Association of Dietary Patterns With Risk of Colorectal Cancer Subtypes Classified by *Fusobacterium nucleatum* in Tumor Tissue

Raaj S. Mehta, MD; Reiko Nishihara, PhD; Yin Cao, ScD; Mingyang Song, MD, ScD; Kosuke Mima, MD, PhD; Zhi Rong Qian, MD, PhD; Jonathan A. Nowak, MD, PhD; Keisuke Kosumi, MD, PhD; Tsuyoshi Hamada, MD, PhD; Yohei Masugi, MD, PhD; Susan Bullman, PhD; David A. Drew, PhD; Aleksandar D. Kostic, PhD; Teresa T. Fung, ScD, RDN; Wendy S. Garrett, MD, PhD; Curtis Huttenhower, PhD; Kana Wu, MD, PhD; Jeffrey A. Meyerhardt, MD, MPH; Xuehong Zhang, MD, ScD; Walter C. Willett, MD, DrPH; Edward L. Giovannucci, MD, ScD; Charles S. Fuchs, MD, MPH; Andrew T. Chan, MD, MPH; Shuji Ogino, MD, PhD

**IMPORTANCE** *Fusobacterium nucleatum* appears to play a role in colorectal carcinogenesis through suppression of the host's immune response to tumor. Evidence also suggests that diet influences intestinal *F* nucleatum. However, the role of *F* nucleatum in mediating the relationship between diet and the risk of colorectal cancer is unknown.

**OBJECTIVE** To test the hypothesis that the associations of prudent diets (rich in whole grains and dietary fiber) and Western diets (rich in red and processed meat, refined grains, and desserts) with colorectal cancer risk may differ according to the presence of *F* nucleatum in tumor tissue.

**DESIGN, SETTING, AND PARTICIPANTS** A prospective cohort study was conducted using data from the Nurses' Health Study (June 1, 1980, to June 1, 2012) and the Health Professionals Follow-up Study (June 1, 1986, to June 1, 2012) on a total of 121,700 US female nurses and 51,529 US male health professionals aged 30 to 55 years and 40 to 75 years, respectively (both predominantly white individuals), at enrollment. Data analysis was performed from March 15, 2015, to August 10, 2016.

**EXPOSURES** Prudent and Western diets.

**MAIN OUTCOMES AND MEASURES** Incidence of colorectal carcinoma subclassified by *F* nucleatum status in tumor tissue, determined by quantitative polymerase chain reaction.

**RESULTS** Of the 173,229 individuals considered for the study, 137,217 were included in the analysis, 47,449 were male (34.6%), and mean (SD) baseline age for men was 54.0 (9.8) years and for women, 46.3 (7.2) years. A total of 1,019 incident colon and rectal cancer cases with available *F* nucleatum data were documented over 26 to 32 years of follow-up, encompassing 3,643,562 person-years. The association of prudent diet with colorectal cancer significantly differed by tissue *F* nucleatum status (*P* = .01 for heterogeneity); prudent diet score was associated with a lower risk of *F* nucleatum-positive cancers (*P* = .003 for trend; multivariable hazard ratio of 0.43; 95% CI, 0.25-0.72, for the highest vs the lowest prudent score quartile) but not with *F* nucleatum-negative cancers (*P* = .47 for trend, the corresponding multivariable hazard ratio of 0.95; 95% CI, 0.77-1.17). There was no significant heterogeneity between the subgroups in relation to Western dietary pattern scores.

**CONCLUSIONS AND RELEVANCE** Prudent diets rich in whole grains and dietary fiber are associated with a lower risk for *F* nucleatum-positive colorectal cancer but not *F* nucleatum-negative cancer, supporting a potential role for intestinal microbiota in mediating the association between diet and colorectal neoplasms.
Accumulating evidence suggests that the human gut microbiome is linked to colorectal cancer development.1-4 Fusobacterium nucleatum has been found to be enriched in colorectal cancer tissue relative to normal adjacent colonic tissue and is detected at higher levels in stool among individuals with colorectal cancer compared with those without cancer.1,5-10 Recent experimental data suggest that F nucleatum may contribute to colorectal carcinogenesis through modulation of host immunity and activation of pathways associated with cellular proliferation.9,11,12 Furthermore, a higher amount of F nucleatum in colorectal cancer tissue has been linked to shorter survival, proximal tumor location, and specific tumor molecular features, such as high-level CpG island methylator phenotype and microsatellite instability.13-15

Prudent dietary patterns—rich in fruits, vegetables, and whole grains—have been associated with a lower risk of colorectal cancer and adenoma16-21 as reviewed in a recent systematic meta-analysis.22 In contrast, Western dietary patterns—dominated by red and processed meats—have been linked with colorectal carcinogenesis.16,22 Although mechanisms underlying these diet-cancer associations remain unclear, it is postulated that the gut microbiota may play a mediating role.23 Recently, in a dietary intervention study, stool F nucleatum levels markedly increased after participants were switched from a prudent-style, high-fiber, low-fat diet to a low-fiber, high-fat diet.24 In addition, accumulating data suggest that low fiber consumption and high meat intake may be associated with altered bacterial and metagenomic profiles as well as an inflammatory phenotype determined by serum levels of metabolites.25-28

Based on these findings, we hypothesized that the inverse association between prudent diets and risk of colorectal cancer might be more evident for a cancer subgroup enriched with tissue F nucleatum than for a subgroup without detectable tissue F nucleatum. To test this hypothesis, we used 2 US nationwide, prospective cohort studies: the Nurses’ Health Study (NHS) (June 1, 1980, to June 1, 2012) and the Health Professional Follow-up Study (HPFS). These 2 studies offered a unique opportunity to integrate prospectively collected, regularly updated dietary intake data with tissue microbial features in incident colorectal cancers that occurred over long-term follow-up.

Methods

Study Population

We used data drawn from 2 ongoing prospective cohort studies, the NHS and the HPFS. The NHS began in 1976 among 121,700 US female nurses aged 30 to 55 years at enrollment. The HPFS began in 1986 among 51,529 US male health professionals aged 40 to 75 years at enrollment. In both cohorts, participants have returned questionnaires every 2 years, with follow-up rates exceeding 90%, to provide information about lifestyle and dietary factors, medication use, and diagnoses of colorectal cancer and other diseases. The institutional review board at the Brigham and Women’s Hospital and Harvard T.H. Chan School of Public Health approved this study, and informed consent was obtained from all participants. The study was conducted from June 1, 1980, to June 1, 2012.

Of 173,229 individuals considered for the study, a total of 137,217 individuals (47,449 men and 89,768 women) were included in this analysis. We excluded participants with implausibly high or low caloric intakes (ie, <600 or >3500 kcal/d for women and <800 or >4200 kcal/d for men), missing dietary pattern data, or those with a history of ulcerative colitis or cancer (except for nonmelanoma skin cancer) before baseline (1980 for the NHS and 1986 for the HPFS) (eMethods in the Supplement).

Assessment of Diet

Participants reported average food intake over the preceding year (of each questionnaire return) through semiquantitative food frequency questionnaires, which have been previously validated and described.29 Total nutrient intake was calculated by summing intakes from all foods and adjusted for total energy intake by the residual method. As previously described, total dietary fiber was calculated according to methods from the Association of Official Analytic Chemists.30 For this analysis, we used information from food frequency questionnaires administered in the following years: 1980, 1984, 1986, 1990, 1994, 1998, 2002, 2006, and 2010 for the NHS and 1986, 1990, 1994, 1998, 2002, 2006, and 2010 for the HPFS.

Assessment of Colorectal Cancer Cases

In both cohorts, incident cases of colorectal cancer were reported by participants through the 2012 follow-up for the HPFS and NHS. We identified and confirmed lethal colorectal cancer cases through information from various sources including next of kin, the National Death Index, death certificates, and medical records. A study physician (including J.A.M. and C.S.F.), blinded to exposure information, reviewed records and extracted data on histologic type, anatomical location, and stage. The cohort study groups attempted to collect formalin-fixed, paraffin-embedded (FFPE) tissue specimens from hospitals throughout the United States as previously detailed.9 Cases with available tissue data (n = 1019) for the present study were similar to those without tissue data (n = 2241) regarding patient and clinical characteristics (eMethods in the Supplement).
**F nucleatum Analysis**

We extracted DNA from colorectal cancer tissue obtained from sections of FFPE tumor blocks (QiAmp DNA FFPE tissue kits; Qiagen). We performed a real-time polymerase chain reaction (PCR) assay using custom TaqMan primer/probe sets (Applied Biosystems) for the nusG gene of *F. nucleatum*. The interassay coefficient of variation of cycle threshold values from each of 5 selected specimens in 5 different batches was less than 1% for all targets in the validation study. *Fusobacterium nucleatum* positivity was defined as a detectable level of *F. nucleatum* DNA within 45 PCR cycles, and *F. nucleatum* negativity was defined as an undetectable level with a proper amplification of human reference gene SLC02A1 (HGNC: 10955).

**Statistical Analysis**

All statistical tests were 2-sided. To account for multiple testing for the 2 primary hypotheses (related to prudent and Western dietary scores) associated with the 2 tumor subtype variables, we adjusted the 2-sided α level to .01 (approximately .05/4) by simple Bonferroni correction in our primary and secondary analysis.

Two maximally uncorrelated dietary patterns—one named prudent and another named Western—were derived by principal component analysis, as previously described and validated with good reproducibility. Factor loadings were derived based on the correlations between food groups and the 2 derived factors. Each participant was assigned a factor score, determined by adding the reported frequencies of food group intakes weighted by the factor loadings. These factor scores were then standardized to have a mean (SD) of 0 (1). To capture long-term habitual consumption, we calculated the cumulative mean of the prudent (or Western) dietary pattern scores from preceding food frequency questionnaires up to each questionnaire cycle. Then, the cumulative average score was categorized into sex-specific quartiles and used as the primary exposure variable.

Using Cox proportional hazards regression models, we computed hazard ratios (HRs) to examine the association of the prudent or Western dietary score with incidence of colorectal cancer. To test for trend with the Wald test, participants with high prudent scores in the HPFS and NHANES tended to smoke less, exercise more, and have greater rates of lower gastrointestinal endoscopy, whereas Western pattern scores were associated with behaviors typically considered unhealthy (eTable 2 in the Supplement).

**Results**

Of the 137,217 individuals included in the analysis, 47,449 were male (34.6%); mean (SD) baseline age for men was 54.0 (9.8) years and for women, 46.3 (7.2) years. Two major, uncorrelated dietary patterns were identified by factor analysis. The prudent dietary pattern was characterized by high intake of vegetables, fruits, whole grains, and legumes, and the Western dietary pattern was characterized by red and processed meats, refined grains, and desserts (eTable 1 in the Supplement). Consistent with prior analyses, participants with high prudent scores in the HPFS and NHS tended to smoke less, exercise more, and have greater rates of lower gastrointestinal endoscopy, whereas Western pattern scores were associated with behaviors typically considered unhealthy (eTable 2 in the Supplement).

After 26 years (in HPFS) and 32 years (in NHS) of follow-up encompassing 3,643,562 person-years, we documented 1,019 incident colorectal cancers with available data on tissue *F. nucleatum* status. Among these cancer cases, there were 125 (12.3%) *F. nucleatum*-positive tumors and 894 (87.7%) *F. nucleatum*-negative tumors. We examined the association of prudent and Western dietary pattern scores with the incidence of overall colorectal cancer. Western dietary pattern scores showed a trend toward associations with overall risk of colorectal cancer in the HPFS (eTable 3 in the Supplement) and the combined cohort (Table); however, statistical significance was not reached with the adjusted level of .01. We did not observe significant heterogeneity in the associations of the dietary scores with colorectal cancer risk between the 2 cohorts (*P* > .21). To maximize statistical power, we used the combined cohort for further analyses.

We then tested our primary hypothesis that the association of prudent and Western diets with colorectal cancer incidence might differ according to the presence of *F. nucleatum* in tumor tissue. Notably, the association between prudent dietary pattern and risk of colorectal cancer significantly differed by tumor *F. nucleatum* status (*P* = .01 for heterogeneity).
We found a significant inverse association of prudent dietary scores with *F. nucleatum*-positive cancer risk (P = .003 for trend) but not with *F. nucleatum*-negative cancer risk (P = .47 for trend). Comparing participants in the highest prudent dietary score quartile with those in the lowest quartile, the multivariable HR for *F. nucleatum*-positive tumors was 0.43 (95% CI, 0.25-0.72); in contrast, the corresponding HR for *F. nucleatum*-negative tumors was 0.95 (95% CI, 0.77-1.17). We found similar differential associations by *F. nucleatum* status in men (HPFS) and women (NHS), although statistical power was limited (eTable 4 in the Supplement). Because we observed that the fraction of colorectal cancers enriched with *F. nucleatum* gradually decreased from cecum to rectum,33 we conducted exploratory analyses stratified by tumor location (eTable 6 in the Supplement). The differential association of prudent diet score with colorectal cancer by tissue *F. nucleatum* status appeared to be consistent in both proximal and distal cancer strata.

When we examined the association of the Western dietary pattern with colorectal cancer subgroups according to *F. nucleatum* status, we observed similar associations in men (HPFS) and women (NHS), although limited statistical power precluded definitive conclusions (eTable 5 in the Supplement).
tumor *F. nucleatum* status, although Western dietary pattern scores appeared to be more strongly associated with *F. nucleatum*-positive cancer risk, there was no significant heterogeneity between the subgroups (*P* = .23 for heterogeneity) (Table).

In a secondary analysis, we sought to determine whether specific food groups might explain the observed differential associations between prudent dietary patterns and risk of colorectal cancer according to *F. nucleatum* status. We examined the top 4 dominantly contributing food groups to the prudent diet pattern (vegetables, fruits, legumes, and whole grains) in relation to the risk of colorectal cancer according to *F. nucleatum* status. We observed no significant heterogeneity (with the adjusted α of .01).

Finally, to further determine whether any specific macronutrient components of the prudent dietary pattern might explain the observed differential associations according to *F. nucleatum* status, we explored associations of fiber, fat, and protein intake with colorectal cancer subgroups (eTable 8 in the Supplement). There appeared to be heterogeneity in the differential association of fiber intake with cancer subgroups classified by *F. nucleatum* status (*P* = .02 for heterogeneity), similar to the findings for prudent dietary pattern scores. Comparing participants in the highest quartile of fiber intake (>26 g/d for men and >19 g/d for women) with those in the lowest quartile (<18 g/d for men and <13 g/d for women), the multivariable HR for *F. nucleatum*-positive tumors was 0.54 (95% CI, 0.32-0.92); in contrast, the corresponding HR for *F. nucleatum*-negative tumors was 1.13 (95% CI, 0.92-1.40). In further exploratory analyses, we found that intakes of cereal-derived fiber might be differentially associated with colorectal cancer according to *F. nucleatum* status (*P* = .01 for heterogeneity) (eTable 9 in the Supplement). We did not observe such heterogeneity for fat or protein.

**Discussion**

In the 2 US nationwide prospective cohorts, we found that participants with higher long-term prudent dietary pattern scores were associated with a lower risk of *F. nucleatum*-positive colorectal cancers but not *F. nucleatum*-negative cancers. Our data also suggest that higher intakes of dietary fiber, one of the components of the prudent diet, may be associated with a lower risk of *F. nucleatum*-positive colorectal cancer but not *F. nucleatum*-negative cancer. These findings support the hypothesis that the possible cancer-preventive effects of prudent diets rich in dietary fiber may be mediated by modulation of specific species in the gut microbiota and subsequent alteration of the amount of *F. nucleatum* in local colonic tissue. To our knowledge, our study represents the first to examine the intersection of diet and incidence of colorectal cancer subgroups according to microbial status in human tumor tissue.

The potential role of diet in modulating the risk of a variety of diseases, including colorectal cancer, has been widely recognized.23,34 According to the World Cancer Research Fund and American Institute for Cancer Research, foods with fiber including whole grains are one of the strongest factors linked to decreasing the risk of colorectal cancer.35 However, there has been considerable heterogeneity in the epidemiologic data associating prudent dietary patterns and the major components of the prudent diet with colorectal cancer.36 Our results here suggest that the inconsistency in the association of prudent dietary patterns (and components of the diet) with lower colorectal cancer risk may be in part attributable to differential associations with cancer subgroups according to *F. nucleatum* in tumor tissue. In addition, given recent findings between increasing amounts of *F. nucleatum* DNA in colorectal cancer tissue and worsened survival,14 our data lend additional support to the promotion of healthy diets to reduce mortality from colorectal cancer.

The precise mechanism by which prudent diets rich in dietary fiber may lower *F. nucleatum*-enriched cancer incidence remains unclear. Accumulating evidence suggests that long-term dietary fiber intake has a profound effect on the gut microbiome, specifically through promotion of microbial diversity and by lowering levels of inflammatory metabolites.25,37-40 A recent study showed that a 2-week feeding intervention switching rural-dwelling South Africans from a high-fiber, low-fat diet to a low-fiber, high-fat diet was associated with an increase in *F. nucleatum* measured by PCR in the stool.24 In addition, some have hypothesized that the variation observed in *F. nucleatum* levels in colorectal cancers collected from Spain, Vietnam, Japan, and the United States may be attributable to differences in dietary practices in these countries.5,41 Furthermore, in a cross-sectional study, participants with advanced adenoma were associated with lower dietary fiber intakes as well as distinct fecal microbiome communities compared with healthy controls.42 It is plausible that an abundance of microbiota-accessible carbohydrates from prudent diets may influence bacterial fermentation of dietary fiber, resulting in altered levels of short-chain fatty acids. These changes may alter pH, increase transit time of gut contents, or lead to differences in local immune surveillance, which are less hospitable for nonnative species, such as *F. nucleatum*, to establish themselves in the colonic niche and potentiate colorectal carcinogenesis.24,25,43,44 Taken together, these data provide evidence of substantial influences of diet on the gut microbiome, which may in turn influence tumorigenesis.

There are several strengths in this study. First, our dietary data were prospectively collected and have been well validated.29 Second, our data were detailed and updated such that we could examine long-term effects of overall dietary patterns, specific food groups, and macronutrients in relation to colorectal cancer risk. Third, we collected detailed data on multiple potential confounders, although residual confounding cannot be excluded. Finally, our molecular pathological epidemiology (MPE) research46 provides refined risk estimates for specific cancer subgroups, such as *F. nucleatum*-positive cancer, and thereby offers insights into pathogenesis and causality. Molecular subtyping in the MPE approach can gather pathogenetically similar cases, and thus can enhance statistical inference (even with a relatively small number of cases).46 The present study represents emerging unique microbial MPE research in which the microbial feature in tumor tissue can serve as a pathogenic signature.
**Limitations**

We acknowledge limitations of this study. First, this study was observational, and residual confounding may be an issue. Nevertheless, adjustment for a variety of known risk factors for colorectal cancer showed no substantial effect on the results. Second, the diet data were derived from food frequency questionnaires and subject to measurement errors. Nonetheless, studies have shown that food frequency questionnaires can capture long-term dietary intakes than detailed diet diaries in a limited period.\(^4\) Third, with the use of FFPE tissue specimens, routine histopathological procedures might have influenced performance characteristics of our PCR assay to detect *F. nucleatum*. Nonetheless, we conducted a rigorous validation study that showed high precision of our PCR assay to detect *F. nucleatum*.\(^9\) Moreover, our assay has previously been shown to have high specificity for *F. nucleatum*.\(^6\) Fourth, we could not collect FFPE blocks from all colorectal cancer cases in the cohorts; nonetheless, cases with available tissue were generally similar to those without tissue with regard to patient characteristics. Fifth, because our participants were all health professionals and most were white, generalizability of the findings to other populations needs to be examined in future studies.

**Conclusions**

This study has shown that a prudent diet is associated with a lower risk of *F. nucleatum*-positive colorectal cancer but not *F. nucleatum*-negative cancer. Our data generate new hypotheses about how the intestinal microbiota may mediate the association between diet and colorectal neoplasms. Further studies are needed to confirm these findings and determine the potential utility of characterization of *F. nucleatum* in colonic mucosa, tumor, or stool as a biomarker for personalized nutritional, probiotic, or antibiotic interventions. In addition, our findings underscore the importance of future large-scale, prospective studies that examine the gut microbiota to understand the complex intersection of diet, the gut microbiome, and carcinogenesis.\(^49\)
Diet and Risk of Colorectal Cancer Subtypes Classified by F. nucleatum


