
eMethods. Detailed Methodology

eTable 1. Factor Loading Matrix for Dietary Patterns in the Health Professionals Follow-up Study (HPFS) and the Nurses’ Health Study (NHS)

eTable 2. Age-Standardized Characteristics of Participants in the Health Professionals Follow-up Study (HPFS, in 1994) and the Nurses’ Health Study (NHS, in 1990) According to Prudent and Western Dietary Score Quartiles (Q1 to Q4)

eTable 3. Hazard Ratios (HRs) of Incident Colorectal Cancer, Overall, According to Prudent or Western Dietary Pattern Score Quartiles in the Health Professionals Follow-up Study and the Nurses’ Health Study

eTable 4. Cohort-Specific Hazard Ratios (HRs) of Incident Colorectal Cancer Subgroups by *F. nucleatum* status According to Prudent Dietary Score Quartiles in the Health Professionals Follow-up Study (1986-2012) and the Nurses’ Health Study (1980-2012)

eTable 5. Hazard ratios (HRs) of Incident Colorectal Cancer—by Low, High, or No Detectable Levels of *F. Nucleatum* in Tumor Tissue—According to Prudent or Western Dietary Score Quartiles in the Combined Cohort of the Health Professionals Follow-up Study (1986-2012) and the Nurses’ Health Study (1980-2012)

eTable 6. Relative Risks (RRs) of Incident Colorectal Cancer Subgroups, Jointly Classified by *F. Nucleatum* Status and Anatomic Subsite, According to Prudent Dietary Pattern Scores Quartiles in the Combined Cohort of the Health Professionals Follow-up Study (1986-2012) and the Nurses’ Health Study (1980-2012)

eTable 7. Hazard Ratios (HRs) of Incident Colorectal Cancer by *F. nucleatum* Status According to The Top Four Food Items (by Factor Loadings) in the Prudent Dietary Pattern in the Combined Cohort

eTable 8. Hazard Ratios (HRs) of Incident Colorectal Cancer by *F. nucleatum* Status According to Three Major Macronutrients in the Combined Cohort

eTable 9. Hazard Ratios (HRs) of Incident Colorectal Cancer Subgroups by *F. nucleatum* Status According to Intakes of Three Major Subclasses of Dietary Fiber in the Combined Cohort

This supplementary material has been provided by the authors to give readers additional information about their work.
eMethods. Detailed Methodology

Prior to analysis, a prospective, written protocol was submitted and approved by the research groups that oversaw the Nurses’ Health Study and the Health Professionals Follow-up Study.

At baseline in the NHS, we excluded 1,001 participants with a prior history of cancer, 1,044 participants with a history of ulcerative colitis, and 430 participants with recorded intakes less than 600 calories per day. At baseline in the HPFS, we excluded 1,997 participants with a prior history of cancer and 475 participants with a prior history of ulcerative colitis.

Normal and tumor sections from all colorectal carcinoma cases in this study were reviewed by a pathologist (S.O.). For various characteristics, colorectal cancer patients with available tumor tissue data (n = 1,019) were generally similar to patients without available tumor tissue data (n = 2,241) (median age 66.8 vs 67.1; current smoker, 15% vs. 15%; mean body-mass index, 26.3 vs. 26.2 kg/m²; previous lower gastrointestinal endoscopy, 34% vs. 36%; mean red meat intake [servings per day], 0.55 vs. 0.54; mean dietary fiber intake, 20.0 vs 19.7 g per day; mean alcohol intake, 8.5 vs. 8.2 g per day). Tumors with and without available tissue were also similar by clinical factors, including grade and stage [proportion of poorly differentiated tumors, 16% vs. 17%; mean pT stage, 2.7 vs. 2.5; proportion with no regional lymph node metastases, 59% vs. 49%; proportion of distant metastasis (M stage), 11% vs. 17%].

Food items from the FFQ were classified into approximately 40 food groups. Factor analysis was performed using an orthogonal rotation procedure to produce two maximally uncorrelated factors, selected based upon the largest eigenvalues. All analyses were adjusted for total caloric intake (kcal per day) and stratified by age (in months), year of questionnaire return and sex (in the analysis using combined cohorts). In multivariable analysis, we adjusted for potential confounders including body mass index (kg/m²), pack-years of smoking (never, 0 to 4 pack-years, 5 to 19 pack-years, 20 to 39 pack-years,
or >40 years), family history of colorectal cancer in any first-degree relative (yes or no), previous lower gastrointestinal endoscopy (yes or no), postmenopausal hormone use (for women only; never, past, or current), physical activity [quintiles of metabolic-equivalent task (MET)-hours per week], and regular aspirin or NSAID use (≥2 tablets per week; yes or no). Prior to pooling data from the two cohorts, we examined the possible heterogeneity between cohorts, using the Q statistic for the association between the prudent (or Western) dietary score and overall incidence of colorectal cancer.

Participants whose *F. nucleatum* tumor status was unknown and those who died of causes other than colorectal cancer were censored during the 26-32 years of follow-up. In addition, in the incidence analysis of one subtype, incidences of other tumor subtypes were treated as censored data. We assessed the proportional hazards assumption by including the product term between age and each covariable (including the exposure of interest, prudent dietary pattern scores) into the Cox model, and testing the statistical significance of the term by Wald test. No deviation from proportional hazards assumption was detected at the α of 0.05 level. We did not conduct a formal power calculation prior to the current analysis. We recognized that participants might have varied their diets over the study period. Thus, we utilized time-varying covariates such that each individual participant contributed multiple person-times with differing dietary data provided on each questionnaire.

cTable 1. Factor Loading Matrix for Dietary Patterns in the Health Professionals Follow-up Study (HPFS) and the Nurses’ Health Study (NHS)

<table>
<thead>
<tr>
<th>Food item</th>
<th>HPFS</th>
<th>NHS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prudent pattern</td>
<td>Western pattern</td>
</tr>
<tr>
<td>Other vegetables</td>
<td>0.69</td>
<td>0.65</td>
</tr>
<tr>
<td>Yellow vegetables</td>
<td>0.65</td>
<td>0.65</td>
</tr>
<tr>
<td>Leafy green vegetables</td>
<td>0.63</td>
<td>0.64</td>
</tr>
<tr>
<td>Cruciferous vegetables</td>
<td>0.63</td>
<td>0.61</td>
</tr>
<tr>
<td>Fruits</td>
<td>0.58</td>
<td>0.58</td>
</tr>
<tr>
<td>Legumes</td>
<td>0.62</td>
<td>0.57</td>
</tr>
<tr>
<td>Tomatoes</td>
<td>0.52</td>
<td>0.48</td>
</tr>
<tr>
<td>Fish</td>
<td>0.46</td>
<td>0.51</td>
</tr>
<tr>
<td>Item</td>
<td>Factor 1</td>
<td>Factor 2</td>
</tr>
<tr>
<td>-----------------------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>Whole grains</td>
<td>0.38</td>
<td>0.41</td>
</tr>
<tr>
<td>Potatoes</td>
<td>0.33</td>
<td>0.34</td>
</tr>
<tr>
<td>Poultry</td>
<td>0.32</td>
<td>0.41</td>
</tr>
<tr>
<td>Water</td>
<td></td>
<td>0.32</td>
</tr>
<tr>
<td>Salad dressing</td>
<td></td>
<td>0.33</td>
</tr>
<tr>
<td>Low-fat dairy</td>
<td></td>
<td>0.32</td>
</tr>
<tr>
<td>Red meat</td>
<td>0.66</td>
<td>0.61</td>
</tr>
<tr>
<td>Processed meat</td>
<td>0.61</td>
<td>0.58</td>
</tr>
<tr>
<td>High fat dairy</td>
<td>0.51</td>
<td>0.50</td>
</tr>
<tr>
<td>French fries</td>
<td>0.49</td>
<td>0.46</td>
</tr>
<tr>
<td>Eggs</td>
<td>0.47</td>
<td>0.41</td>
</tr>
<tr>
<td>Desserts²</td>
<td>0.43</td>
<td>0.45</td>
</tr>
<tr>
<td>Condiments⁴</td>
<td>0.39</td>
<td>0.36</td>
</tr>
<tr>
<td>Refined grains</td>
<td>0.38</td>
<td>0.38</td>
</tr>
<tr>
<td>Butter</td>
<td>0.38</td>
<td>0.50</td>
</tr>
<tr>
<td>Mayonnaise</td>
<td>0.36</td>
<td>0.34</td>
</tr>
<tr>
<td>Margarine</td>
<td>0.34</td>
<td>0.32</td>
</tr>
<tr>
<td>Snacks⁵</td>
<td>0.34</td>
<td></td>
</tr>
<tr>
<td>Pizza</td>
<td>0.33</td>
<td>0.36</td>
</tr>
<tr>
<td>Creamy soups</td>
<td>0.31</td>
<td>0.32</td>
</tr>
<tr>
<td>Sugar-sweetened beverages</td>
<td>0.31</td>
<td>0.33</td>
</tr>
</tbody>
</table>

¹ Only items with correlation coefficients >0.30 are presented. With the orthogonal rotation used, correlations are identical to factor loading matrix.
² Other vegetables include corn, onion, eggplant, celery, green peppers, and mixed vegetables.
³ Desserts include chocolate, candy bars, cookies, brownies, cake, pie, and pastries.
⁴ Condiments include soy sauce, non-dairy creamer, Worcestershire sauce, red chili sauce, and pepper.
⁵ Snacks include chips, popcorn, and crackers.
eTable 2. Age-Standardized Characteristics of Participants in the Health Professionals Follow-up Study (HPFS, in 1994) and the Nurses’ Health Study (NHS, in 1990) According to Prudent and Western Dietary Score Quartiles (Q1 to Q4)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Health Professionals Follow-up Study</th>
<th>Nurses’ Health Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prudent Q1</td>
<td>Prudent Q4</td>
</tr>
<tr>
<td>Age (mean)</td>
<td>61.1</td>
<td>61.3</td>
</tr>
<tr>
<td>Pack-years smoked (mean)</td>
<td>16.1</td>
<td>11.0</td>
</tr>
<tr>
<td>Current smoker, %</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>Family history of colorectal cancer, %</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>History of previous endoscopy, %</td>
<td>17</td>
<td>21</td>
</tr>
<tr>
<td>Current multivitamin use, %</td>
<td>39</td>
<td>47</td>
</tr>
<tr>
<td>Regular aspirin/NSAID use, %</td>
<td>42</td>
<td>45</td>
</tr>
<tr>
<td>Postmenopause, %</td>
<td>78</td>
<td>79</td>
</tr>
<tr>
<td>Menopausal hormone use, %</td>
<td>32</td>
<td>36</td>
</tr>
<tr>
<td>Body mass index, kg/m² (mean)</td>
<td>26.1</td>
<td>25.8</td>
</tr>
<tr>
<td>Physical activity, MET-hours/week (mean)</td>
<td>26.4</td>
<td>44.8</td>
</tr>
<tr>
<td>Unprocessed red meat, svg/day</td>
<td>0.59</td>
<td>0.52</td>
</tr>
<tr>
<td>Processed red meat, svg/day</td>
<td>0.33</td>
<td>0.20</td>
</tr>
<tr>
<td>Poultry, svg/day</td>
<td>0.30</td>
<td>0.52</td>
</tr>
<tr>
<td>Fruit, svg/day</td>
<td>1.64</td>
<td>3.86</td>
</tr>
<tr>
<td>Vegetable, svg/day</td>
<td>1.84</td>
<td>5.54</td>
</tr>
<tr>
<td>Alcohol, g/day</td>
<td>11.2</td>
<td>10.6</td>
</tr>
</tbody>
</table>

© 2017 American Medical Association. All rights reserved.
<table>
<thead>
<tr>
<th></th>
<th>450</th>
<th>617</th>
<th>627</th>
<th>449</th>
<th>376</th>
<th>490</th>
<th>499</th>
<th>367</th>
</tr>
</thead>
<tbody>
<tr>
<td>Folate, μg/day</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calcium, mg/day</td>
<td>864</td>
<td>975</td>
<td>1024</td>
<td>831</td>
<td>900</td>
<td>1062</td>
<td>1151</td>
<td>835</td>
</tr>
<tr>
<td>Vitamin D, IU/day</td>
<td>408</td>
<td>498</td>
<td>541</td>
<td>378</td>
<td>309</td>
<td>376</td>
<td>420</td>
<td>271</td>
</tr>
<tr>
<td>Dietary fiber, g/day</td>
<td>18</td>
<td>29</td>
<td>28</td>
<td>19</td>
<td>15</td>
<td>22</td>
<td>21</td>
<td>16</td>
</tr>
</tbody>
</table>

1Continuous variables are described as means.
2 Regular users are defined as ≥2 standard (325-mg) tablets of aspirin or ≥ 2 tablets of non-steroidal anti-inflammatory drugs (NSAIDs) per week.
3 The percentages of postmenopausal participants as well as menopausal hormone use are among women only.
4 Physical activity is represented by the product sum of the metabolic equivalent task (MET) score of each specific recreational activity and hours spent on that activity per week. MET-hours/week values were assessed in 1988 for the NHS.
### eTable 3. Hazard Ratios (HRs) of Incident Colorectal Cancer, Overall, According to Prudent or Western Dietary Pattern Score Quartiles in the Health Professionals Follow-up Study and the Nurses’ Health Study

<table>
<thead>
<tr>
<th></th>
<th>Quartile 1</th>
<th>Quartile 2</th>
<th>Quartile 3</th>
<th>Quartile 4</th>
<th>( P_{\text{trend}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prudent dietary pattern</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Health Professionals Follow-up Study</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Person-years</td>
<td>260,140</td>
<td>258,863</td>
<td>260,286</td>
<td>258,542</td>
<td></td>
</tr>
<tr>
<td>No. of cases (n=388)</td>
<td>85</td>
<td>105</td>
<td>112</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td>Age-adjusted HR (95% CI)(^1)</td>
<td>1 (referent)</td>
<td>1.08 (0.81-1.44)</td>
<td>1.03 (0.77-1.398)</td>
<td>0.74 (0.53-1.02)</td>
<td>0.04</td>
</tr>
<tr>
<td>Multivariable HR (95% CI)(^2)</td>
<td>1 (referent)</td>
<td>1.11 (0.83-1.49)</td>
<td>1.07 (0.80-1.45)</td>
<td>0.80 (0.56-1.11)</td>
<td>0.10</td>
</tr>
<tr>
<td><strong>Nurses’ Health Study</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Person-years</td>
<td>653,429</td>
<td>648,813</td>
<td>652,109</td>
<td>651,380</td>
<td></td>
</tr>
<tr>
<td>No. of cases (n=631)</td>
<td>165</td>
<td>143</td>
<td>156</td>
<td>167</td>
<td></td>
</tr>
<tr>
<td>Age-adjusted HR (95% CI)(^1)</td>
<td>1 (referent)</td>
<td>0.84 (0.67-1.06)</td>
<td>0.84 (0.67-1.05)</td>
<td>0.83 (0.66-1.05)</td>
<td>0.17</td>
</tr>
<tr>
<td>Multivariable HR (95% CI)(^2)</td>
<td>1 (referent)</td>
<td>0.88 (0.70-1.10)</td>
<td>0.89 (0.71-1.12)</td>
<td>0.91 (0.71-1.16)</td>
<td>0.52</td>
</tr>
<tr>
<td><strong>Western dietary pattern</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Health Professionals Follow-up Study</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Person-years</td>
<td>259,064</td>
<td>259,719</td>
<td>259,567</td>
<td>259,480</td>
<td></td>
</tr>
<tr>
<td>No. of cases (n=388)</td>
<td>83</td>
<td>104</td>
<td>96</td>
<td>105</td>
<td></td>
</tr>
<tr>
<td>Age-adjusted HR (95% CI)(^1)</td>
<td>1 (referent)</td>
<td>1.42 (1.06-1.91)</td>
<td>1.45 (1.05-1.98)</td>
<td>1.80 (1.25-2.59)</td>
<td>0.003</td>
</tr>
<tr>
<td>Multivariable HR (95% CI)(^2)</td>
<td>1 (referent)</td>
<td>1.37 (1.01-1.84)</td>
<td>1.35 (0.98-1.86)</td>
<td>1.62 (1.11-2.37)</td>
<td>0.02</td>
</tr>
<tr>
<td><strong>Nurses’ Health Study</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Person-years</td>
<td>651,592</td>
<td>650,806</td>
<td>650,898</td>
<td>652,436</td>
<td></td>
</tr>
<tr>
<td>No. of cases (n=631)</td>
<td>161</td>
<td>171</td>
<td>147</td>
<td>152</td>
<td></td>
</tr>
<tr>
<td>Age-adjusted HR (95% CI)(^1)</td>
<td>1 (referent)</td>
<td>1.18 (0.95-1.47)</td>
<td>1.09 (0.86-1.40)</td>
<td>1.31 (1.00-1.73)</td>
<td>0.09</td>
</tr>
<tr>
<td>Multivariable HR (95% CI)(^2)</td>
<td>1 (referent)</td>
<td>1.13 (0.91-1.41)</td>
<td>1.01 (0.79-1.29)</td>
<td>1.13 (0.85-1.51)</td>
<td>0.55</td>
</tr>
</tbody>
</table>

\(^1\) Stratified by age and calendar year and adjusted for total caloric intake

\(^2\) As above, and additionally adjusted for family history of colorectal cancer in any first-degree relative, history of previous endoscopy, pack-years of smoking (never, 0-4, 5-19, 20-39, or >40), body mass index (kg/m2), physical activity (MET-hours/week), regular aspirin or NSAID use (≥2 tablets/week), menopausal hormone therapy status (never, past current) (women only), and total caloric intake (kcal/day).
Tests for trend were conducted using the median value of each category as a continuous variable.

Abbreviations: CI, confidence interval; HR, hazard ratio; MET, metabolic equivalent task; NSAID, non-steroidal anti-inflammatory drug.
Table 4. Cohort-Specific Hazard Ratios (HRs) of Incident Colorectal Cancer Subgroups by *F. nucleatum* status According to Prudent Dietary Score Quartiles in the Health Professionals Follow-up Study (1986-2012) and the Nurses’ Health Study (1980-2012)

<table>
<thead>
<tr>
<th>Prudent dietary score</th>
<th>Quartile 1</th>
<th>Quartile 2</th>
<th>Quartile 3</th>
<th>Quartile 4</th>
<th>Pr(\text{trend})</th>
<th>\text{Pr(\text{heterogeneity})}</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Health Professionals Follow-up Study</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum</em> (+) colorectal cancer (n=39)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)$^1$</td>
<td>1 (referent)</td>
<td>0.82 (0.32-1.73)</td>
<td>0.74 (0.32-1.73)</td>
<td>0.31 (0.11-0.90)</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum</em> (-) colorectal cancer (n=349)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)$^1$</td>
<td>1 (referent)</td>
<td>1.16 (0.85-1.59)</td>
<td>1.13 (0.82-1.55)</td>
<td>0.87 (0.61-1.24)</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td><strong>Nurses’ Health Study</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum</em> (+) colorectal cancer (n=86)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)$^1$</td>
<td>1 (referent)</td>
<td>0.47 (0.25-0.87)</td>
<td>0.68 (0.4-1.18)</td>
<td>0.49 (0.27-0.89)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum</em> (-) colorectal cancer (n=545)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)$^1$</td>
<td>1 (referent)</td>
<td>0.97 (0.76-1.25)</td>
<td>0.94 (0.73-1.2)</td>
<td>1.01 (0.78-1.3)</td>
<td>0.97</td>
<td></td>
</tr>
</tbody>
</table>

1 Stratified by age and calendar year and adjusted for total caloric intake, family history of colorectal cancer in any first-degree relative, history of previous endoscopy, pack-years of smoking (never, 0-4, 5-19, 20-39, or >40), body mass index (kg/m²), physical activity (MET-hours/week), regular aspirin or NSAID use (≥2 tablets/week), menopausal hormone therapy status (never, past current) (women only), and total caloric intake (kcal/day).

2 Tests for trend were conducted using the median value of each category as a continuous variable.

3 We tested for heterogeneity by using a likelihood ratio test, comparing a model that allows separate associations for the two colorectal cancer subgroups (i.e., *F. nucleatum*-positive and negative subgroups) with a model that assumes a common association.

Abbreviations: CI, confidence interval; HR, hazard ratio; MET, metabolic equivalent task; NSAID, non-steroidal anti-inflammatory drug.
**Table 5.** Hazard ratios (HRs) of Incident Colorectal Cancer—by Low, High, or No Detectable Levels of *F. Nucleatum* in Tumor Tissue—According to Prudent or Western Dietary Score Quartiles in the Combined Cohort of the Health Professionals Follow-up Study (1986-2012) and the Nurses’ Health Study (1980-2012)

<table>
<thead>
<tr>
<th>Quotient</th>
<th>Quartile 1</th>
<th>Quartile 2</th>
<th>Quartile 3</th>
<th>Quartile 4</th>
<th>(P_{\text{trend}})</th>
<th>(P_{\text{heterogeneity}})</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prudent dietary pattern</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum</em>-high colorectal cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases (n=60)</td>
<td>20</td>
<td>10</td>
<td>20</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)</td>
<td>1 (referent)</td>
<td>0.47 (0.22-1.00)</td>
<td>0.92 (0.49-1.73)</td>
<td>0.43 (0.20-0.92)</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum</em>-low colorectal cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases (n=65)</td>
<td>23</td>
<td>16</td>
<td>14</td>
<td>12</td>
<td></td>
<td>0.03</td>
</tr>
<tr>
<td>Multivariable HR (95% CI)</td>
<td>1 (referent)</td>
<td>0.65 (0.34-1.23)</td>
<td>0.52 (0.26-1.01)</td>
<td>0.43 (0.22-0.87)</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum</em>-negative colorectal cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases (n=894)</td>
<td>207</td>
<td>222</td>
<td>234</td>
<td>231</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)</td>
<td>1 (referent)</td>
<td>1.04 (0.86-1.26)</td>
<td>1.00 (0.83-1.22)</td>
<td>0.95 (0.77-1.17)</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td><strong>Western dietary pattern</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum</em>-high colorectal cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases (n=60)</td>
<td>13</td>
<td>18</td>
<td>17</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)</td>
<td>1 (referent)</td>
<td>1.44 (0.70-2.95)</td>
<td>1.49 (0.70-3.09)</td>
<td>1.16 (0.52-2.60)</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum</em>-low colorectal cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases (n=65)</td>
<td>12</td>
<td>15</td>
<td>16</td>
<td>22</td>
<td></td>
<td>0.19</td>
</tr>
<tr>
<td>Multivariable HR (95% CI)</td>
<td>1 (referent)</td>
<td>1.29 (0.60-2.78)</td>
<td>1.50 (0.70-3.20)</td>
<td>2.25 (1.08-4.66)</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum</em>-negative colorectal cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases (n=894)</td>
<td>219</td>
<td>242</td>
<td>210</td>
<td>223</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)</td>
<td>1 (referent)</td>
<td>1.20 (0.99-1.44)</td>
<td>1.08 (0.88-1.33)</td>
<td>1.25 (0.99-1.58)</td>
<td>0.12</td>
<td></td>
</tr>
</tbody>
</table>

1 Stratified by age, calendar year, and gender and adjusted for total caloric intake (kcal/day), family history of colorectal cancer in any first-degree relative, history of previous endoscopy, pack-years of smoking (never, 0-4, 5-19, 20-39, or ≥40), body mass index (kg/m²), physical activity (MET-hours/week), and regular aspirin or NSAID use (≥2 tablets/week).

2 Tests for trend were conducted using the median value of each quartile category as a continuous variable.

3 We tested for heterogeneity by using a likelihood ratio test, comparing a model that allows separate associations for the two colorectal cancer subgroups (i.e., *F. nucleatum*-positive and negative subgroups) with a model that assumes a common association.

Abbreviations: CI, confidence interval; HR, hazard ratio; MET, metabolic equivalent task; NSAID, non-steroidal anti-inflammatory drug.
**Table 6.** Relative Risks (RRs) of Incident Colorectal Cancer Subgroups, Jointly Classified by *F. Nucleatum* Status and Anatomic Subsite, According to Prudent Dietary Pattern Scores Quartiles in the Combined Cohort of the Health Professionals Follow-up Study (1986-2012) and the Nurses’ Health Study (1980-2012)

<table>
<thead>
<tr>
<th>Anatomic subsite¹</th>
<th><em>F. nucleatum</em> status</th>
<th>Quartile 1</th>
<th>Quartile 2</th>
<th>Quartile 3</th>
<th>Quartile 4</th>
<th>P trend³</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proximal colon cancer (n=496)</td>
<td><em>F. nucleatum</em> (+) colorectal cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases (n=79)</td>
<td>28</td>
<td>12</td>
<td>23</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)²</td>
<td>1 (referent)</td>
<td>0.38 (0.19-0.76)</td>
<td>0.70 (0.40-1.23)</td>
<td>0.46 (0.24-0.85)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum</em> (-) colorectal cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases (n=417)</td>
<td>89</td>
<td>98</td>
<td>110</td>
<td>120</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)²</td>
<td>1 (referent)</td>
<td>1.05 (0.78-1.40)</td>
<td>1.06 (0.80-1.41)</td>
<td>1.10 (0.82-1.47)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distal colon and rectal cancer (n=515)</td>
<td><em>F. nucleatum</em> (+) colorectal cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases (n=44)</td>
<td>14</td>
<td>14</td>
<td>10</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)²</td>
<td>1 (referent)</td>
<td>0.98 (0.46-2.06)</td>
<td>0.65 (0.28-1.48)</td>
<td>0.38 (0.14-1.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum</em> (-) colorectal cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases (n=471)</td>
<td>117</td>
<td>123</td>
<td>123</td>
<td>108</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)²</td>
<td>1 (referent)</td>
<td>1.03 (0.80-1.33)</td>
<td>0.96 (0.74-1.24)</td>
<td>0.81 (0.61-1.07)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

¹ Tumors were classified as proximal if they were removed from the cecum to the transverse colon, distal if they were removed from the splenic flexure to the sigmoid colon, and rectal if they were removed from the rectosigmoid junction to the anal canal (excluding anal squamous cell carcinoma).

² Stratified by age, calendar year, and gender and adjusted for family history of colorectal cancer in any first-degree relative, history of previous endoscopy, pack-years of smoking (never, 0-4, 5-19, 20-39, or ≥40), body mass index (kg/m²), physical activity (MET-hours/week), regular aspirin or NSAID use (≥2 tablets/week), and total caloric intake (kcal/day).

³ Tests for trend were conducted using the median value of each category as a continuous variable.

Abbreviations: CI, confidence interval; HR, hazard ratio; MET, metabolic equivalent task; NSAID, non-steroidal anti-inflammatory drug.
**Table 7.** Hazard Ratios (HRs) of Incident Colorectal Cancer by *F. nucleatum* Status According to The Top Four Food Items (by Factor Loadings) in the Prudent Dietary Pattern in the Combined Cohort

<table>
<thead>
<tr>
<th>Food Item</th>
<th>Quartile 1</th>
<th>Quartile 2</th>
<th>Quartile 3</th>
<th>Quartile 4</th>
<th>( P_{\text{trend}} )</th>
<th>( P_{\text{heterogeneity}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vegetables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum (+)</em> colorectal cancer (n=125)</td>
<td>29</td>
<td>34</td>
<td>31</td>
<td>31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)(^\dagger)</td>
<td>1 (referent)</td>
<td>0.95 (0.58-1.57)</td>
<td>0.80 (0.47-1.33)</td>
<td>0.70 (0.42-1.18)</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum (-)</em> colorectal cancer (n=894)</td>
<td>156</td>
<td>216</td>
<td>261</td>
<td>261</td>
<td></td>
<td>0.03</td>
</tr>
<tr>
<td>Multivariable HR (95% CI)(^\dagger)</td>
<td>1 (referent)</td>
<td>1.20 (0.97-1.48)</td>
<td>1.33 (1.08-1.64)</td>
<td>1.25 (1.01-1.56)</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td><strong>Fruits</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum (+)</em> colorectal cancer (n=125)</td>
<td>31</td>
<td>31</td>
<td>37</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)(^\dagger)</td>
<td>1 (referent)</td>
<td>0.77 (0.47-1.28)</td>
<td>0.84 (0.52-1.37)</td>
<td>0.55 (0.32-0.95)</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum (-)</em> colorectal cancer (n=894)</td>
<td>192</td>
<td>234</td>
<td>231</td>
<td>237</td>
<td></td>
<td>0.22</td>
</tr>
<tr>
<td>Multivariable HR (95% CI)(^\dagger)</td>
<td>1 (referent)</td>
<td>1.03 (0.85-1.25)</td>
<td>0.91 (0.74-1.11)</td>
<td>0.85 (0.68-1.05)</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td><strong>Legumes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum (+)</em> colorectal cancer (n=125)</td>
<td>23</td>
<td>40</td>
<td>38</td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)(^\dagger)</td>
<td>1 (referent)</td>
<td>1.15 (0.68-1.96)</td>
<td>0.94 (0.55-1.60)</td>
<td>0.98 (0.54-1.79)</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum (-)</em> colorectal cancer (n=894)</td>
<td>151</td>
<td>265</td>
<td>283</td>
<td>195</td>
<td></td>
<td>0.60</td>
</tr>
<tr>
<td>Multivariable HR (95% CI)(^\dagger)</td>
<td>1 (referent)</td>
<td>1.36 (1.11-1.67)</td>
<td>1.26 (1.02-1.55)</td>
<td>1.16 (0.92-1.47)</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td><strong>Whole grains</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum (+)</em> colorectal cancer (n=125)</td>
<td>26</td>
<td>42</td>
<td>37</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)(^\dagger)</td>
<td>1 (referent)</td>
<td>1.07 (0.60-1.76)</td>
<td>0.95 (0.57-1.58)</td>
<td>0.56 (0.31-1.03)</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum (-)</em> colorectal cancer (n=894)</td>
<td>151</td>
<td>264</td>
<td>247</td>
<td>232</td>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td>Multivariable HR (95% CI)(^\dagger)</td>
<td>1 (referent)</td>
<td>1.31 (1.06-1.61)</td>
<td>1.23 (0.99-1.52)</td>
<td>1.17 (0.94-1.45)</td>
<td>0.56</td>
<td></td>
</tr>
</tbody>
</table>

\(^\dagger\) Stratified by age, calendar year, and gender and adjusted for family history of colorectal cancer in any first-degree relative, history of previous endoscopy, pack-years of smoking (never, 0-4, 5-19, 20-39, or >40), body mass index (kg/m²), physical activity (MET-hours/week), regular aspirin or NSAID use (≥2 tablets/week), and total caloric intake (kcal/day).

\(^2\) Tests for trend were conducted using the median value of each quartile category as a continuous variable.

\(^3\) We tested for heterogeneity by using a likelihood ratio test, comparing a model that allows separate associations for the two colorectal cancer subgroups (i.e., *F. nucleatum*-positive and negative subgroups) with a model that assumes a common association.

Abbreviations: CI, confidence interval; HR, hazard ratio; MET, metabolic equivalent task; NSAID, non-steroidal anti-inflammatory drug.
Table 8. Hazard Ratios (HRs) of Incident Colorectal Cancer by *F. nucleatum* Status According to Three Major Macronutrients in the Combined Cohort

<table>
<thead>
<tr>
<th></th>
<th>Quartile 1</th>
<th>Quartile 2</th>
<th>Quartile 3</th>
<th>Quartile 4</th>
<th><em>P</em>&lt;sub&gt;trend&lt;/sub&gt;</th>
<th><em>P</em>&lt;sub&gt;heterogeneity&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fiber</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum</em> (+) colorectal cancer (n=125)</td>
<td>30</td>
<td>27</td>
<td>41</td>
<td>27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)&lt;sup&gt;1&lt;/sup&gt;</td>
<td>1 (referent)</td>
<td>0.68 (0.40-1.16)</td>
<td>0.87 (0.53-1.41)</td>
<td>0.54 (0.32-0.92)</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum</em> (-) colorectal cancer (n=894)</td>
<td>164</td>
<td>230</td>
<td>236</td>
<td>264</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)&lt;sup&gt;1&lt;/sup&gt;</td>
<td>1 (referent)</td>
<td>1.14 (0.93-1.40)</td>
<td>1.08 (0.87-1.32)</td>
<td>1.13 (0.92-1.40)</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td><strong>Fat</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum</em> (+) colorectal cancer (n=125)</td>
<td>28</td>
<td>28</td>
<td>31</td>
<td>28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)&lt;sup&gt;1&lt;/sup&gt;</td>
<td>1 (referent)</td>
<td>1.41 (0.86-2.30)</td>
<td>1.27 (0.75-2.12)</td>
<td>1.63 (0.96-2.78)</td>
<td>0.10</td>
<td>0.29</td>
</tr>
<tr>
<td><em>F. nucleatum</em> (-) colorectal cancer (n=894)</td>
<td>237</td>
<td>234</td>
<td>230</td>
<td>193</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)&lt;sup&gt;1&lt;/sup&gt;</td>
<td>1 (referent)</td>
<td>1.05 (0.87-1.26)</td>
<td>1.12 (0.92-1.35)</td>
<td>1.13 (0.92-1.37)</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td><strong>Protein</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum</em> (+) colorectal cancer (n=125)</td>
<td>32</td>
<td>36</td>
<td>37</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)&lt;sup&gt;1&lt;/sup&gt;</td>
<td>1 (referent)</td>
<td>1.01 (0.62-1.63)</td>
<td>1.08 (0.67-1.73)</td>
<td>0.68 (0.39-1.18)</td>
<td>0.27</td>
<td>0.72</td>
</tr>
<tr>
<td><em>F. nucleatum</em> (-) colorectal cancer (n=894)</td>
<td>225</td>
<td>243</td>
<td>246</td>
<td>180</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)&lt;sup&gt;1&lt;/sup&gt;</td>
<td>1 (referent)</td>
<td>1.06 (0.89-1.28)</td>
<td>1.08 (0.90-1.30)</td>
<td>0.81 (0.66-0.99)</td>
<td>0.05</td>
<td></td>
</tr>
</tbody>
</table>

1 Stratified by age, calendar year, and gender and adjusted for family history of colorectal cancer in any first-degree relative, history of previous endoscopy, pack-years of smoking (never, 0-4, 5-19, 20-39, or >40), body mass index (kg/m<sup>2</sup>), physical activity (MET-hours/week), regular aspirin or NSAID use (≥2 tablets/week), and total caloric intake (kcal/day).

2 Tests for trend were conducted using the median value of each quartile category as a continuous variable.

3 We tested for heterogeneity by using a likelihood ratio test, comparing a model that allows separate associations for the two colorectal cancer subgroups (i.e., *F. nucleatum*-positive and negative subgroups) with a model that assumes a common association.

Abbreviations: CI, confidence interval; HR, hazard ratio; MET, metabolic equivalent task; NSAID, non-steroidal anti-inflammatory drug.
Table 9. Hazard Ratios (HRs) of Incident Colorectal Cancer Subgroups by *F. nucleatum* Status According to Intakes of Three Major Subclasses of Dietary Fiber in the Combined Cohort

<table>
<thead>
<tr>
<th></th>
<th>Quartile 1</th>
<th>Quartile 2</th>
<th>Quartile 3</th>
<th>Quartile 4</th>
<th>$P_{\text{trend}}$</th>
<th>$P_{\text{heterogeneity}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cereal-derived fiber</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum (+)</em> colorectal cancer (n=125)</td>
<td>29</td>
<td>31</td>
<td>34</td>
<td>31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)$^1$</td>
<td>1 (referent)</td>
<td>0.75 (0.45-1.26)</td>
<td>0.69 (0.41-1.15)</td>
<td>0.58 (0.34-0.99)</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum (-)</em> colorectal cancer (n=894)</td>
<td>166</td>
<td>223</td>
<td>242</td>
<td>263</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)$^1$</td>
<td>1 (referent)</td>
<td>1.10 (0.89-1.35)</td>
<td>1.09 (0.88-1.34)</td>
<td>1.13 (0.92-1.40)</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td><strong>Vegetable-derived fiber</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum (+)</em> colorectal cancer (n=125)</td>
<td>33</td>
<td>30</td>
<td>29</td>
<td>33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)$^1$</td>
<td>1 (referent)</td>
<td>0.74 (0.45-1.22)</td>
<td>0.67 (0.41-1.12)</td>
<td>0.71 (0.44-1.16)</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum (-)</em> colorectal cancer (n=894)</td>
<td>166</td>
<td>230</td>
<td>250</td>
<td>248</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)$^1$</td>
<td>1 (referent)</td>
<td>1.25 (1.02-1.53)</td>
<td>1.29 (1.05-1.57)</td>
<td>1.24 (1.01-1.52)</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td><strong>Fruit-derived fiber</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum (+)</em> colorectal cancer (n=125)</td>
<td>29</td>
<td>32</td>
<td>36</td>
<td>28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)$^1$</td>
<td>1 (referent)</td>
<td>0.88 (0.53-1.47)</td>
<td>0.90 (0.55-1.48)</td>
<td>0.71 (0.41-1.20)</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum (-)</em> colorectal cancer (n=894)</td>
<td>195</td>
<td>228</td>
<td>243</td>
<td>228</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>Multivariable HR (95% CI)$^1$</td>
<td>1 (referent)</td>
<td>0.99 (0.81-1.20)</td>
<td>0.98 (0.80-1.19)</td>
<td>0.85 (0.70-1.04)</td>
<td>0.08</td>
<td></td>
</tr>
</tbody>
</table>

1 Stratified by age, calendar year, and gender and adjusted for family history of colorectal cancer in any first-degree relative, history of previous endoscopy, pack-years of smoking (never, 0-4, 5-19, 20-39, or ≥40), body mass index (kg/m²), physical activity (MET-hours/week), regular aspirin or NSAID use (≥2 tablets/week), and total caloric intake (kcal/day).

2 Tests for trend were conducted using the median value of each category as a continuous variable.

3 We tested for heterogeneity by using a likelihood ratio test, comparing a model that allows separate associations for the two colorectal cancer subgroups (i.e., *F. nucleatum*-positive and negative subgroups) with a model that assumes a common association.

Abbreviations: CI, confidence interval; HR, hazard ratio; MET, metabolic equivalent task; NSAID, non-steroidal anti-inflammatory drug.