events (%)	179 (0.35)	115 (0.37)	133 (0.49)	
Unadjusted HR (95% CI)	1 (reference)	1.08 (0.85 to 1.36)	1.42 (1.14 to 1.78)	0.007
Min. adjusted HR (95% CI)	1 (reference)	0.83 (0.65 to 1.05)	1.03 (0.82 to 1.30)	0.18
Fully adjusted HR (95% CI)	1 (reference)	1.00 (0.74 to 1.35)	1.12 (0.81 to 1.53)	0.72

Minimal adjustments are for age, gender, and geographic region. Fully adjusted model includes age, gender, geographic region, education, alcohol intake, smoking status, BMI, total energy intake and location.

Supplementary Table 12 - Association of Dairy Intake and Development of IBD

	<1 serving/day	1 to <2 servings/day	≥2 servings/day	P trend
No. of participants	64,198	22,995	27,388	
No. of events (%)	153 (0.24)	73 (0.32)	205 (0.75)	
Unadjusted HR (95% CI)	1 (reference)	1.33 (1.01 to 1.76)	3.07 (2.49 to 3.79)	<0.001
Min. adjusted HR (95% CI)	1 (reference)	1.06 (0.80 to 1.41)	1.50 (1.15 to 1.96)	0.006
Fully adjusted HR (95% CI)	1 (reference)	1.14 (0.78 to 1.68)	1.26 (0.86 to 1.83)	0.49

Minimal adjustments are for age, gender, and geographic region. Fully adjusted model includes age, gender, geographic region, education, alcohol intake, smoking status, BMI, total energy intake and location.

Supplementary Table 13 - Association of Starch Intake and Development of IBD

	<200g/day	200 to <400g/day	≥400g/day	P trend
No. of participants	9,770	42,878	63,389	
No. of events (%)	48 (0.49)	198 (0.46)	221 (0.35)	
Unadjusted HR (95% CI)	1 (reference)	0.93 (0.68 to 1.27)	0.70 (0.51 to 0.95)	0.005
Min. adjusted HR (95% CI)	1 (reference)	1.22 (0.89 to 1.67)	1.23 (0.89 to 1.71)	0.43
Fully adjusted HR (95% CI)	1 (reference)	1.41 (0.98 to 2.04)	1.50 (0.95 to 2.38)	0.16

Minimal adjustments are for age, gender, and geographic region. Fully adjusted model includes age, gender, geographic region, education, alcohol intake, smoking status, BMI, total energy intake and location.

Supplementary Table 14 - Association of Fruit, Vegetable & Legume Intake and Development of IBD

	<2 servings/day	2 to <6 servings/day	≥6 servings/day	P trend
No. of participants	15,354	64,473	34,754	
No. of events (%)	40 (0.26)	194 (0.30)	197 (0.57)	
Unadjusted HR (95% CI)	1 (reference)	1.16 (0.83 to 1.63)	2.36 (1.68 to 3.32)	<0.001
Min. adjusted HR (95% CI)	1 (reference)	1.02 (0.72 to 1.43)	1.23 (0.86 to 1.77)	0.21
Fully adjusted HR (95% CI)	1 (reference)	0.97 (0.59 to 1.60)	1.31 (0.77 to 2.22)	0.11

Minimal adjustments are for age, gender, and geographic region.   Fully adjusted model includes age, gender, geographic region, education, alcohol intake, smoking status, BMI, total energy intake and location.

Supplementary Table 15 - Association of Fruit Intake and Development of IBD

	<1 serving/day	1 to <4 servings/day	≥4 servings/day	P trend
No. of participants	55,344	46,265	12,972	
No. of events (%)	152 (0.27)	225 (0.49)	54 (0.42)	
Unadjusted HR (95% CI)	1 (reference)	1.90 (1.54 to 2.33)	1.69 (1.24 to 2.30)	<0.001
Min. adjusted HR (95% CI)	1 (reference)	1.21 (0.97 to 1.52)	0.91 (0.65 to 1.27)	0.07
Fully adjusted HR (95% CI)	1 (reference)	1.06 (0.79 to 1.43)	0.79 (0.50 to 1.23)	0.31

Minimal adjustments are for age, gender, and geographic region.   Fully adjusted model includes age, gender, geographic region, education, alcohol intake, smoking status, BMI, total energy intake and location.

Supplementary Table 16 - Association of Vegetable Intake and Development of IBD

	<5 servings/day	5 to <7 servings/day	≥7 servings/day	P trend
No. of participants	99,966	6,771	7,844	
No. of events (%)	313 (0.31)	40 (0.59)	78 (0.99)	
Unadjusted HR (95% CI)	1 (reference)	1.87 (1.35 to 2.60)	3.17 (2.47 to 4.06)	<0.001
Min. adjusted HR (95% CI)	1 (reference)	1.21 (0.86 to 1.70)	1.33 (1.00 to 1.76)	0.11
Fully adjusted HR (95% CI)	1 (reference)	1.43 (0.98 to 2.09)	1.38 (1.00 to 1.89)	0.06

Minimal adjustments are for age, gender, and geographic region.   Fully adjusted model includes age, gender, geographic region, education, alcohol intake, smoking status, BMI, total energy intake and location.

Supplementary Table 17 - Association of Legume Intake and Development of IBD

	<1 serving/week	1 to <3 servings/week	≥3 servings/week	P trend
No. of participants	25,015	33,463	54,197	
No. of events (%)	105 (0.42)	156 (0.47)	169 (0.31)	
Unadjusted HR (95% CI)	1 (reference)	1.11 (0.87 to 1.42)	0.73 (0.57 to 0.93)	<0.001
Min. adjusted HR (95% CI)	1 (reference)	0.97 (0.75 to 1.24)	0.72 (0.57 to 0.92)	0.008
Fully adjusted HR (95% CI)	1 (reference)	0.97 (0.69 to 1.36)	1.20 (0.87 to 1.66)	0.25

Minimal adjustments are for age, gender, and geographic region.   Fully adjusted model includes age, gender, geographic region, education, alcohol intake, smoking status, BMI, total energy intake and location.

Supplementary Table 18 - Association of Fried Foods Intake and Development of IBD

IBD			
	0 servings	<1 serving/day	≥1 serving/day	P trend
No. of participants	18,290	56,606	39,685	
No. of events (%)	44 (0.24)	163 (0.29)	224 (0.56)	
Unadjusted HR (95% CI)	1 (reference)	1.21 (0.86 to 1.68)	2.50 (1.81 to 3.45)	<0.001
Min. adjusted HR (95% CI)	1 (reference)	1.00 (0.72 to 1.41)	1.26 (0.88 to 1.79)	0.12
Fully adjusted HR (95% CI)	1 (reference)	2.35 (1.21 to 4.55)	3.02 (1.51 to 6.03)	0.006
Crohn's disease			
No. of participants	18,290	56,606	39,685	
No. of events (%)	12 (0.07)	28 (0.05)	50 (0.13)	
Unadjusted HR (95% CI)	1 (reference)	0.76 (0.39 to 1.50)	2.04 (1.09 to 3.83)	
Min. adjusted HR (95% CI)	1 (reference)	0.54 (0.27 to 1.08)	0.65 (0.31 to 1.36)	
Fully adjusted HR (95% CI)	1 (reference)	0.68 (0.18 to 2.56)	1.19 (0.30 to 4.66)	
Ulcerative colitis			
No. of participants	18,290	56,606	39,685	
No. of events (%)	32 (0.17)	135 (0.24)	174 (0.44)	
Unadjusted HR (95% CI)	1 (reference)	1.37 (0.93 to 2.02)	2.67 (1.83 to 3.89)	
Min. adjusted HR (95% CI)	1 (reference)	1.18 (0.80 to 1.74)	1.49 (0.99 to 2.23)	
Fully adjusted HR (95% CI)	1 (reference)	3.13 (1.43 to 6.82)	3.76 (1.67 to 8.51)	

Heterogeneity of results from Crohn's disease and ulcerative colitis - Chi² p-value=0.186; I^2 = 37.6%.

Minimal adjustments are for age, gender, and geographic region.   Fully adjusted model includes age, gender, geographic region, education, alcohol intake, smoking status, BMI, total energy intake and location.

Supplementary Figure 1 – Study flow diagram for participant selection
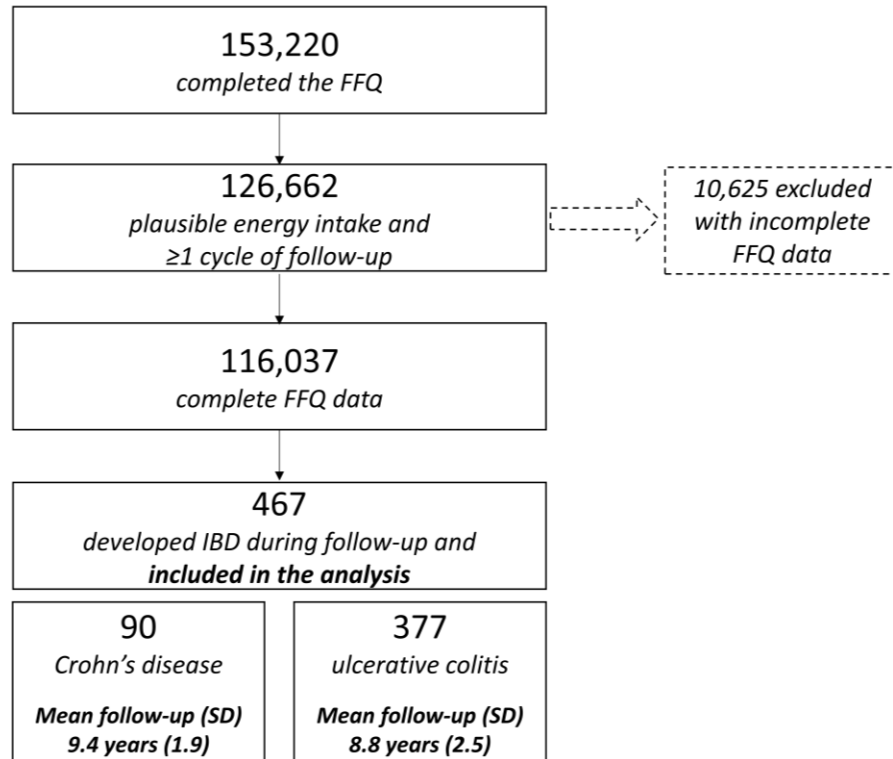
Supplementary Figure 2 – Forest plot of ultra-processed food intake by region comparing ≥1 serving per day with < 1 serving per day

Processed Food Intake by Region	Events (%)		HR (95% CI)
Europe and North America ≥1 servings/day	169 (1.26)		1.46 (0.90 to 2.37)
South America ≥1 servings/day	19 (0.18)		1.14 (0.55 to 2.37)
Africa ≥1 servings/day	8 (0.48)		1.21 (0.46 to 3.21)
Middle East ≥1 servings/day	68 (1.41)		2.26 (1.46 to 3.49)
South Asia ≥1 servings/day	1 (0.07)		0.91 (0.12 to 7.17)
China ≥1 servings/day	2 (0.08)		1.82 (0.41 to 8.10)
Heterogeneity between groups: p = 0.69 Overall, IV (I^2 = 0.0%, p = 0.69)			1.51 (1.05 to 1.97)

1   5