MDPI

Article

Healthy Diet, Polygenic Risk Score, and Upper Gastrointestinal Cancer Risk: A Prospective Study from UK Biobank

Wenmin Liu [1,†], Tianpei Wang [1,2,3,†], Meng Zhu [1,2] and Guangfu Jin [1,2,*]

1. Department of Epidemiology, School of Public Health, Nanjing Medical University, Nanjing 211166, China
2. Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Personalized Medicine and China International Cooperation Center for Environment and Human Health, Nanjing Medical University, Nanjing 211166, China
3. Public Health Institute of Gusu School, The Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou 215000, China
* Correspondence: guangfujin@njmu.edu.cn
† These authors contributed equally to this work.

Abstract: Dietary and genetic factors are considered to be associated with UGI cancer risk. However, examinations of the effect of healthy diet on UGI cancer risk and the extent to which healthy diet modifies the impact of genetic susceptibility on UGI cancer remains limited. Associations were analyzed through Cox regression of the UK Biobank data (n = 415,589). Healthy diet, based on "healthy diet score," was determined according to fruit, vegetables, grains, fish, and meat consumption. We compared adherence to healthy diet and the risk of UGI cancer. We also constructed a UGI polygenic risk score (UGI-PRS) to assess the combined effect of genetic risk and healthy diet. For the results high adherence to healthy diet reduced 24% UGI cancer risk (HR $_{\text{high-quality diet}}$: 0.76 (0.62–0.93), p = 0.009). A combined effect of high genetic risk and unhealthy diet on UGI cancer risk was observed, with HR reaching 1.60 (1.20–2.13), p = 0.001). Among participants with high genetic risk, the absolute five-year incidence risk of UGI cancer was significantly reduced, from 0.16% to 0.10%, by having a healthy diet. In summary, healthy diet decreased UGI cancer risk, and individuals with high genetic risk can attenuate UGI cancer risk by adopting a healthy diet.

Keywords: UGI cancer; dietary pattern; polygenic risk score; prospective cohort; UK Biobank

1. Introduction

Upper gastrointestinal (UGI) cancer, including esophageal cancer (ESC) and gastric cancer (GC), account for 1.7 million new cancer cases and 1.3 million deaths each year worldwide [1]. Previous studies have identified several common environmental risk factors for UGI cancer, including tobacco [2] and alcohol consumption [3], obesity [4], physical activity [5], and dietary factors [6]. Dietary components have received an increasing amount of attention as a potentially modifiable factor [7,8].

It was estimated that 5.1–5.9% of cancer cases each year worldwide can be attributed directly to poor diet [9]. As recently reported by the World Cancer Research Fund International/American Institute for Cancer Research, the role of individual dietary components on UGI cancer risk remains controversial and limited [10]. Rather than individual dietary components, people consume diverse foods together, and the resulting complex combination of dietary components is likely to have interactive or synergistic effects [11]. In this context, dietary pattern analysis has been recommended as an approach because it considers the complexity of overall diet and can potentially facilitate public health interventions [12]. In recent years, cancer prevention guidelines have shifted from reductionist or nutrition-centric approaches to more holistic dietary concepts characterized by dietary patterns. Holistic dietary concepts emphasize how food as a whole can prevent chronic disease, associating nutrients, foods or food groups with health rather than studying the

role played by nutrient/food interactions in health [13–15]. Adherence to a dietary pattern can be assessed using a priori method, which is constructed on the basis of a predefined set of criteria (generally based on guidelines) to measure diet quality in a given population [16], which would be easier to make comparisons between different studies and populations. A meta-analysis of the association of GC risk with dietary patterns indicated that Western dietary patterns (generally considered unhealthy, characterized by an increased consumption of meat, high-fat dairy products, sweets, and starchy foods) were associated with a higher GC risk, while prudent dietary patterns (generally considered healthy, characterized by higher intake of vegetables and fruits) played a protective factor [17]. A case-control study suggested that adherence to a healthy dietary pattern represented by high loadings of vegetables and fruits was associated with a lower risk of GC [18]. However, there is no large-scale prospective cohort study that systematically investigates the association between dietary patterns and UGI cancer risk.
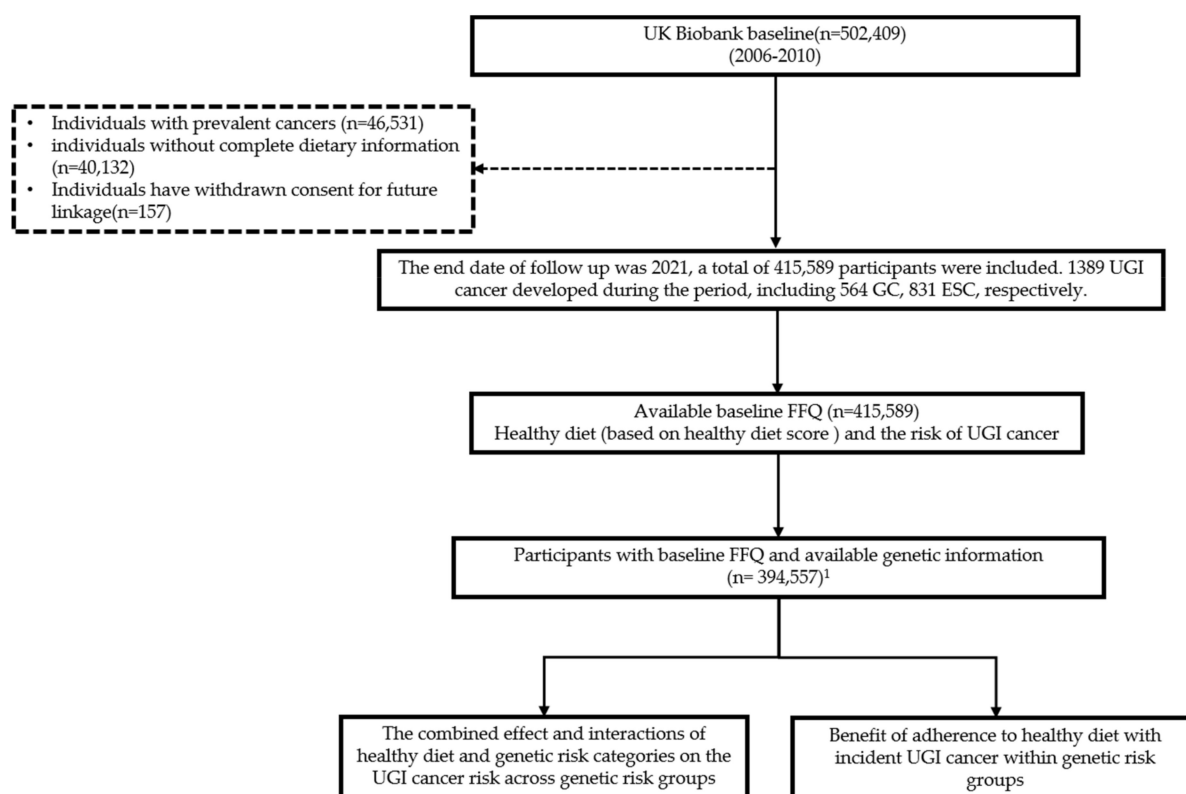
Accumulating evidence has shown that genetic factors have major roles in the development of UGI cancer [19,20]. Recent genome-wide association studies (GWAS) have identified dozens of genetic variants associated with UGI cancer risk [21,22]. The PRSs, gathering genetic contribution and effects of all UGI cancer-associated genetic variants, have been proven to effectively predict incident cases of ESC and GC [23,24]. Both dietary factors and genetic risk play an essential role in the development of the disease. A Gene-Diet Interaction Study from the UK Biobank showed that, compared with those in the lowest intraocular pressure (IOP) polygenic risk score (PRS) quartile who consumed no caffeine, those in the highest IOP PRS quartile who consumed $\geq$321 mg/day showed a 3.90-fold higher glaucoma prevalence [25]. Moreover, one current study suggested that genetic factors modified the association between diet and cardiovascular disease (CVD) [26]. However, previous studies have typically focused on the separate effects of dietary factors and genetic factors on UGI cancer risk. Few studies provided insight into the combined effect of dietary factors and genetic factors on UGI cancer risk. It is unclear whether there is a gene-diet combined effect or interaction in the risk of UGI cancer development, as well as the extent to which participants with a high genetic risk of UGI cancer can offset that risk by adhering to a healthy diet.

In this study, we conducted dietary pattern analysis based on examining the adherence to healthy diet and investigated the association of adherence to healthy diet with UGI cancer risk using UK Biobank data. We also tested the hypothesis that dietary factors and genetic factors jointly contribute to incident UGI cancer and that adopting a healthy diet can attenuate UGI cancer risk for individuals at high genetic risk.

2. Materials and Methods

2.1. Study Design and Participants

UK Biobank is a large, population-based prospective study with genetic and phenotypic data. Between 2006 and 2010, UK Biobank recruited over 500,000 participants from the general population who were aged 40–69 years. Participants were recruited at 22 assessment centers located throughout England, Wales, and Scotland [27]. Participants completed a touch-screen questionnaire, took physical measurements, and provided biological samples at assessment centers. The basic collection details are described elsewhere [28,29]. We excluded participants with prevalent cancer (*n* = 46,531), those who were missing any dietary information data (*n* = 40,132), and individuals who had withdrawn consent for future linkage (*n* = 157), leaving 415,589 participants (193,083 men and 222,506 women) included in the study. First, we examined the association between the degree of adherence to healthy diet defined by healthy diet score and UGI cancer risk. Then, we compared the combined effect and interactions of healthy diet and genetic risk categories on UGI cancer risk across genetic risk groups. Last, we compared the benefit of adherence to a healthy diet within genetic risk groups (Figure 1).

Figure 1. Study design and workflow. [1] For healthy diet and genetic risk on UGI cancer risk across and within genetic risk group analysis, participants without available genetic information were excluded (*n* = 21,032).

2.2. Exposure Measurement

2.2.1. Dietary Intake Assessment

The touch-screen questionnaire, self-completed at baseline, was used to collect the frequency of consumption of the following 12 food items over the previous year with FFQ: beef, lamb, pork, processed meat, oily fish, non-oily fish, fresh fruit, dried fruit, raw vegetables, cooked vegetables, cereal, and bread. We also created new data fields based on food items: (1) Red meat intake, (2) Total fish intake, (3) Total vegetables intake, (4) Total fruit intake, (5) Whole grains intake, and (6) Refined grains intake. We summed beef, lamb and pork intake to create red meat intake. We also summed oily fish and non-oily fish intake to generate total fish intake. To calculate total vegetables and fruit consumption respectively, we aggregated cooked vegetables and salad/raw vegetable intake as total vegetables intake, and fresh fruit and dried fruit as total fruit intake. We divided grains into whole grains and refined grains according to the type of bread and cereal mainly consumed. We defined wholemeal or wholegrain bread, bran cereal, oat cereal, and muesli as whole grains; white bread, brown bread, other bread, biscuit cereal, and other cereals as refined grains. We categorized the 12 food items into 7 food groups, including red meat, processed meat, total fish, total fruit, total vegetables, whole grains and refined grains. We also defined serving size for each baseline food items. For bread and cereal, data were provided for weekly consumption, which were converted into daily consumption. Detailed serving size and coding for each food item/food group are shown in Table S1.

2.2.2. Healthy Diet Score Estimation

We adopted seven dietary factors and cut-offs according to recommendations for dietary priorities on cardiometabolic health [30], that is, increasing fruit, vegetables, whole grains, and fish consumption, and decreasing red meat, processed meat, and refined grains intake. The healthy diet score was calculated using the seven dietary components: Total fruit

≥ 4 servings/day; Total vegetables ≥ 4 servings/day; Total fish ≥ 2 servings/week; Processed meat ≤ 1 serving/week; Red meat ≤ 1.5 servings/week; Whole grains ≥ 3 servings/day; Refined grains ≤ 1.5 servings/day. Each favorable dietary factor was given one point (Table S2). The score ranged from 0 to 7; we defined score 0–1 as low-quality diet, 2–4 as intermediate-quality diet, and 5–7 as high-quality diet, according to data distribution characteristics. Next, we categorized the scores into unfavorable diet (healthy diet score < 4) and favorable diet (healthy diet score ≥ 4).

2.3. PRS Calculation and UGI-PRS Construction

Genotyping process and single nucleotide polymorphisms (SNPs) used in the UKB research have been described elsewhere in detail [31,32]. We extracted variants with $p < 5 \times 10^{-8}$ and minor allele frequency (MAF) ≥ 0.01 from GWAS with the largest sample size in European ancestry [23,33]. For variants that were not available in the UKB genotyping data, their strong correlated SNPs ($r^2 > 0.8$) were included in the present study. If more than one variant correlated in the same locus were reported, the SNPs with the smallest reported p-value were selected by using the linkage disequilibrium clumping procedure (at $r^2 < 0.2$) in PLINK. We excluded SNPs with allele mismatches or MAF differences > 0.10, compared with those in the European population of 1000 Genomes, and palindromic SNPs (A/T, G/C) with an MAF ≥ 0.45. Finally, we estimated site-specific PRS based on 13 SNPs and 3 SNPs for ESC and GC, respectively (Table S3). No SNPs were shared or in high LD ($r^2 > 0.6$) with each other in more than one site-specific PRS. Firstly, site-specific PRS was created following an additive model [34], generated by multiplying the genotype dosage of each risk allele by its respective effect size, summing all alleles together. Then, we built a UGI-PRS to assess UGI cancer risk by summing site-specific PRSs weighted by ESC and GC age-standardized incidence rate in the UK population [35]. Cancer site-specific PRS has been proven to effectively identify individuals with high risk of overall cancers and gastrointestinal cancer risk [36,37]. The UGI-PRS was divided into three levels of genetic risk: low (lowest quintile), moderate (quintiles 2–4), and high (top quintile).

2.4. Outcome Assessment

The outcomes in the study were first primary incident events due to UGI cancer (ESC and GC), which is identified through the national cancer registries of England, Wales, and Scotland, coded by the 10th revision of the International Classification of Diseases (ICD-10), as (C15) and (C16) for ESC and GC, respectively. After four years of baseline recruitment (2006–2010), UGI cancer risk in participants was assessed from baseline up to the UGI cancer diagnosis, death, completion of follow-up, or loss to follow-up, whichever occurred first. The time of risk was calculated according to date the participant attended the assessment center (Data Field: 53), date of cancer diagnosis (Data Field: 40005) and the end date of follow-up. The end date of follow-up was updated to September 2018 for Scotland and to June 2021 for England and Wales. For participants who developed a UGI cancer, time at risk was the interval between the date of cancer diagnosis and the date of attending assessment. For participants without UGI cancer, time at risk was calculated by the end date of follow-up minus date of attending assessment center.

2.5. Statistical Analysis

Cox proportional hazard models were used to investigate the associations between healthy diet and UGI cancer risk and to estimate hazards ratios (HRs) and 95% confidence intervals (CIs) with the time of follow-up used as the timeline variable. The proportional hazard assumptions were checked using Schoenfeld residuals. We determined UGI cancer risk for participants among healthy diet score categories (low-quality diet, intermediate-quality diet, and high-quality diet group). We also compared the UGI cancer risk for per two-point increase in healthy diet score. Furthermore, we investigated the combined effect and interactions of dietary and genetic factors on UGI cancer risk according to healthy diet and genetic risk categories to explore the extent to which healthy diet modified

the associations between genetic susceptibility and UGI cancer risk across genetic risk groups. We examined the results for potential additive and multiplicative interaction between healthy diet and genetic risk [38]. The additive interaction was evaluated using two indexes: the relative excess risk due to the interaction (RERI) and the attributable proportion due to the interaction (AP) [39]. The 95% CIs of the RERI and AP were generated by drawing 5000 bootstrap samples from the estimation data set [40]. If there was no additive interaction, the CIs of the RERI and AP would include 0. In addition, we used RHR (ratio of HR) to evaluate the gene–diet multiplicative interactions by setting variable cross-product terms of the healthy diet with the genetic risk in the models. The 95% CIs of RHR would contain 1 if there was no multiplicative interaction. We also calculated the absolute risk as the percentage of incident UGI cancer cases occurring in each genetic risk group to compare the benefit of adherence to a healthy diet with incident UGI cancer within genetic risk groups. The absolute risk reduction was calculated according to the given groups UGI cancer incidences difference, and then the difference in five-year event rates was extrapolated among given groups. The calculation of 95% CIs for the absolute risk reduction were calculated by drawing 1000 bootstrap samples from the estimation dataset.

Two models were applied in our analyses: minimally adjusted model, adjusted for age at recruitment, sex, Townsend deprivation index, assessment center (10 regions) and ethnic background; fully adjusted model, additionally adjusted for BMI (kg/m^2, <25, 25–29.9, ≥30), glycosylated hemoglobin (HbA1c, mmol/mol, quintiles), smoking status (never, former, current, unknown), alcohol intake frequency (never/rare, twice or less per week, at least three times per week, unknown), education (college or university degree, no degree, unknown), multimorbidity (None, ≥1, unknown), physical activity (<600 MET minutes/week, 600–3000 MET minutes/week, >3000 MET minutes/week) [41] and family cancer history (yes, no, unknown) (Table S4). We additionally adjusted the top 10 genetic principal components of ancestry in the analysis including genetic risk. Missing data were coded as missing proxies (unknown) for categorical variables, while those for continuous variables were imputed with sex-specific median values.

We performed the following sensitivity analysis to further investigate the robustness of our results: (1) excluded participants who reported that they had made a major change in their diet in the past 5 years due to illness in the past 5 years (*n* = 41,292); (2) excluded participants followed up for less than two years (*n* = 1648); (3) excluded non-white participants (*n* = 21,680).

All statistical analyses were performed with R software for version 4.2.0 (R Core Team, Auckland, CA, USA). All *p* values were two-sided and *p* < 0.05 was considered statistically significant.

3. Results

3.1. Participants and Characteristics

A total of 415,589 participants (53.54% women) had available dietary data of this study. The median follow-up period was 12.12 (interquartile range: 11.32–12.84) years for UGI cancer incidence. A total of 1389 UGI cancer developed during the period, including 564 GC and 831 ESC. The baseline characteristics of participants are shown in Table 1. For 1389 UK Biobank participants (mean [SD] age, 61.21 [6.29] years; 27.93% women) with incidents of UGI cancer had a mean (SD) BMI of 28.61 (5.19) kg/m^2. Of all participants, the 16.99% with UGI cancer were current smokers, and 23.18% UGI cancer participants consumed alcohol at least three times per week. The 414,200 participants (mean [SD] age, 56.17 [8.09] years; 53.63% women) had a mean (SD) BMI of 27.39 (4.75) kg/m^2 without UGI cancer. A total of 10.25% participants with UGI cancer were current smokers, and 18.29% UGI cancer participants consumed alcohol at least three times per week.

Table 1. Baseline characteristics of participants in UK Biobank [1].

	Participants	
	With Incident UGI Cancer (*n* = 1389)	**Without UGI Cancer (*n* = 414,200)**
Age at baseline, y	61.21 ± 6.29	56.17 ± 8.09
Female	388 (27.93)	222,118 (53.63)
Townsend deprivation index, means ± SD	−1.04 ± 3.23	−1.40 ± 3.03
BMI, means ± SD	28.61 ± 5.19	27.39 ± 4.75
HbA1c, mmol/mol, means ± SD	38.12 ± 7.98	35.94 ± 6.47
Physical activity, MET minutes/week		
<600	259 (18.65)	63,772 (15.4)
600–3000	774 (55.72)	244,531 (59.04)
>3000	356 (25.63)	105,897 (25.57)
Ethnicity		
White	1346 (96.9)	392,733 (94.82)
Nonwhite	38 (2.74)	20,197 (4.88)
Unknown	5 (0.36)	1270 (0.31)
Education		
College or university degree	352 (25.34)	139,657 (33.72)
No degree	1023 (73.65)	271,190 (65.47)
Unknown	14 (1.01)	3353 (0.81)
Smoking status		
Never	506 (36.43)	228,680 (55.21)
Former	640 (46.08)	141,909 (34.26)
Current	236 (16.99)	42,454 (10.25)
Unknown	7 (0.5)	1157 (0.28)
Alcohol intake frequency		
Never/rare	621 (44.71)	184,431 (44.53)
Twice or less per week	445 (32.04)	153,836 (37.14)
At least three times per week	322 (23.18)	75,752 (18.29)
Unknown	1 (0.07)	181 (0.04)
Health status		
Multimorbidity, n (%)		
None	216 (15.55)	107,012 (25.84)
≥1	1172 (84.38)	306,847 (74.08)
Unknown	1 (0.07)	341 (0.08)
Family cancer history		
yes	832 (59.9)	257,969 (62.28)
no	380 (27.36)	109,695 (26.48)
Unknown	177 (12.74)	46,536 (11.24)

[1] Values are presented as mean ± SD or *n* (%) unless otherwise indicated.

3.2. Healthy Diet and the Risk of UGI Cancer

The association between adherence to healthy diet and UGI cancer risk was shown in Table 2. Individuals with a high-quality diet that included high intake of fruit, vegetables, fish and whole grains and reduced amount of red meat, processed meat and refined grains had a lower risk of UGI cancer incidents compared with those in low-quality diet group, with HR of 0.76 (95% CI: 0.62–0.93, *p* = 0.009). Having a two-point increase in healthy diet score was associated with a higher UGI cancer risk, with HR of 0.90 (95% CI: 0.83–0.97, *p* = 0.006). Similar results were noted in a series of sensitivity analyses (Table S5).

Table 2. Associations between healthy diet score and the risk of UGI cancer.

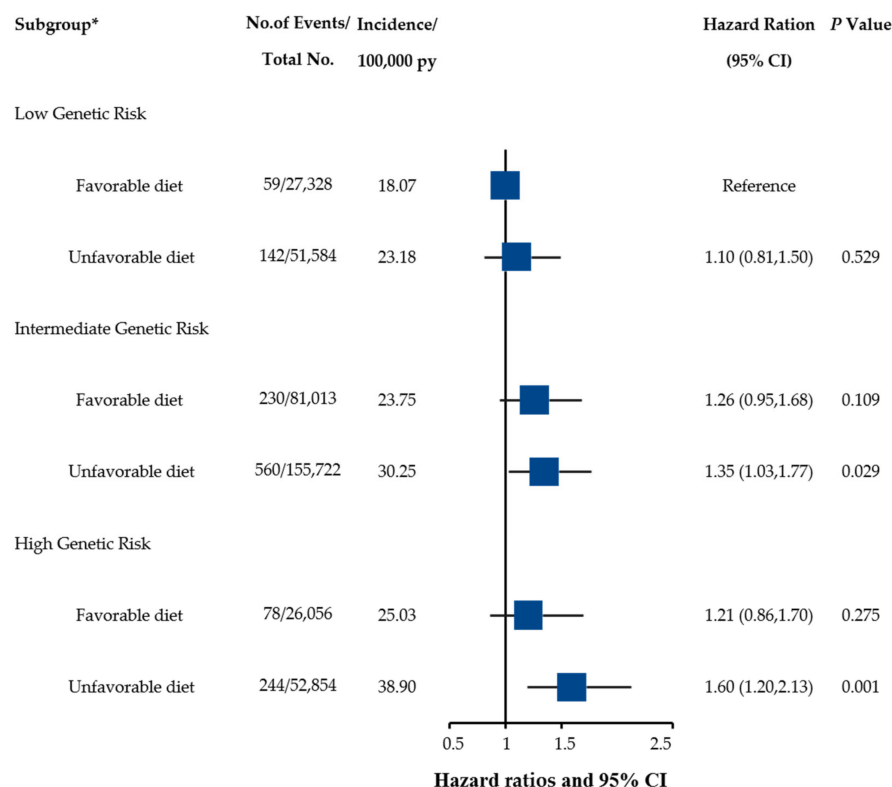| Healthy Diet	Total No. (Cases)	Minimally Adjusted Model [1]		Fully Adjusted Model [2]	
		HR (95% CI)	*p* Value	HR (95% CI)	*p* Value
Healthy diet score [3]					
Low-quality diet (0–1)	64,171 (304)	1.00 (ref)		1.00 (ref)	
Intermediate-quality diet (2–4)	297,417 (943)	0.81 (0.71, 0.92)	0.001	0.87 (0.77, 1.00)	0.047

Table 2. *Cont.*

Healthy Diet	Total No. (Cases)	Minimally Adjusted Model [1]		Fully Adjusted Model [2]	
		HR (95% CI)	*p* Value	HR (95% CI)	*p* Value
High-quality diet (5–7)	54,001 (142)	0.66 (0.54, 0.81)	<0.001	0.76 (0.62, 0.93)	0.009
Per two-point score increase	415,589 (1389)	0.84 (0.78, 0.91)	<0.001	0.90 (0.83, 0.97)	0.006
p for trend			<0.001		0.007

Definition of abbreviations: HR, hazard ratio; 95% CI: 95% confidence interval; ref, reference. 1 Minimally adjusted model: adjusted for age at recruitment, sex, assessment center (10 regions), Townsend deprivation index and ethnicity. 2 Fully adjusted model: minimally adjusted model additionally adjusted for education, BMI, glycosylated hemoglobin (HbAlc), smoking status, alcohol intake frequency, physical activity, multimorbidity and family history of cancer. 3 Healthy diet score: using available data from UK Biobank Food Frequency Questionnaire at baseline; Health diet score ranged from 0 to 7. Fruits: ≥4 servings/day; Vegetables: ≥4 servings/day; Fish: ≥2 servings/week; Processed meats: ≤1 serving/week; Unprocessed red meats: ≤1.5 servings/week; Whole grains: ≥3 servings/day; Refined grains: ≤1.5 servings/day.

3.3. Combined Effect and Interactions of Healthy Diet and Genetic Risk on UGI Cancer Risk

We determined that participants who had an unhealthy diet and were in a high genetic risk group had an approximately 1.60-fold risk of UGI cancer risk, with HR reaching 1.60 (95% CI: 1.20–2.13, *p* = 0.001), when compared with participants with a healthy diet and low genetic risk (Figure 2). The results of the sensitivity analysis did not change materially (Figure S1A–C). The RERI, AP, and RHR were not significant, which indicated no additive and multiplicative interactions of healthy diet and genetic risk on the risk of UGI cancer (Table 3).

Subgroup*	No.of Events/ Total No.	Incidence/ 100,000 py	Hazard Ration (95% CI)	*P* Value
Low Genetic Risk				
Favorable diet	59/27,328	18.07	Reference	
Unfavorable diet	142/51,584	23.18	1.10 (0.81,1.50)	0.529
Intermediate Genetic Risk				
Favorable diet	230/81,013	23.75	1.26 (0.95,1.68)	0.109
Unfavorable diet	560/155,722	30.25	1.35 (1.03,1.77)	0.029
High Genetic Risk				
Favorable diet	78/26,056	25.03	1.21 (0.86,1.70)	0.275
Unfavorable diet	244/52,854	38.90	1.60 (1.20,2.13)	0.001

Hazard ratios and 95% CI

Figure 2. Risk of incident UGI cancer according to healthy diet and genetic risk categories in the UKB cohort. The HRs were estimated using Cox proportional hazard models with adjustment for age at recruitment, sex, assessment center (10 regions), ethnicity, Townsend deprivation index, education, BMI, glycosylated hemoglobin (HbA1c), smoking status, alcohol intake frequency, physical activity, multimorbidity, family history of cancer, and the first 10 principal components of ancestry. * For healthy diet and genetic risk on UGI cancer risk across and within genetic risk group analysis, participants without available genetic information were excluded (*n* = 21,032). Unfavorable diet (healthy diet score < 4) and Favorable diet (healthy diet score ≥ 4).

Table 3. Interaction between diet and genetic risk [1].

	PRS *	
	Intermediate	**High**
RERI (95% CI)	−0.01 (−0.47–0.31)	0.28 (−0.23–0.67)
AP (95% CI)	−0.01 (−0.29–0.26)	0.18 (−0.13–0.45)
RHR (95% CI)	1.03 (0.73–1.45)	0.84 (0.56–1.24)

Definition of abbreviations: RERI = relative excess risk due to the interaction; AP = attributable proportion due to the interaction; RHR = ratio of hazard ratio. * Defined by PRS: low (lowest quintile), intermediate (quintiles 2–4), and high (quintile 5). [1] Cox proportional hazards regression is adjusted for age at recruitment, sex, assessment center (10 regions), Townsend deprivation index, ethnicity, education, BMI, glycosylated hemoglobin (HbAlc), smoking status, alcohol intake frequency, physical activity, multimorbidity, and family history of cancer.

3.4. Benefits of Adherence to a Healthy Diet with UGI Cancer Risk

In further stratification analyses with an unhealthy dietary pattern as the reference group according to genetic risk categories, we found that in the intermediate and high genetic risk groups, similar risk reduction for UGI cancer were observed in those who adhered to a healthy dietary pattern compared to those who adhered to an unhealthy dietary pattern. Among participants with an intermediate genetic risk, the absolute five-year incidence risk of UGI cancer were 0.13 for participants with an unhealthy dietary pattern versus 0.11 for those with a healthy dietary pattern. Similarly, for individuals with high genetic risk, the absolute five-year incidence risk of UGI cancer decreased from 0.16 for participants with an unhealthy dietary pattern to 0.10 for those with a healthy dietary pattern (Table 4). The results of sensitivity analyses were similarly (Table S6).

Table 4. UGI cancer risk associated with healthy diet by genetic risk level in the UKB cohort [1].

Dietary Pattern	Low Genetic Risk		Intermediate Genetic Risk		High Genetic Risk	
	Unfavorable	**Favorable**	**Unfavorable**	**Favorable**	**Unfavorable**	**Favorable**
No. of cases/Person-years	142/61,2672	59/326,606	560/185,1465	230/968,645	244/627,221	78/311,605
HR (95% CI)	Ref.	0.85 (0.63–1.17)	Ref.	0.94 (0.80–1.10)	Ref.	0.78 (0.60–1.01)
p value		0.323		0.417		0.057
Absolute risk (%)-5 years (95% CI)	0.10 (0.07–0.12)	0.08 (0.05–0.10)	0.13 (0.12–0.15)	0.11 (0.09–0.12)	0.16 (0.13–0.19)	0.10 (0.07–0.13)
Absolute risk reduction (%)-5 years (95% CI)	Ref.	0.02 (−0.06–0.49)	Ref.	0.03 (0.01–0.05)	Ref.	0.06 (0.02–0.09)

[1] Cox proportional hazards regression is adjusted for age at recruitment, sex, assessment center (10 regions), Townsend deprivation index, ethnicity, education, BMI, glycosylated hemoglobin (HbAlc), smoking status, alcohol intake frequency, physical activity, multimorbidity and family history of cancer. Unfavorable dietary pattern (healthy diet score < 4) and Favorable dietary pattern (healthy diet score ≥ 4).

4. Discussion

In this large, prospective study using UK Biobank, we investigated dietary pattern analyses based on healthy diet and UGI cancer risk. We found that improving the quality of healthy diet was associated with a lower risk of UGI cancer. Across genetic risk groups, analysis further showed that individuals with high genetic risk and an unhealthy dietary pattern were at a greater risk of UGI cancer compared to those with low genetic risk and a healthy dietary pattern. Within genetic risk groups, analysis indicated that adherence to a healthy dietary pattern was consistently associated with a decreased absolute five-year incidence risk of UGI cancer in intermediate and high genetic risk groups.

Current studies suggested that dietary patterns analyses are regarded as good ways to explore diet and cancer risk. A systematic review and meta-analysis from prospective cohort studies supported an association between healthy dietary patterns and decreased risks of colon and breast cancer [42]. One study that focused on nutrition and breast cancer showed that adherence to a healthy dietary pattern might improve overall survival after diagnosis of breast cancer [43]. We performed dietary pattern analyses based on healthy diet score and the risk of UGI cancer. A systematic review and meta-analysis on dietary

patterns and gastric cancer risk indicated that there is an approximately two-fold difference in GC risk between a 'prudent/healthy' diet, and a 'Western/unhealthy' diet [17]. A population-based case-control study suggested that a diet high in fruit and vegetables may decrease the risk of ESC cancer [44]. Another systematic review and meta-analysis suggested that a healthy dietary pattern was significantly associated with a decreased risk of ESC [45]. Our study also found similar results, i.e., that adherence to a healthy diet reduced the UGI cancer risk. We also compared the benefit of adherence to a healthy dietary pattern within genetic risk groups based on the calculation of absolute five-year incidence risk of UGI cancer. We found that individuals with intermediate and high genetic risk who adopted a healthy diet had a decreased risk of developing UGI cancer. For participants with high genetic risk, the absolute five-year incidence risk of UGI cancer was significantly reduced from 0.16% to 0.10% by having a healthy diet. Taken together, our findings along with previous evidence not only demonstrated the significance of adherence to healthy diet, but also provided collective support for public health interventions to promote a healthy dietary pattern for everyone, especially people with intermediate or high genetic risks, which will ultimately lead to a reduction of UGI cancer burden.

It has been estimated that ESC and GC could be prevented in 54% and 59% of patients in the UK, respectively [46]. It is important to understand the contribution of modifiable risk factors to UGI cancer and how they affect or add to the inherited genetic factors. At present, several studies have summarized the association between diet and nutrition and the UGI cancer risk; however, reported meta-analytic estimates from observational studies may not represent causality. Instead, they may result from common biases across studies, such as exposure measurement error, residual confounding, and publication bias, and thereby weaken the strength of the scientific evidence [47–49]. In addition, few studies have focused on the combined effect and interactions of gene–diet on the risk of UGI cancer. We systematically and comprehensively investigated the association between modifiable dietary factors with UGI cancer risk and tested the hypothesis that UGI cancer risk can be modified or reduced by adopting a healthy diet in a large prospective cohort study.

UK Biobank is a large, general population-based prospective cohort, which provides health outcomes and a wide range of potential confounders, including diet. One of the inevitable problems with large sample studies is that p values are more likely to be statistically different. In detail, a statistical p value is the distance between the data and the null hypothesis measured by an estimate of the parameter of interest. This distance is usually measured in terms of the standard deviation (standard error). The standard error shrinks as the sample size increases; in a very large sample, the standard error becomes very small, which leads to a statistically significant distance between the estimate and the null hypothesis that may be negligible. Therefore, to reduce type I errors, the null hypothesis cannot be rejected by the p-value alone in a large sample study. These problems can be solved by additionally reporting effect sizes and 95% confidence intervals (CI) [50]. In our study, we provided 95% CI as well as p values to more cautiously infer the association between healthy diet and UGI cancer.

The present study has several limitations. First, participants in the UK Biobank are of European descent; therefore, the summary statistics should be generalized to the general population with caution. Secondly, the use of self-reported recall of FFQ could introduce some level of recall bias. Third, it is generally accepted that associations between nutrients and disease should only be considered primary if the effects are independent of energy intake [51]. We were not able to adjust for total energy intake because the baseline touchscreen brief FFQ only covered some commonly consumed foods. Therefore, our findings may be biased by the differences in body size, physical activity, and metabolic efficiency resulting from energy intake. Last, covariates were evaluated only once at baseline, and changes during the follow-up or competitive risk of other illnesses may have an effect on risk estimates.

5. Conclusions

Our findings confirm and broaden the results from previous studies. Healthy diet was associated with a lower risk of UGI cancer. Dietary factors and genetic risk had a combined effect on risk of UGI cancer. Individuals with high genetic risk can attenuate UGI cancer risk by adopting a healthy dietary pattern.

References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *Cancer J. Clin.* **2021**, *71*, 209–249. [CrossRef] [PubMed]
2. Wang, S.M.; Katki, H.A.; Graubard, B.I.; Kahle, L.L.; Chaturvedi, A.; Matthews, C.E.; Freedman, N.D.; Abnet, C.C. Population Attributable Risks of Subtypes of Esophageal and Gastric Cancers in the United States. *Am. J. Gastroenterol.* **2021**, *116*, 1844–1852. [CrossRef]
3. Lu, L.; Mullins, C.S.; Schafmayer, C.; Zeißig, S.; Linnebacher, M. A global assessment of recent trends in gastrointestinal cancer and lifestyle-associated risk factors. *Cancer Commun.* **2021**, *41*, 1137–1151. [CrossRef] [PubMed]
4. Avgerinos, K.I.; Spyrou, N.; Mantzoros, C.S.; Dalamaga, M. Obesity and cancer risk: Emerging biological mechanisms and perspectives. *Metab. Clin. Exp.* **2019**, *92*, 121–135. [CrossRef] [PubMed]
5. Behrens, G.; Jochem, C.; Keimling, M.; Ricci, C.; Schmid, D.; Leitzmann, M.F. The association between physical activity and gastroesophageal cancer: Systematic review and meta-analysis. *Eur. J. Epidemiol.* **2014**, *29*, 151–170. [CrossRef]
6. Navarro Silvera, S.A.; Mayne, S.T.; Risch, H.A.; Gammon, M.D.; Vaughan, T.; Chow, W.H.; Dubin, J.A.; Dubrow, R.; Schoenberg, J.; Stanford, J.L.; et al. Principal component analysis of dietary and lifestyle patterns in relation to risk of subtypes of esophageal and gastric cancer. *Ann. Epidemiol.* **2011**, *21*, 543–550. [CrossRef]
7. Gonzalez, C.A.; Riboli, E. Diet and cancer prevention: Contributions from the European Prospective Investigation into Cancer and Nutrition (EPIC) study. *Eur. J. Cancer* **2010**, *46*, 2555–2562. [CrossRef]
8. Abnet, C.C.; Corley, D.A.; Freedman, N.D.; Kamangar, F. Diet and upper gastrointestinal malignancies. *Gastroenterology* **2015**, *148*, 1234–1243.e4. [CrossRef]
9. GBD 2019 Cancer Risk Factors Collaborators. The global burden of cancer attributable to risk factors, 2010–2019: A systematic analysis for the Global Burden of Disease Study 2019. *Lancet* **2022**, *400*, 563–591. [CrossRef]
10. World Cancer Research Fund; American Institute for Cancer Research. Diet, Nutrition, Physical Activity and Cancer: A Global Perspective. Continuous Update Project Expert Report 2018. Available online: http://dietandcancerreport.org (accessed on 11 March 2019).
11. Hu, F.B. Dietary pattern analysis: A new direction in nutritional epidemiology. *Curr. Opin. Lipidol.* **2002**, *13*, 3–9. [CrossRef]
12. Sacks, F.M.; Obarzanek, E.; Windhauser, M.M.; Svetkey, L.P.; Vollmer, W.M.; McCullough, M.; Karanja, N.; Lin, P.H.; Steele, P.; Proschan, M.A.; et al. Rationale and design of the Dietary Approaches to Stop Hypertension trial (DASH). A multicenter controlled-feeding study of dietary patterns to lower blood pressure. *Ann. Epidemiol.* **1995**, *5*, 108–118. [CrossRef]
13. Fardet, A.; Rock, E. Toward a new philosophy of preventive nutrition: From a reductionist to a holistic paradigm to improve nutritional recommendations. *Adv. Nutr.* **2014**, *5*, 430–446. [CrossRef]
14. Fardet, A.; Rock, E. Perspective: Reductionist Nutrition Research Has Meaning Only within the Framework of Holistic and Ethical Thinking. *Adv. Nutr.* **2018**, *9*, 655–670. [CrossRef]
15. Fardet, A.; Rock, E. Exclusive reductionism, chronic diseases and nutritional confusion: The degree of processing as a lever for improving public health. *Crit. Rev. Food Sci. Nutr.* **2022**, *62*, 2784–2799. [CrossRef]
16. Steck, S.E.; Murphy, E.A. Dietary patterns and cancer risk. *Nat. Rev. Cancer* **2020**, *20*, 125–138. [CrossRef]
17. Bertuccio, P.; Rosato, V.; Andreano, A.; Ferraroni, M.; Decarli, A.; Edefonti, V.; La Vecchia, C. Dietary patterns and gastric cancer risk: A systematic review and meta-analysis. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **2013**, *24*, 1450–1458. [CrossRef]
18. Kim, J.H.; Lee, J.; Choi, I.J.; Kim, Y.I.; Kim, J. Dietary patterns and gastric cancer risk in a Korean population: A case-control study. *Eur. J. Nutr.* **2021**, *60*, 389–397. [CrossRef]
19. Mucci, L.A.; Hjelmborg, J.B.; Harris, J.R.; Czene, K.; Havelick, D.J.; Scheike, T.; Graff, R.E.; Holst, K.; Möller, S.; Unger, R.H.; et al. Familial Risk and Heritability of Cancer Among Twins in Nordic Countries. *JAMA* **2016**, *315*, 68–76. [CrossRef]
20. Sun, W.Y.; Yang, H.; Wang, X.K.; Fan, J.H.; Qiao, Y.L.; Taylor, P.R. The Association Between Family History of Upper Gastrointestinal Cancer and the Risk of Death from Upper Gastrointestinal Cancer-based on Linxian Dysplasia Nutrition Intervention Trial (NIT) Cohort. *Front. Oncol.* **2022**, *12*, 897534. [CrossRef]
21. Gharahkhani, P.; Fitzgerald, R.C.; Vaughan, T.L.; Palles, C.; Gockel, I.; Tomlinson, I.; Buas, M.F.; May, A.; Gerges, C.; Anders, M.; et al. Genome-wide association studies in oesophageal adenocarcinoma and Barrett's oesophagus: A large-scale meta-analysis. *Lancet Oncol.* **2016**, *17*, 1363–1373. [CrossRef]
22. Yan, C.; Zhu, M.; Ding, Y.; Yang, M.; Wang, M.; Li, G.; Ren, C.; Huang, T.; Yang, W.; He, B.; et al. Meta-analysis of genome-wide association studies and functional assays decipher susceptibility genes for gastric cancer in Chinese populations. *Gut* **2020**, *69*, 641–651. [CrossRef] [PubMed]
23. Kunzmann, A.T.; Cañadas Garre, M.; Thrift, A.P.; McMenamin, Ú.C.; Johnston, B.T.; Cardwell, C.R.; Anderson, L.A.; Spence, A.D.; Lagergren, J.; Xie, S.H.; et al. Information on Genetic Variants Does Not Increase Identification of Individuals at Risk of Esophageal Adenocarcinoma Compared to Clinical Risk Factors. *Gastroenterology* **2019**, *156*, 43–45. [CrossRef] [PubMed]
24. Jin, G.; Lv, J.; Yang, M.; Wang, M.; Zhu, M.; Wang, T.; Yan, C.; Yu, C.; Ding, Y.; Li, G.; et al. Genetic risk, incident gastric cancer, and healthy lifestyle: A meta-analysis of genome-wide association studies and prospective cohort study. *Lancet Oncol.* **2020**, *21*, 1378–1386. [CrossRef] [PubMed]
25. Kim, J.; Aschard, H.; Kang, J.H.; Lentjes, M.A.H.; Do, R.; Wiggs, J.L.; Khawaja, A.P.; Pasquale, L.R. Intraocular Pressure, Glaucoma, and Dietary Caffeine Consumption: A Gene-Diet Interaction Study from the UK Biobank. *Ophthalmology* **2021**, *128*, 866–876. [CrossRef]

26. Zhang, H.; Zeng, Y.; Yang, H.; Hu, Y.; Hu, Y.; Chen, W.; Ying, Z.; Sun, Y.; Qu, Y.; Li, Q.; et al. Familial factors, diet, and risk of cardiovascular disease: A cohort analysis of the UK Biobank. *Am. J. Clin. Nutr.* **2021**, *114*, 1837–1846. [CrossRef]
27. Collins, R. What makes UK Biobank special? *Lancet* **2012**, *379*, 1173–1174. [CrossRef]
28. Palmer, L.J. UK Biobank: Bank on it. *Lancet* **2007**, *369*, 1980–1982. [CrossRef]
29. Sudlow, C.; Gallacher, J.; Allen, N.; Beral, V.; Burton, P.; Danesh, J.; Downey, P.; Elliott, P.; Green, J.; Landray, M.; et al. UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **2015**, *12*, e1001779. [CrossRef]
30. Lourida, I.; Hannon, E.; Littlejohns, T.J.; Langa, K.M.; Hyppönen, E.; Kuzma, E.; Llewellyn, D.J. Association of Lifestyle and Genetic Risk With Incidence of Dementia. *JAMA* **2019**, *322*, 430–437. [CrossRef]
31. Bycroft, C.; Freeman, C.; Petkova, D.; Band, G.; Elliott, L.T.; Sharp, K.; Motyer, A.; Vukcevic, D.; Delaneau, O.; O'Connell, J.; et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **2018**, *562*, 203–209. [CrossRef]
32. Choi, S.H.; Weng, L.C.; Roselli, C.; Lin, H.; Haggerty, C.M.; Shoemaker, M.B.; Barnard, J.; Arking, D.E.; Chasman, D.I.; Albert, C.M.; et al. Association Between Titin Loss-of-Function Variants and Early-Onset Atrial Fibrillation. *JAMA* **2018**, *320*, 2354–2364. [CrossRef]
33. Helgason, H.; Rafnar, T.; Olafsdottir, H.S.; Jonasson, J.G.; Sigurdsson, A.; Stacey, S.N.; Jonasdottir, A.; Tryggvadottir, L.; Alexiusdottir, K.; Haraldsson, A.; et al. Loss-of-function variants in ATM confer risk of gastric cancer. *Nat. Genet.* **2015**, *47*, 906–910. [CrossRef]
34. Dai, J.; Lv, J.; Zhu, M.; Wang, Y.; Qin, N.; Ma, H.; He, Y.Q.; Zhang, R.; Tan, W.; Fan, J.; et al. Identification of risk loci and a polygenic risk score for lung cancer: A large-scale prospective cohort study in Chinese populations. *Lancet Respir. Med.* **2019**, *7*, 881–891. [CrossRef]
35. Office for National Statistics. Cancer Registration Statistics, England. Available online: https://www.ons.gov.uk/peoplepopulationand-community/healthandsocialcare/conditionsanddiseases/datasets/cancerregistrationstatisticscancerregistrationstatisticsengland (accessed on 26 April 2019).
36. Zhu, M.; Wang, T.; Huang, Y.; Zhao, X.; Ding, Y.; Zhu, M.; Ji, M.; Wang, C.; Dai, J.; Yin, R.; et al. Genetic Risk for Overall Cancer and the Benefit of Adherence to a Healthy Lifestyle. *Cancer Res.* **2021**, *81*, 4618–4627. [CrossRef]
37. Liu, Y.; Yan, C.; Yin, S.; Wang, T.; Zhu, M.; Liu, L.; Jin, G. Genetic risk, metabolic syndrome, and gastrointestinal cancer risk: A prospective cohort study. *Cancer Med.* **2022**, *12*, 597–605. [CrossRef]
38. Knol, M.J.; VanderWeele, T.J. Recommendations for presenting analyses of effect modification and interaction. *Int. J. Epidemiol.* **2012**, *41*, 514–520. [CrossRef]
39. Li, R.; Chambless, L. Test for additive interaction in proportional hazards models. *Ann. Epidemiol.* **2007**, *17*, 227–236. [CrossRef]
40. Assmann, S.F.; Hosmer, D.W.; Lemeshow, S.; Mundt, K.A. Confidence intervals for measures of interaction. *Epidemiology* **1996**, *7*, 286–290. [CrossRef]
41. Arthur, R.S.; Wang, T.; Xue, X.; Kamensky, V.; Rohan, T.E. Genetic Factors, Adherence to Healthy Lifestyle Behavior, and Risk of Invasive Breast Cancer Among Women in the UK Biobank. *J. Natl. Cancer Inst.* **2020**, *112*, 893–901. [CrossRef]
42. Grosso, G.; Bella, F.; Godos, J.; Sciacca, S.; Del Rio, D.; Ray, S.; Galvano, F.; Giovannucci, E.L. Possible role of diet in cancer: Systematic review and multiple meta-analyses of dietary patterns, lifestyle factors, and cancer risk. *Nutr. Rev.* **2017**, *75*, 405–419. [CrossRef]
43. De Cicco, P.; Catani, M.V.; Gasperi, V.; Sibilano, M.; Quaglietta, M.; Savini, I. Nutrition and Breast Cancer: A Literature Review on Prevention, Treatment and Recurrence. *Nutrients* **2019**, *11*, 1514. [CrossRef] [PubMed]
44. Chen, H.; Ward, M.H.; Graubard, B.I.; Heineman, E.F.; Markin, R.M.; Potischman, N.A.; Russell, R.M.; Weisenburger, D.D.; Tucker, K.L. Dietary patterns and adenocarcinoma of the esophagus and distal stomach. *Am. J. Clin. Nutr.* **2002**, *75*, 137–144. [CrossRef] [PubMed]
45. Liu, X.; Wang, X.; Lin, S.; Yuan, J.; Yu, I.T. Dietary patterns and oesophageal squamous cell carcinoma: A systematic review and meta-analysis. *Br. J. Cancer* **2014**, *110*, 2785–2795. [CrossRef] [PubMed]
46. Cancer Research UK. Cancer Statistics for the UK. Available online: https://www.cancerresearchuk.org/health-professional/cancer-statistics-for-the-uk (accessed on 5 April 2019).
47. Ioannidis, J.P. Why most discovered true associations are inflated. *Epidemiology* **2008**, *19*, 640–648. [CrossRef]
48. Ioannidis, J.P. Why most published research findings are false. *PLoS Med.* **2005**, *2*, e124. [CrossRef]
49. Dwan, K.; Gamble, C.; Williamson, P.R.; Kirkham, J.J. Systematic review of the empirical evidence of study publication bias and outcome reporting bias—An updated review. *PLoS ONE* **2013**, *8*, e66844. [CrossRef]
50. Lin, M.; Lucas, H.C.; Shmueli, G. Research Commentary—Too Big to Fail: Large Samples and the p-Value Problem. *Inf. Syst. Res.* **2013**, *24*, 906–917. [CrossRef]
51. Willett, W.C. *Nutritional Epidemiology*; Oxford University Press: New York, NY, USA, 1998.