



# Calcium intake and colon cancer risk subtypes by tumor molecular characteristics

NaNa Keum<sup>1,2</sup> · Li Liu<sup>1,3,4</sup> · Tsuyoshi Hamada<sup>3</sup> · Zhi Rong Qian<sup>3</sup> · Jonathan A. Nowak<sup>5</sup> · Yin Cao<sup>6</sup> · Annacarolina da Silva<sup>3</sup> · Keisuke Kosumi<sup>3</sup> · Mingyang Song<sup>1,7,8</sup> · Daniel Nevo<sup>9,10</sup> · Molin Wang<sup>9,10</sup> · Andrew T. Chan<sup>7,8,11</sup> · Jeffrey A. Meyerhardt<sup>3</sup> · Charles S. Fuchs<sup>12,13,14</sup> · Kana Wu<sup>1</sup> · Shuji Ogino<sup>1,3,5,10,15</sup> · Reiko Nishihara<sup>1,3,5,9,10</sup> · Xuehong Zhang<sup>11</sup>

Received: 15 November 2018 / Accepted: 28 March 2019 / Published online: 8 April 2019  
© Springer Nature Switzerland AG 2019

## Abstract

**Background** A preventive potential of high calcium intake against colorectal cancer has been indicated for distal colon cancer, which is inversely associated with high-level CpG island methylator phenotype (CIMP), high-level microsatellite instability (MSI), and *BRAF* and *PIK3CA* mutations. In addition, *BRAF* mutation is strongly inversely correlated with *KRAS* mutation. We hypothesized that the association between calcium intake and colon cancer risk might vary by these molecular features.

**Methods** We prospectively followed 88,506 women from the Nurses' Health Study and 47,733 men from the Health Professionals Follow-up Study for up to 30 years. Duplication-method Cox proportional cause-specific hazards regression was used to estimate multivariable hazard ratios (HRs), and 95% confidence intervals (95% CIs) for the associations between calcium intake and the risk of colon cancer subtypes. By Bonferroni correction, the  $\alpha$ -level was adjusted to 0.01.

**Results** Based on 853 colon cancer cases, the inverse association between dietary calcium intake and colon cancer risk differed by CIMP status ( $p_{\text{heterogeneity}} = 0.01$ ). Per each 300 mg/day increase in intake, multivariable HRs were 0.84 (95% CI 0.76–0.94) for CIMP-negative/low and 1.12 (95% CI 0.93–1.34) for CIMP-high. Similar differential associations were suggested for MSI subtypes ( $p_{\text{heterogeneity}} = 0.02$ ), with the corresponding HR being 0.86 (95% CI 0.77–0.95) for non-MSI-high and 1.10 (95% CI 0.92–1.32) for MSI-high. No differential associations were observed by *BRAF*, *KRAS*, or *PIK3CA* mutations.

**Conclusion** The inverse association between dietary calcium intake and colon cancer risk may be specific to CIMP-negative/low and possibly non-MSI-high subtypes.

**Keywords** Dietary calcium · Colorectal cancer · Colon cancer · MSI · CIMP

---

Use of standardized official symbols: We use HUGO (Human Genome Organisation)—approved official symbols for genes and gene products, including *BRAF*, *CASR*, *KRAS*, and *PIK3CA*; all of which are described at [www.genenames.org](http://www.genenames.org). The official symbols are italicized, to differentiate from nonitalicized colloquial names that are used along with the official symbols. This format enables readers to familiarize the official symbols for genes and gene products together with common colloquial names.

---

NaNa Keum, Li Liu, and Tsuyoshi Hamada have contributed equally as co-first authors.

---

Reiko Nishihara, and Xuehong Zhang have contributed equally as co-last authors.

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s10552-019-01165-3>) contains supplementary material, which is available to authorized users.

---

Extended author information available on the last page of the article

## Abbreviations

BMI	Body mass index
CASR	Calcium sensing receptor
CIMP	CpG island methylator phenotype
FFQ	Food frequency questionnaire
MSI	Microsatellite instability
NHS	Nurses' Health Study
HPFS	Health Professionals Follow-up Study

## Introduction

Calcium has long been suggested as a potential chemopreventive agent against colorectal cancer [1]. Most observational studies found a modest but statistically significant inverse association between calcium intake and colorectal

cancer risk [2–5]. Yet, evidence from randomized controlled trials on calcium supplements and colorectal neoplasms remains inconclusive [6–8]. Colorectal cancer is an etiologically heterogeneous disease induced by diverse combinations of genomic or epigenomic alterations [9], but most previous studies have evaluated total colorectal cancer in aggregate. Hence, epidemiologic studies integrating major molecular events observed in colorectal cancer may provide additional insights into the role of calcium in colorectal carcinogenesis.

By cancer site, a statistically significant inverse association with calcium intake was observed for colon cancer (primarily distal colon cancer) but generally not for rectal cancer [1–3, 10, 11]. In the distal colon cancer, tumor molecular features such as high-level CpG island methylator phenotype (CIMP), high-level microsatellite instability (MSI), *BRAF*, and *PIK3CA* mutations are less prevalent than in proximal colon cancer [12–15]. In addition, *BRAF* mutation is strongly inversely correlated with *KRAS* mutation in colorectal cancer [16]. In light of this biological evidence, we hypothesized that the association between calcium intake and colon cancer risk might differ according to tumor molecular features according to status of CIMP, MSI, *BRAF*, *KRAS* or *PIK3CA* mutations. Given sizable overlap between the CIMP and MSI pathways [17, 18], we also examined the association according to their joint categories.

## Methods

### Study population

Participants were identified from two ongoing prospective cohort studies in the U.S., the Nurses' Health Study (NHS), and the Health Professionals Follow-up Study (HPFS). The NHS was established in 1976, enrolling 121,700 female registered nurses aged 30–55 years. The HPFS was established in 1986, enrolling 51,529 male health professionals aged 40–75 years. In each cohort, follow-up rates have exceeded 90% in each 2-year cycle. Through a baseline questionnaire and biennial follow-up questionnaires, participants provided updated information concerning demographics, lifestyle factors, and medical history. The institutional review boards at the Brigham and Women's Hospital and the Harvard T.H. Chan School of Public Health approved this study, and return of completed questionnaires was considered to imply informed consent.

Calcium intake was first assessed in 1980 in NHS and 1986 in HPFS. At baseline, we excluded participants who had cancer (except for non-melanoma skin cancer), inflammatory bowel disease, and missing data on calcium

intake. The final analytic cohort included 88,506 women and 47,733 men (Supplemental Fig. 1).

### Assessments of calcium intake

Dietary intakes during the preceding year were assessed through a validated semiquantitative food frequency questionnaire (FFQ) [19, 20], which was administered in 1980, 1984, and 1986 in NHS, 1986 in HPFS, and every 4 years thereafter through 2010 in each cohort. Information on multivitamin and other supplement use was collected via biennial questionnaires starting from the baseline.

As previously described, dietary calcium intake (i.e., calcium from food sources) was estimated by calculating, for each calcium-containing food on FFQ, the product of its consumption frequency and calcium content of its specified portion size, and then by summing the products across the calcium-containing foods [10]. As the major contributor to dietary calcium intake [21], dairy calcium intake was separately derived by summing calcium contributions from dairy foods (milk, yogurt, cheese, butter, cream, etc.) and foods with dairy ingredients (e.g., baked goods, chocolates, ice cream, etc.) [2, 10]. Supplemental calcium intake was estimated from multivitamins and calcium supplements. Total calcium intake was calculated by summing dietary and supplemental calcium intake.

The correlation coefficients comparing FFQ and the reference (multiple one-week dietary records) ranged from 0.60 to 0.70 for energy-adjusted total and dietary calcium intakes in each cohort [10, 20, 22].

### Assessment of covariates

From the baseline and follow-up questionnaires, we collected information on age, race, family history of colorectal cancer, history of sigmoidoscopy/colonoscopy, aspirin use, smoking, physical activity [23], and intakes of alcohol, red/processed meat, and folate. Body mass index (BMI) was calculated based on reported height and weight. From known predictors of plasma 25-hydroxy vitamin D [25(OH)D], we calculated 25(OH)D score, which statistically significantly predicted reduced colorectal cancer risk in our cohort [24].

### Ascertainments of incident colorectal cancer and death

We used colon cancer as the primary outcome and colorectal cancer as the secondary. Incident colorectal cancer was reported by participants on biennial follow-up questionnaires through 2010. Unreported fatal colorectal cancer and death were ascertained based on reports from family or postal

authorities or the National Death Index. After obtaining permission from participants or next-of-kin, study physicians blinded to participants' exposure status reviewed medical records to confirm the diagnosis and to extract information on tumor characteristics including anatomic location. These methods confirmed over 90% of incident colorectal cancer [25] and > 98% of deaths in the cohorts [26].

## Ascertainments of tumor molecular markers

### Tumor tissue collection and genomic DNA extraction

Formalin-fixed paraffin-embedded tissue blocks were collected from hospitals where participants with colon cancer underwent tumor resection. Histopathologic features were reviewed by the study pathologist (S.O.). DNA was extracted from tumor tissue as previously described [27].

### DNA methylation analysis for CIMP

Using a bisulfite-treated DNA and real-time PCR (Methyl-Light assay) [28], we quantified DNA methylation in eight CIMP-specific promoters [29]. Consistent with previous definition [18], tumors with  $\geq 6/8$  methylated makers were classified as CIMP-high, and those with  $< 6/8$  methylated markers as CIMP-negative/low [29].

### MSI analysis

MSI status was determined using 10 microsatellite markers [30]. Tumors with instability in  $\geq 30\%$  of the markers were classified as MSI-high, and those with instability  $< 30\%$  of the markers as non-MSI-high.

### Sequencing of BRAF, KRAS, and PIK3CA

Polymerase chain reaction (PCR) and pyrosequencing targeted for *BRAF* (codon 600) [30] *KRAS* (codons 12, 13, 61, and 146) [31, 32], and *PIK3CA* (exons 9 and 20) [33, 34] were performed as previously described.

### Statistical analysis

Evidence suggests that calcium intake in the distant past (e.g., at least a decade) is likely etiologically relevant [2, 3]. To address this temporal association, our primary analyses were based on a latency analysis, where we examined calcium intake with colon cancer diagnosed 8–12 years after the intake assessment [35]. Thus, accrual of person-time of follow-up started 8 years after the date of baseline questionnaire return until the date of colorectal cancer diagnosis, death from any cause, or end of follow-up (June 2010 for NHS, January 2010 for HPFS), whichever came first. We

conducted sensitivity analyses by including all incident colon cancer cases during the entire follow-up period.

In modeling each source of calcium intake, the cutoffs in categorical analysis and the unit increment of 300 mg/day in linear trend analysis were set as reported in our previous publication [2]. Individuals with a high-energy intake are likely to have a high calcium intake. To reduce extraneous variation in calcium intake attributable to energy intake, calcium intake at each questionnaire cycle was adjusted for energy intake using the residual method [36]. Further, to minimize random measurement error in calcium intake, we calculated the cumulative average of the energy-adjusted calcium intakes from the baseline questionnaire up to the most recent follow-up questionnaire excluding the lagging period. For instance, we associated energy-adjusted calcium intakes averaged up to 1990 questionnaires with colon cancer incidence occurred between 1998 and 2000.

We calculated hazard ratios (HRs) and 95% confidence intervals (95% CIs) for the association between calcium intake (total, dietary, supplemental, and dairy) and colon cancer risk overall and by molecular subtypes using duplication-method Cox proportional cause-specific hazards regression [37]. We censored colon cancer with missing tumor marker data at the time of the diagnosis and adjusted for time-varying covariates when appropriate (for the list of adjusted variables, see footnotes of Tables). We observed no violation of the proportional hazard assumption from the Wald test performed on an additionally added interaction term between continuous calcium intake and continuous age. By Cochran's Q test [38], we observed no statistically significant heterogeneity in the linear trend by gender, except a few subtypes in the analysis with supplemental calcium (see footnote of Table 5). Thus, we conducted primary analyses in a pooled cohort of NHS and HPFS to maximize statistical power; and in secondary analyses, we examined the associations separately in each cohort. We evaluated the heterogeneity by colon cancer subtypes using the likelihood ratio test, by comparing the model in which a linear association with calcium intake was allowed to vary by tumor subtypes with the model in which we assumed a common association [37]. Tumor subtypes tested were selected a priori based on evidence indicative of differential frequencies of occurrence across the segments of colon. We repeated comparable analyses using the secondary outcome (i.e., colorectal cancer).

In the cohorts, not all colon cancer cases provided tumor tissues, and missing tumor subtype information could occur nonrandomly. As a sensitivity analysis to address the concern for potential selection bias, we assessed the associations between calcium intakes and colon cancer subtypes using inverse probability weighting [39]. For each molecular marker, the probability of having tumor-subtype information was estimated by fitting a logistic regression among all colon cancer cases within

each cohort, using age, and year of colon cancer diagnosis and tumor stage, grade, and location as the model covariates.

All statistical tests were two-sided. Given multiple hypothesis testing performed in the heterogeneity tests by five tumor markers, we applied Bonferroni correction and adjusted the significance level to 0.01 ( $=0.05/5$ ). Analyses were performed using SAS 9.3 (SAS Institute, Cary, NC).

## Results

In the primary (8–12 years lagged) analysis with up to 30 years of follow-up, we accumulated 2,354,710 person-years and documented 1,843 incident colon cancer cases, of which 853 cases had information on at least one tumor marker analyzed. Participants with higher total calcium intakes tended to be older, undergo sigmoidoscopy/colonoscopy, use aspirin regularly, take multivitamins, smoke

less, and engage in more physical activities (Table 1). They were likely to consume less alcohol, red meat, and processed meat, but to consume more folate and to have higher predicted vitamin D score. Dietary calcium was the major contributor to total calcium intake, with supplemental calcium becoming an important source for those with total calcium intake of  $\geq 1200$  mg/day.

## Total calcium

A linear inverse association did not vary statistically significantly by any of the five tumor molecular markers tested in the pooled cohort (Tables 2 and 3). Within men, differential linear associations were suggested by *BRAF* mutation status ( $p_{\text{heterogeneity}} = 0.02$ ; Supplemental Table 1), with HR per 300 mg/day increase in intake being 0.95 (95% CI 0.85, 1.07) for wild-type and 1.32 (95% CI 1.05, 1.66) for mutant.

**Table 1** Age-standardized characteristics of person-years by total calcium intake in men (1986–2010) and women (1980–2010)

Characteristics <sup>a</sup>	Total calcium intake (mg/day)				
	0–599	600–799	800–999	1,000–1,199	$\geq 1,200$
No. person-years	490,392	628,765	513,794	342,430	379,329
Calcium intake by sources, mg/day					
Total calcium intake	491 (83)	700 (57)	894 (57)	1092 (57)	1,454 (233)
Dietary calcium	482 (84)	657 (89)	787 (137)	903 (195)	1,047 (293)
Dairy calcium	213 (81)	360 (101)	482 (145)	595 (204)	731 (302)
Supplemental calcium	14 (33)	47 (75)	109 (130)	188 (189)	391 (340)
Age, years	62 (9.3)	64 (9.5)	65 (9.4)	66 (9.4)	68 (9.3)
Men, %	26	30	27	24	28
Caucasian, %	95	97	98	98	98
Family history of colorectal cancer, %	14	14	15	15	14
History of sigmoidoscopy/colonoscopy, %	31	39	42	44	46
Regular aspirin use, <sup>b</sup> %	38	41	43	43	44
Multivitamin use, %	41	50	58	62	66
Ever smokers, %	54	51	50	49	48
BMI, kg/m <sup>2</sup>	25.4 (4.3)	25.6 (4.2)	25.5 (4.2)	25.3 (4.2)	25.3 (4.2)
Physical activity, MET-hours/week	17.7 (20.5)	20.8 (21.9)	21.6 (21.9)	22.0 (22.0)	23.8 (23.9)
Dietary intake					
Total energy, kcal/day	1,716 (492)	1,791 (477)	1,810 (473)	1,818 (464)	1,730 (453)
Alcohol, g/day	9.2 (13.5)	7.8 (10.8)	6.8 (9.6)	6.2 (8.9)	5.5 (8.3)
Red meat, servings/week	2.3 (1.5)	2.0 (1.2)	1.8 (1.2)	1.7 (1.1)	1.4 (1.0)
Processed meat, servings/week	1.1 (1.4)	1.0 (1.2)	0.9 (1.1)	0.8 (1.1)	0.7 (1.0)
Folate, $\mu\text{g/day}$	377 (184)	439 (190)	482 (193)	513 (199)	587 (261)
Predicted plasma 25-hydroxy vitamin D score, ng/mL	7.4 (10.4)	8.9 (10.5)	8.7 (10.1)	8.5 (9.6)	10.2 (9.9)

BMI body mass index, MET metabolic equivalent task score

<sup>a</sup>Values were mean (SD) or percentage and all values, except age, were standardized to the age distribution of the study population during follow-up

<sup>b</sup>Regular aspirin use was defined as  $\geq 2$  tablets/week

**Table 2** HR and 95% CI of total calcium intake and colon cancer risk overall and by molecular subtypes

	Total calcium intake (mg/day)					Per 300	$P_{\text{trend}}^a$	$P_{\text{heterogeneity}}^b$
	0–599	600–799	800–999	1,000–1,199	≥ 1,200			
<b>Overall</b>								NR
Cases, <i>n</i>	407	533	373	264	266	1,843		
Age-adjusted <sup>c</sup>	1 (ref)	0.93 (0.81, 1.05)	0.77 (0.66, 0.87)	0.78 (0.67, 0.92)	0.63 (0.54, 0.74)	0.88 (0.84, 0.92)	< 0.001	
Multivariable <sup>d</sup>	1 (ref)	0.97 (0.84, 1.10)	0.83 (0.72, 0.97)	0.88 (0.74, 1.04)	0.74 (0.62, 0.88)	0.92 (0.88, 0.97)	< 0.001	
<b>CIMP status negative/low</b>								0.18
Cases, <i>n</i>	137	186	126	71	84	604		
Age-adjusted <sup>c</sup>	1 (ref)	1.00 (0.80, 1.25)	0.81 (0.63, 1.03)	0.69 (0.51, 0.92)	0.64 (0.49, 0.85)	0.89 (0.83, 0.96)	0.002	
Multivariable <sup>d</sup>	1 (ref)	1.02 (0.81, 1.28)	0.86 (0.66, 1.10)	0.74 (0.55, 1.00)	0.74 (0.55, 0.99)	0.93 (0.86, 1.01)	0.08	
<b>High</b>								
Cases, <i>n</i>	30	51	41	33	34	189		
Age-adjusted <sup>c</sup>	1 (ref)	1.18 (0.75, 1.86)	1.05 (0.65, 1.68)	1.15 (0.70, 1.90)	0.98 (0.59, 1.61)	0.98 (0.86, 1.11)	0.75	
Multivariable <sup>d</sup>	1 (ref)	1.20 (0.76, 1.90)	1.11 (0.68, 1.79)	1.25 (0.75, 2.07)	1.12 (0.67, 1.87)	1.03 (0.90, 1.18)	0.66	
<b>MSI status</b>								0.24
<b>Non-MSI-high</b>								
Cases, <i>n</i>	136	205	134	76	90	641		
Age-adjusted <sup>c</sup>	1 (ref)	1.09 (0.88, 1.36)	0.84 (0.66, 1.07)	0.71 (0.53, 0.95)	0.67 (0.51, 0.88)	0.88 (0.82, 0.95)	< 0.001	
Multivariable <sup>d</sup>	1 (ref)	1.12 (0.90, 1.40)	0.91 (0.71, 1.17)	0.79 (0.59, 1.06)	0.79 (0.59, 1.06)	0.93 (0.86, 1.01)	0.08	
<b>High</b>								
Cases, <i>n</i>	31	47	43	31	31	183		
Age-adjusted <sup>c</sup>	1 (ref)	1.04 (0.66, 1.64)	1.05 (0.66, 1.68)	1.08 (0.65, 1.79)	0.86 (0.52, 1.42)	0.96 (0.85, 1.10)	0.58	
Multivariable <sup>d</sup>	1 (ref)	1.07 (0.68, 1.69)	1.13 (0.70, 1.81)	1.20 (0.72, 2.00)	1.02 (0.61, 1.71)	1.02 (0.89, 1.17)	0.76	
<b>BRAF mutation</b>								0.10
<b>(–)</b>								
Cases, <i>n</i>	140	210	137	77	95	659		
Age-adjusted <sup>c</sup>	1 (ref)	1.08 (0.87, 1.34)	0.83 (0.65, 1.05)	0.70 (0.53, 0.92)	0.68 (0.52, 0.89)	0.88 (0.82, 0.94)	< 0.001	
Multivariable <sup>d</sup>	1 (ref)	1.11 (0.89, 1.38)	0.89 (0.69, 1.14)	0.76 (0.57, 1.02)	0.80 (0.60, 1.06)	0.92 (0.85, 0.99)	0.04	
<b>(+)</b>								
Cases, <i>n</i>	30	42	41	31	27	171		
Age-adjusted <sup>c</sup>	1 (ref)	0.98 (0.61, 1.57)	1.08 (0.67, 1.73)	1.13 (0.68, 1.89)	0.82 (0.48, 1.39)	0.99 (0.87, 1.13)	0.90	
Multivariable <sup>d</sup>	1 (ref)	1.00 (0.62, 1.61)	1.15 (0.71, 1.86)	1.26 (0.75, 2.11)	0.96 (0.56, 1.64)	1.05 (0.91, 1.20)	0.53	
<b>KRAS mutation</b>								0.57
<b>(–)</b>								
Cases, <i>n</i>	117	152	120	68	84	541		
Age-adjusted <sup>c</sup>	1 (ref)	0.93 (0.73, 1.19)	0.86 (0.66, 1.11)	0.71 (0.53, 0.96)	0.71 (0.53, 0.94)	0.91 (0.85, 0.99)	0.02	
Multivariable <sup>d</sup>	1 (ref)	0.96 (0.75, 1.23)	0.92 (0.71, 1.20)	0.79 (0.58, 1.08)	0.83 (0.61, 1.13)	0.96 (0.89, 1.05)	0.37	
<b>(+)</b>								
Cases, <i>n</i>	54	100	59	40	39	292		
Age-adjusted <sup>c</sup>	1 (ref)	1.31 (0.94, 1.83)	0.90 (0.62, 1.31)	0.91 (0.60, 1.38)	0.71 (0.47, 1.08)	0.88 (0.79, 0.98)	0.02	
Multivariable <sup>d</sup>	1 (ref)	1.33 (0.95, 1.86)	0.96 (0.66, 1.41)	0.99 (0.65, 1.51)	0.82 (0.53, 1.25)	0.93 (0.83, 1.03)	0.18	
<b>PIK3CA mutation</b>								0.58
<b>(–)</b>								
Cases, <i>n</i>	133	195	139	80	100	647		
Age-adjusted <sup>c</sup>	1 (ref)	1.04 (0.83, 1.30)	0.86 (0.67, 1.09)	0.72 (0.55, 0.96)	0.74 (0.56, 0.96)	0.91 (0.85, 0.98)	0.01	
Multivariable <sup>d</sup>	1 (ref)	1.07 (0.86, 1.35)	0.93 (0.72, 1.19)	0.80 (0.60, 1.08)	0.87 (0.65, 1.16)	0.96 (0.89, 1.04)	0.34	
<b>(+)</b>								
Cases, <i>n</i>	22	40	27	23	16	128		
Age-adjusted <sup>c</sup>	1 (ref)	1.18 (0.70, 1.99)	0.96 (0.54, 1.69)	1.23 (0.68, 2.23)	0.63 (0.33, 1.21)	0.87 (0.74, 1.02)	0.09	
Multivariable <sup>d</sup>	1 (ref)	1.21 (0.72, 2.06)	1.04 (0.59, 1.85)	1.35 (0.74, 2.45)	0.74 (0.38, 1.43)	0.92 (0.78, 1.08)	0.30	

**Table 2** (continued)

CIMP CpG island methylator phenotype, MSI microsatellite instability, NR not relevant

<sup>a</sup>The *p* value for linear trend across calcium intake was from the Wald test on the continuous term of calcium intake

<sup>b</sup>The *p* value for heterogeneity across tumor markers was from the likelihood ratio test comparing the model in which a linear association with calcium intake was allowed to vary by tumor subtypes with the model in which a common association was assumed

<sup>c</sup>Age-adjusted analysis was stratified by age (continuous), questionnaire cycle (continuous), and sex (men vs. women)

<sup>d</sup>Multivariable analysis was stratified by age (continuous), questionnaire cycle (continuous), and sex (men vs. women); adjusted for Caucasian (yes vs. no), family history of colorectal cancer (yes vs. no), history of sigmoidoscopy/colonoscopy (yes vs. no), regular aspirin use (yes vs. no), smoking (0, 1–9.9, 10+ pack-years), BMI (<25, 25–27.4, 27.5–29.9, 30+ kg/m<sup>2</sup>), physical activity (<3, 3–26.9, 27+ MET-hours/week), 25-hydroxyvitamin D scores (quintiles), and intakes of energy (quintiles), alcohol (quintiles), red and processed meat (quintiles), and folate (quintiles)

**Table 3** Multivariable HR<sup>a</sup> and 95% CI of calcium intake and colon cancer risk by the joint CIMP and MSI status

Sources of calcium	CIMP and MSI status				<i>p</i> <sup>b</sup> <sub>heterogeneity</sub>
	CIMP-negative/low; non-MSI-high	CIMP-negative/low; MSI-high	CIMP-high; non-MSI-high	CIMP-high; MSI-high	
Cases, n	544	38	50	136	
Total calcium					
Per 300 mg/day	0.94 (0.87, 1.03)	0.92 (0.68, 1.24)	0.99 (0.76, 1.27)	1.04 (0.89, 1.21)	0.72
<i>p</i> <sup>c</sup> <sub>trend</sub>	0.17	0.57	0.92	0.63	
Dietary calcium					
Per 300 mg/day	0.86 (0.77, 0.96)	0.77 (0.50, 1.19)	0.97 (0.67, 1.39)	1.19 (0.96, 1.47)	0.05
<i>p</i> <sup>c</sup> <sub>trend</sub>	0.01	0.24	0.85	0.11	
Supplemental calcium					
Per 300 mg/day	1.03 (0.91, 1.16)	1.07 (0.71, 1.62)	1.06 (0.74, 1.51)	0.91 (0.71, 1.17)	0.81
<i>p</i> <sup>c</sup> <sub>trend</sub>	0.66	0.74	0.76	0.45	

CIMP CpG island methylator phenotype, MSI microsatellite instability

<sup>a</sup>Multivariable HR was adjusted for the same set of variables as denoted in Table 2; additionally, dietary and supplemental calcium intakes were mutually adjusted for

<sup>b</sup>The *p* value for heterogeneity across tumor markers was from the likelihood ratio test comparing the model in which a linear association with calcium intake was allowed to vary by tumor subtypes with the model in which a common association was assumed

<sup>c</sup>The *p* value for linear trend across calcium intake was from the Wald test on the continuous term of calcium intake

## Dietary calcium

The linear associations of calcium intake with incidence of colon cancer subtypes differed statistically significantly by CIMP status after Bonferroni's correction ( $p_{\text{heterogeneity}} = 0.01$ ; Table 4). Per 300 mg/day increase in intake, multivariable HRs were 0.84 (95% CI 0.76, 0.94) for CIMP-negative/low colon cancer and 1.12 (95% CI 0.93, 1.34) for CIMP-high colon cancer. Similarly, evidence was suggestive of differential associations by MSI status, albeit not statistically significant after Bonferroni correction ( $p_{\text{heterogeneity}} = 0.02$ ; Table 4). The corresponding HRs were 0.86 (95% CI 0.77, 0.95) for non-MSI-high colon cancer and 1.10 (95% CI 0.92, 1.32) for MSI-high colon cancer. Consistent patterns of the associations according to MSI and CIMP status were suggested within each cohort of men and

women (Supplemental Table 2). When colon cancer was jointly classified by CIMP and MSI statuses, while heterogeneity was not statistically significant with Bonferroni correction ( $p_{\text{heterogeneity}} = 0.05$ ; Table 3), an inverse association of dietary calcium was marked for the subtype with both CIMP-negative/low and non-MSI-high.

According to subtypes defined by *BRAF*, *KRAS* and *PIK3CA* mutations, generally no statistically significant heterogeneity was indicated (Table 4). Yet, an inverse association with dietary calcium was suggested more strongly for *BRAF* wild-type colon cancer than *BRAF* mutant, especially among men ( $p_{\text{heterogeneity}} = 0.02$ ; Supplemental Table 2).

Across the five molecular markers, the aforementioned patterns of associations appeared to persist in the analyses with dairy calcium (Supplemental Tables 3–5).

**Table 4** HR and 95% CI of dietary calcium intake and colon cancer risk overall and by molecular subtypes

	Dietary calcium intake (mg/day)				Per 300	$P^a_{\text{trend}}$	$P^b_{\text{heterogeneity}}$
	0–599	600–749	750–899	≥ 900			
Overall							NR
Cases, $n$	590	537	360	356	1843		
Age-adjusted <sup>c</sup>	1 (ref)	0.95 (0.84, 1.07)	0.87 (0.76, 1.00)	0.72 (0.63, 0.82)	0.86 (0.81, 0.91)	<0.001	
Multivariable <sup>d</sup>	1 (ref)	0.99 (0.87, 1.11)	0.94 (0.82, 1.07)	0.79 (0.68, 0.91)	0.89 (0.84, 0.95)	0.001	
CIMP status							0.01
Negative/low							
Cases, $n$	198	173	134	99	604		
Age-adjusted <sup>c</sup>	1 (ref)	0.95 (0.77, 1.16)	1.01 (0.81, 1.26)	0.61 (0.48, 0.78)	0.81 (0.73, 0.90)	<0.001	
Multivariable <sup>d</sup>	1 (ref)	0.97 (0.79, 1.20)	1.07 (0.85, 1.35)	0.67 (0.52, 0.86)	0.84 (0.76, 0.94)	0.003	
High							
Cases, $n$	49	61	32	47	189		
Age-adjusted <sup>c</sup>	1 (ref)	1.32 (0.90, 1.92)	0.95 (0.61, 1.49)	1.23 (0.82, 1.85)	1.08 (0.90, 1.29)	0.43	
Multivariable <sup>d</sup>	1 (ref)	1.37 (0.94, 2.01)	1.01 (0.65, 1.59)	1.35 (0.90, 2.04)	1.12 (0.93, 1.34)	0.23	
MSI status							0.02
non-MSI-high							
Cases, $n$	200	199	130	112	641		
Age-adjusted <sup>c</sup>	1 (ref)	1.06 (0.87, 1.29)	0.94 (0.75, 1.17)	0.67 (0.53, 0.84)	0.82 (0.74, 0.91)	<0.001	
Multivariable <sup>d</sup>	1 (ref)	1.09 (0.89, 1.34)	1.01 (0.80, 1.27)	0.74 (0.58, 0.94)	0.86 (0.77, 0.95)	0.004	
High							
Cases, $n$	50	51	37	45	183		
Age-adjusted <sup>c</sup>	1 (ref)	1.05 (0.71, 1.56)	1.06 (0.69, 1.63)	1.12 (0.75, 1.68)	1.06 (0.88, 1.27)	0.54	
Multivariable <sup>d</sup>	1 (ref)	1.10 (0.74, 1.63)	1.13 (0.73, 1.74)	1.23 (0.82, 1.87)	1.10 (0.92, 1.32)	0.31	
<i>BRAF</i> mutation							0.07
(–)							
Cases, $n$	203	199	136	121	659		
Age-adjusted <sup>c</sup>	1 (ref)	1.03 (0.85, 1.26)	0.96 (0.77, 1.19)	0.70 (0.56, 0.88)	0.83 (0.75, 0.92)	<0.001	
Multivariable <sup>d</sup>	1 (ref)	1.07 (0.87, 1.30)	1.02 (0.82, 1.28)	0.77 (0.60, 0.97)	0.87 (0.78, 0.96)	0.01	
(+)							
Cases, $n$	50	52	31	38	171		
Age-adjusted <sup>c</sup>	1 (ref)	1.10 (0.75, 1.63)	0.93 (0.59, 1.46)	1.01 (0.66, 1.54)	1.02 (0.84, 1.23)	0.87	
Multivariable <sup>d</sup>	1 (ref)	1.15 (0.78, 1.70)	0.99 (0.63, 1.56)	1.11 (0.72, 1.71)	1.06 (0.87, 1.28)	0.57	
<i>KRAS</i> mutation							0.84
(–)							
Cases, $n$	171	153	113	104	541		
Age-adjusted <sup>c</sup>	1 (ref)	0.94 (0.76, 1.18)	0.96 (0.76, 1.22)	0.75 (0.58, 0.95)	0.87 (0.78, 0.97)	0.01	
Multivariable <sup>d</sup>	1 (ref)	0.98 (0.78, 1.22)	1.03 (0.80, 1.31)	0.82 (0.63, 1.06)	0.91 (0.81, 1.02)	0.10	
(+)							
Cases, $n$	84	97	55	56	292		
Age-adjusted <sup>c</sup>	1 (ref)	1.22 (0.91, 1.64)	0.93 (0.66, 1.31)	0.77 (0.55, 1.09)	0.86 (0.74, 0.99)	0.04	
Multivariable <sup>d</sup>	1 (ref)	1.25 (0.93, 1.68)	0.99 (0.70, 1.40)	0.84 (0.59, 1.19)	0.89 (0.77, 1.03)	0.13	
<i>PIK3CA</i> mutation							0.42
(–)							
Cases, $n$	202	195	134	116	647		
Age-adjusted <sup>c</sup>	1 (ref)	1.02 (0.84, 1.25)	0.95 (0.76, 1.18)	0.69 (0.55, 0.87)	0.86 (0.78, 0.95)	0.003	
Multivariable <sup>d</sup>	1 (ref)	1.06 (0.87, 1.30)	1.02 (0.81, 1.28)	0.77 (0.60, 0.98)	0.90 (0.81, 1.00)	0.05	
(+)							
Cases, $n$	30	42	22	34	128		
Age-adjusted <sup>c</sup>	1 (ref)	1.38 (0.86, 2.21)	1.02 (0.58, 1.77)	1.29 (0.78, 2.12)	0.95 (0.77, 1.19)	0.68	

**Table 4** (continued)

	Dietary calcium intake (mg/day)				Per 300	$P^a_{\text{trend}}$	$P^b_{\text{heterogeneity}}$
	0–599	600–749	750–899	≥ 900			
Multivariable <sup>d</sup>	1 (ref)	1.43 (0.89, 2.29)	1.10 (0.63, 1.91)	1.41 (0.85, 2.34)	0.99 (0.80, 1.23)	0.94	

CIMP CpG island methylator phenotype, MSI microsatellite instability; NR, not relevant

<sup>a</sup>The  $p$  value for linear trend across calcium intake was from the Wald test on the continuous term of calcium intake

<sup>b</sup>The  $p$  value for heterogeneity across tumor markers was from the likelihood ratio test comparing the model in which a linear association with calcium intake was allowed to vary by tumor subtypes with the model in which a common association was assumed

<sup>c</sup>Age-adjusted analysis was adjusted for the same set of variables as denoted in Table 2

<sup>d</sup>Multivariable analysis was adjusted for supplemental calcium intake (<200, 200–299, 300–499, ≥500 mg/day) in addition to the same set of variables as denoted in Table 2

## Supplemental calcium

Supplemental calcium intake was not statistically significantly associated with overall colon cancer risk after adjusting for potential confounders (Table 5). Across the tumor subtypes tested, an inverse linear trend was suggested more consistently in women than in men, but no statistically significant heterogeneity was indicated within each cohort (Supplemental Table 6) and in the pooled analyses (Tables 3 and 5). Yet, among women, an inverse association was pronounced for *PIK3CA* mutant colon cancer ( $p_{\text{heterogeneity}} = 0.02$ ).

## Sensitivity analysis

In analyses of colorectal cancer risk by tumor markers, the overall pattern of heterogeneity generally persisted (Supplemental Tables 7–10). However, the differential associations by CIMP status became marginally insignificant after Bonferroni correction ( $p_{\text{heterogeneity}} = 0.02$ , Supplemental Table 8). Without considering the latency (Supplemental Table 11), although an inverse association became weaker for all sources of calcium and the heterogeneity tests were not statistically significant, we observed similarly differential associations of dietary calcium intake according to CIMP or MSI status. For all sources of calcium, after using inverse probability weighting to adjust for censoring of colon cancer due to missing subtype information, we observed consistent results on the associations between calcium intakes and the risk of colon cancer subtypes (Supplemental Table 12).

## Discussion

In these large, prospective cohort studies of women and men, higher dietary calcium intake was associated with lower risk of CIMP-negative/low colon cancer, but not of CIMP-high subtype. An inverse association with dietary calcium intake

was also suggested for non-MSI-high colon cancer but not for MSI-high subtype. Dietary calcium intake was not statistically significantly differentially associated with colon cancer risk by *BRAF*, *KRAS*, or *PIK3CA* mutations, although an inverse association appeared pronounced for *BRAF* wild-type than for *BRAF* mutant tumors. Taken together, our study suggests that there may be etiologic heterogeneity in colon cancer molecular subtypes in relation to calcium intake.

While observational studies generally indicate that sources of calcium do not modify its effect on colorectal cancer risk [3], our findings were most evident with dietary calcium. Bioavailability of calcium between dairy products, the major source of calcium intake in our cohorts, and calcium supplements are comparable [40]. However, in our cohort, dietary calcium intake was the primary determinant of inter-individual difference in total calcium intake. Dietary calcium intake was more consistent over time, while supplemental calcium intake increased considerably only recently especially among women [3]. In addition, it was better estimated by FFQ than supplemental calcium intake [10, 20]. These factors may partially explain more apparent heterogeneous associations observed for dietary calcium intake than for supplemental calcium intake or total calcium intake that incorporates supplemental calcium by definition.

The CIMP-high status correlates with MSI-high status, because aberrant CpG island methylation is one mechanism to inactivate DNA mismatch-repair genes, causing MSI-high tumor [18]. Furthermore, CIMP-high colorectal cancer is highly associated with *BRAF* mutation [18, 41] and arises through the serrated pathway rather than the classical adenoma–carcinoma pathway [42, 43]. Consistent with the molecular correlation among CIMP-high, MSI-high, and *BRAF* mutation, a possible inverse association of dietary calcium was pronounced with CIMP-negative/low colon cancer (and modestly suggested for non-MSI-high and *BRAF* wild-type colon cancers), but not with the other counterparts. Our findings are also consistent with the existing literature suggesting a stronger association of calcium intake with risk of

**Table 5** HR and 95% CI of supplemental calcium intake and colon cancer risk overall and by molecular subtypes

	Supplemental calcium intake (mg/day)					$P_{\text{trend}}^a$	$P_{\text{heterogeneity}}^b$
	0–199	200–299	300–499	≥ 500	Per 300		
Overall							NR
Cases, <i>n</i>	1,416	131	156	140	1,843		
Age-adjusted <sup>c</sup>	1 (ref)	0.85 (0.71, 1.02)	0.80 (0.67, 0.95)	0.75 (0.62, 0.89)	0.89 (0.83, 0.95) <sup>d</sup>	< 0.001	
Multivariable <sup>e</sup>	1 (ref)	0.92 (0.76, 1.10)	0.90 (0.76, 1.07)	0.86 (0.72, 1.04)	0.95 (0.89, 1.02)	0.17	
CIMP status							0.67
Negative/low							
Cases, <i>n</i>	464	42	47	51	604		
Age-adjusted <sup>c</sup>	1 (ref)	0.88 (0.64, 1.21)	0.80 (0.59, 1.09)	0.89 (0.66, 1.19)	0.96 (0.86, 1.07) <sup>d</sup>	0.48	
Multivariable <sup>e</sup>	1 (ref)	0.94 (0.68, 1.29)	0.87 (0.64, 1.19)	1.02 (0.75, 1.38)	1.02 (0.91, 1.14) <sup>d</sup>	0.71	
High							
Cases, <i>n</i>	137	12	24	16	189		
Age-adjusted <sup>c</sup>	1 (ref)	0.64 (0.35, 1.15)	1.01 (0.64, 1.57)	0.72 (0.43, 1.22)	0.91 (0.75, 1.11)	0.35	
Multivariable <sup>e</sup>	1 (ref)	0.67 (0.37, 1.22)	1.10 (0.70, 1.72)	0.83 (0.49, 1.41)	0.97 (0.80, 1.19)	0.79	
MSI status							0.61
non-MSI-high							
Cases, <i>n</i>	493	41	56	51	641		
Age-adjusted <sup>c</sup>	1 (ref)	0.81 (0.58, 1.12)	0.91 (0.68, 1.20)	0.82 (0.61, 1.10)	0.94 (0.85, 1.05) <sup>d</sup>	0.29	
Multivariable <sup>e</sup>	1 (ref)	0.87 (0.63, 1.21)	1.01 (0.75, 1.35)	0.96 (0.71, 1.30)	1.01 (0.91, 1.13)	0.80	
High							
Cases, <i>n</i>	137	11	20	15	183		
Age-adjusted <sup>c</sup>	1 (ref)	0.62 (0.33, 1.15)	0.85 (0.53, 1.38)	0.70 (0.41, 1.20)	0.88 (0.71, 1.08)	0.21	
Multivariable <sup>e</sup>	1 (ref)	0.67 (0.36, 1.24)	0.96 (0.59, 1.56)	0.83 (0.48, 1.43)	0.96 (0.78, 1.18)	0.67	
<i>BRAF</i> mutation							0.27
(–)							
Cases, <i>n</i>	510	42	59	48	659		
Age-adjusted <sup>c</sup>	1 (ref)	0.80 (0.58, 1.11)	0.92 (0.70, 1.21)	0.75 (0.55, 1.01)	0.90 (0.81, 1.01)	0.07	
Multivariable <sup>e</sup>	1 (ref)	0.86 (0.62, 1.19)	1.02 (0.77, 1.35)	0.87 (0.64, 1.18)	0.97 (0.86, 1.08)	0.55	
(+)							
Cases, <i>n</i>	124	12	17	18	171		
Age-adjusted <sup>c</sup>	1 (ref)	0.73 (0.40, 1.33)	0.81 (0.48, 1.36)	0.94 (0.57, 1.55)	1.02 (0.84, 1.24)	0.85	
Multivariable <sup>e</sup>	1 (ref)	0.78 (0.43, 1.43)	0.89 (0.53, 1.50)	1.10 (0.66, 1.83)	1.10 (0.90, 1.33)	0.36	
<i>KRAS</i> mutation							0.50
(–)							
Cases, <i>n</i>	410	30	54	47	541		
Age-adjusted <sup>c</sup>	1 (ref)	0.65 (0.45, 0.95)	0.95 (0.71, 1.28)	0.85 (0.62, 1.16)	0.95 (0.85, 1.07)	0.42	
Multivariable <sup>e</sup>	1 (ref)	0.70 (0.48, 1.02)	1.07 (0.79, 1.43)	0.99 (0.72, 1.36)	1.03 (0.91, 1.16)	0.66	
(+)							
Cases, <i>n</i>	226	24	22	20	292		
Age-adjusted <sup>c</sup>	1 (ref)	1.05 (0.68, 1.61)	0.76 (0.49, 1.19)	0.71 (0.44, 1.12)	0.90 (0.76, 1.06) <sup>d</sup>	0.20	
Multivariable <sup>e</sup>	1 (ref)	1.13 (0.73, 1.73)	0.83 (0.53, 1.31)	0.82 (0.51, 1.31)	0.96 (0.81, 1.13) <sup>d</sup>	0.62	
<i>PIK3CA</i> mutation							0.11
(–)							
Cases, <i>n</i>	489	43	57	58	647		
Age-adjusted <sup>c</sup>	1 (ref)	0.79 (0.57, 1.08)	0.84 (0.64, 1.12)	0.89 (0.68, 1.18)	0.96 (0.86, 1.07)	0.42	
Multivariable <sup>e</sup>	1 (ref)	0.85 (0.62, 1.17)	0.94 (0.70, 1.25)	1.05 (0.79, 1.40)	1.03 (0.92, 1.15)	0.58	
(+)							
Cases, <i>n</i>	102	7	15	4	128		
Age-adjusted <sup>c</sup>	1 (ref)	0.66 (0.30, 1.43)	1.12 (0.64, 1.97)	0.27 (0.10, 0.75)	0.77 (0.59, 1.01) <sup>d</sup>	0.06	

**Table 5** (continued)

	Supplemental calcium intake (mg/day)				Per 300	$p^a_{\text{trend}}$	$p^b_{\text{heterogeneity}}$
	0–199	200–299	300–499	≥ 500			
Multivariable <sup>c</sup>	1 (ref)	0.70 (0.32, 1.52)	1.24 (0.71, 2.18)	0.32 (0.12, 0.88)	0.83 (0.63, 1.08) <sup>d</sup>	0.17	

<sup>a</sup>The  $p$  value for linear trend across calcium intake was from the Wald test on the continuous term of calcium intake

<sup>b</sup>The  $p$  value for heterogeneity across tumor markers was from the likelihood ratio test comparing the model in which a linear association with calcium intake was allowed to vary by tumor subtypes with the model in which a common association was assumed

<sup>c</sup>Age-adjusted analysis was stratified by the same set of variables as denoted in Supplementary Table 2

<sup>d</sup> $p_{\text{heterogeneity}}$  by sex was < 0.05

<sup>e</sup>Multivariable analysis was adjusted for dietary calcium intake (< 600, 600–749, 750–899, ≥ 900 mg/day) in addition to the same set of variables as denoted in Table 2

distal colon cancer [2, 10, 11], of which CIMP-negative/low and non-MSI-high subtypes were more prevalent [12, 44].

To date, only few epidemiologic studies have evaluated etiologic heterogeneity of colorectal cancer subtypes by CIMP or MSI status in relation to calcium intake. In a case–case study conducted among 3119 colorectal cancer patients, calcium supplement use within 5 years of cancer diagnosis was not differentially associated with CIMP subtypes [45]. In a case–case study of 58 MSI-high and 278 non-MSI-high colon cancer patients, dietary calcium intake approximately 2 years before cancer diagnosis did not differ statistically significantly between the two subtypes of colon cancer [46]. In contrast, we observed differential associations by CIMP or MSI status. The discrepancy may reflect different study populations including different baseline calcium intake, or possibly the issue of etiologically relevant timing of calcium exposure. A time-lagged analysis suggests that adequate calcium intake may manifest its benefit against colorectal cancer incidence at least 8–12 years after the intake [2]. Therefore, previous studies that did not account for a long latency might have missed potential heterogeneous associations by CIMP or MSI status.

The potential mechanisms by which calcium differentially influences colorectal carcinogenesis by CIMP status may be linked to calcium sensing receptor (*CASR*). Expressed in multiple tissues including colonic epithelial cells, *CASR* is involved not only in controlling calcium homeostasis [47] but also in maintaining a balance across cell proliferation, differentiation, and apoptosis [48]. In our recent study, higher calcium intake was associated with a lower risk of *CASR*-positive but not *CASR*-negative colorectal cancer [49], which suggests a potential role of *CASR* in mediating chemopreventive action of calcium. Given that hypermethylation of the *CASR* promoter is one mechanism to suppress *CASR* expression [50], *CASR* expression levels are likely normal for CIMP-negative/low colon cancer but reduced for CIMP-high subtype. Indeed, we recently found that no or weak *CASR* expression was associated with CIMP-high status [51].

Our study has several strengths. This study represents one of the few epidemiologic studies that evaluated potential heterogeneous associations of calcium intake with colon cancer by major tumor molecular markers. Leveraging molecular pathological epidemiologic design [52], our study provides insights into biological mechanism linking calcium and colon cancer incidence and may help identify molecular subtypes of colon cancer that can benefit from potential chemopreventive action of calcium. The prospective cohort design, long-term follow-up, use of validated and updated measure of calcium intake, adjustment for a variety of potential confounders, and high follow-up rates further enhance the validity of our findings.

Yet, several limitations should be considered. First, testing multiple molecular markers increase the likelihood of chance findings. To address this concern, we applied Bonferroni correction and considered biological plausibility when interpreting our results. Particularly, CIMP status and MSI status are biologically correlated, and differential associations by these subtypes observed in our study are coherent with the molecular correlations. Second, tumor tissue was not available from every case from our cohorts, and relatively small number of cases of subtype-specific colon cancer limited our statistical power to detect a statistically significant association or heterogeneity. Third, residual confounding is likely because of the observational nature of our study. Finally, as health professionals, our participants are more likely to receive colonoscopies in which adenomas are removed, and thus, colon cancers are prevented even among those with the worst risk factors. This may have attenuated our estimated associations, and our findings have limited generalizability to other populations.

In conclusion, the association between calcium intake and colon cancer risk may vary by CIMP or MSI status, with the benefit possibly confined to CIMP-negative/low or non-MSI-high tumors. Given the sparse data on calcium intake and the risk of colon cancer subtypes and some inconsistent results across different calcium sources, our findings need to be confirmed in future studies.

**Acknowledgments** The authors would like to thank the participants and staff of the Health Professionals Follow-up Study and Nurses' Health Study for their valuable contributions. NK, LL, and TH—conceived of the project, performed the statistical analysis, wrote the paper, and had primary responsibility for all parts of the manuscript; and all authors—reviewed and interpreted the data, and read and approved the final manuscript.

**Funding** This work was supported by the U.S. National Institutes of Health (NIH) Grants (P01 CA87969 to M.J. Stampfer; UM1 CA186107 to M.J. Stampfer; P01 CA55075 to W.C. Willett; UM1 CA167552 to W.C. Willett; P50 CA127003 to C.S.F.; R01 CA118553 to C.S.F.; R01 CA 169141 to C.S.F.; R01 CA137178 to A.T.C.; K24 DK098311 to A.T.C.; R35 CA197735 to S.O.; R01 CA151993 to S.O.; K07 CA190673 to R.N.; R03 CA176717, and K07 CA188126 to X.Z.); Nodal Award from the Dana-Farber Harvard Cancer Center (to S.O.); by the American Cancer Society Research Scholar Grant (RSG NEC-130476 to X.Z.); and by grants from the Project P Fund, The Friends of the Dana-Farber Cancer Institute, Bennett Family Fund, and the Entertainment Industry Foundation through National Colorectal Cancer Research Alliance. N.K. was supported by grants from the National Research Foundation of Korea (NRF-2018R1C1B6008822; NRF-2018R1A4A1022589) and the Dongguk University Research Grant of 2017. T.H. was supported by a fellowship grant from the Mitsukoshi Health and Welfare Foundation. K.K. was supported by the Overseas Research Fellowship grant from the Japan Society for the Promotion of Science (JP2017-775). L.L. is supported by the grant from National Natural Science Foundation of China No. 81302491, a scholarship grant from the Chinese Scholarship Council and a fellowship grant from the Huazhong University of Science and Technology. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript. M.S. was supported by the 2017 AACR-AstraZeneca Fellowship in Immuno-oncology Research (17-40-12-SONG).

## Compliance with ethical standards

**Conflict of interest** The authors had no conflicts of interest to declare related to the study.

## References

- Li NF, Yao XG, Zhu J et al (2010) Higher levels of plasma TNF- $\alpha$  and neuropeptide y in hypertensive patients with obstructive sleep apnea syndrome. *Clin Exp Hypertens* 32:54–60
- Zhang X, Keum N, Wu K et al (2016) Calcium intake and colorectal cancer risk: results from the nurses' health study and health professionals follow-up study. *Int J Cancer* 139:2232–2242
- Keum N, Aune D, Greenwood DC, Ju W, Giovannucci EL (2014) Calcium intake and colorectal cancer risk: dose-response meta-analysis of prospective observational studies. *Int J Cancer* 135:1940–1948
- Cho E, Smith-Warner SA, Spiegelman D et al (2004) Dairy foods, calcium, and colorectal cancer: a pooled analysis of 10 cohort studies. *J Natl Cancer Inst* 96:1015–1022
- Huncharek M, Muscat J, Kupelnick B (2009) Colorectal cancer risk and dietary intake of calcium, vitamin D, and dairy products: a meta-analysis of 26,335 cases from 60 observational studies. *Nutr Cancer* 61:47–69
- Bristow SM, Bolland MJ, MacLennan GS et al (2013) Calcium supplements and cancer risk: a meta-analysis of randomised controlled trials. *Br J Nutr* 110:1384–1393
- Baron JA, Barry EL, Mott LA et al (2015) A trial of calcium and vitamin D for the prevention of colorectal adenomas. *N Engl J Med* 373:1519–1530
- Baron JA, Beach M, Mandel JS et al (1999) Calcium supplements for the prevention of colorectal adenomas. Calcium Polyp Prevention Study Group. *N Engl J Med* 340:101–107
- Grady WM, Carethers JM (2008) Genomic and epigenetic instability in colorectal cancer pathogenesis. *Gastroenterology* 135:1079–1099
- Wu K, Willett WC, Fuchs CS, Colditz GA, Giovannucci EL (2002) Calcium intake and risk of colon cancer in women and men. *J Natl Cancer Inst* 94:437–446
- Slattery ML, Neuhausen SL, Hoffman M et al (2004) Dietary calcium, vitamin D, VDR genotypes and colorectal cancer. *Int J Cancer* 111:750–756
- Curtin K, Slattery ML, Samowitz WS (2011) CpG island methylation in colorectal cancer: past, present and future. *Pathol Res Int* 2011:902674
- Yamauchi M, Morikawa T, Kuchiba A et al (2012) Assessment of colorectal cancer molecular features along bowel subsites challenges the conception of distinct dichotomy of proximal versus distal colorectum. *Gut* 61:847–854
- Rosty C, Young JP, Walsh MD et al (2013) PIK3CA activating mutation in colorectal carcinoma: associations with molecular features and survival. *PLoS ONE* 8:e65479
- Phipps AI, Buchanan DD, Makar KW et al (2012) BRAF mutation status and survival after colorectal cancer diagnosis according to patient and tumor characteristics. *Cancer Epidemiol Biomark Prev* 21:1792–1798
- Imamura Y, Morikawa T, Liao X et al (2012) Specific mutations in KRAS codons 12 and 13, and patient prognosis in 1075 BRAF wild-type colorectal cancers. *Clin Cancer Res* 18:4753–4763
- Markowitz SD, Bertagnolli MM (2009) Molecular origins of cancer: molecular basis of colorectal cancer. *N Engl J Med* 361:2449–2460
- Weisenberger DJ, Siegmund KD, Campan M et al (2006) CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat Genet* 38:787–793
- Feskanich D, Rimm EB, Giovannucci EL et al (1993) Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire. *J Am Diet Assoc* 93:790–796
- Rimm EB, Giovannucci EL, Stampfer MJ, Colditz GA, Litin LB, Willett WC (1992) Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. *Am J Epidemiol* 135:1114–1126 (**discussion 27–36**)
- Bailey RL, Dodd KW, Goldman JA et al (2010) Estimation of total usual calcium and vitamin D intakes in the United States. *J Nutr* 140:817–822
- Willett WC, Sampson L, Stampfer MJ et al (1985) Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol* 122:51–65
- Keum N, Cao Y, Oh H et al (2016) Sedentary behaviors and light-intensity activities in relation to colorectal cancer risk. *Int J Cancer* 138:2109–2117
- Giovannucci E, Liu Y, Rimm EB et al (2006) Prospective study of predictors of vitamin D status and cancer incidence and mortality in men. *J Natl Cancer Inst* 98:451–459

25. Hanyuda A, Ogino S, Qian ZR et al (2016) Body mass index and risk of colorectal cancer according to tumor lymphocytic infiltrate. *Int J Cancer* 139:854–868
26. Stampfer MJ, Willett WC, Speizer FE et al (1984) Test of the National Death Index. *Am J Epidemiol* 119:837–839
27. Morikawa T, Shima K, Kuchiba A et al (2012) No evidence for interference of h&e staining in DNA testing: usefulness of DNA extraction from H&E-stained archival tissue sections. *Am J Clin Pathol* 138:122–129
28. Ogino S, Kawasaki T, Brahmandam M et al (2006) Precision and performance characteristics of bisulfite conversion and real-time PCR (MethylLight) for quantitative DNA methylation analysis. *J Mol Diagn* 8:209–217
29. Ogino S, Kawasaki T, Kirkner GJ, Kraft P, Loda M, Fuchs CS (2007) Evaluation of markers for CpG island methylator phenotype (CIMP) in colorectal cancer by a large population-based sample. *J Mol Diagn* 9:305–314
30. Ogino S, Nosho K, Kirkner GJ et al (2009) CpG island methylator phenotype, microsatellite instability, BRAF mutation and clinical outcome in colon cancer. *Gut* 58:90–96
31. Ogino S, Kawasaki T, Brahmandam M et al (2005) Sensitive sequencing method for KRAS mutation detection by Pyrosequencing. *J Mol Diagn* 7:413–421
32. Imamura Y, Lochhead P, Yamauchi M et al (2014) Analyses of clinicopathological, molecular, and prognostic associations of KRAS codon 61 and codon 146 mutations in colorectal cancer: cohort study and literature review. *Mol Cancer* 13(1):135
33. Nosho K, Kawasaki T, Ohnishi M et al (2008) PIK3CA mutation in colorectal cancer: relationship with genetic and epigenetic alterations. *Neoplasia*. 10:534–541
34. Liao XY, Lochhead P, Nishihara R et al (2012) Aspirin Use, Tumor PIK3CA Mutation, and Colorectal-Cancer Survival. *New Engl J Med*. 367:1596–1606
35. Lee JE, Willett WC, Fuchs CS et al (2011) Folate intake and risk of colorectal cancer and adenoma: modification by time. *The American journal of clinical nutrition*. 93:817–825
36. Wolffenbuttel E (1955) Trepthonic eggs (incubated eggs) as medicine; orthobiotic foods, with special reference to auto-antibiosis in cancer. *Rev Bras Med* 12:388–392
37. Wang ML, Spiegelman D, Kuchiba A et al (2016) Statistical methods for studying disease subtype heterogeneity. *Stat Med* 35:782–800
38. Higgins JP, Thompson SG (2002) Quantifying heterogeneity in a meta-analysis. *Stat Med* 21:1539–1558
39. Cole SR, Hernan MA (2008) Constructing inverse probability weights for marginal structural models. *Am J Epidemiol* 168:656–664
40. Weaver CM, Proulx WR, Heaney R (1999) Choices for achieving adequate dietary calcium with a vegetarian diet. *Am J Clin Nutr* 70:543S–548S
41. Nosho K, Irahara N, Shima K et al (2008) Comprehensive bio-statistical analysis of CpG Island methylator phenotype in colorectal cancer using a large population-based sample. *PLoS ONE* 3(11):e3698
42. Kocarnik JM, Shiovitz S, Phipps AI (2015) Molecular phenotypes of colorectal cancer and potential clinical applications. *Gastroenterol Rep* 3:269–276
43. Kudryavtseva AV, Lipatova AV, Zaretsky AR et al (2016) Important molecular genetic markers of colorectal cancer. *Oncotarget* 7:53959–53983
44. Pancione M, Remo A, Colantuoni V (2012) Genetic and epigenetic events generate multiple pathways in colorectal cancer progression. *Pathol Res Int* 2012:509348
45. Weisenberger DJ, Levine AJ, Long TI et al (2015) Association of the colorectal CpG island methylator phenotype with molecular features, risk factors, and family history. *Cancer Epidemiol Biomark Prev* 24:512–519
46. Slattery ML, Anderson K, Curtin K, Ma KN, Schaffer D, Samowitz W (2001) Dietary intake and microsatellite instability in colon tumors. *Int J Cancer* 93:601–607
47. Brown EM (2013) Role of the calcium-sensing receptor in extracellular calcium homeostasis. *Best practice & research. Clin Endocrinol Metabol* 27:333–343
48. Aggarwal A, Prinz-Wohlgenannt M, Tennakoon S et al (2015) The calcium-sensing receptor: a promising target for prevention of colorectal cancer. *Biochim Biophys Acta* 1853:2158–2167
49. Yang W, Liu L, Masugi Y et al (2017) Calcium intake and risk of colorectal cancer according to expression status of calcium-sensing receptor (CASR). *Gut* 67(8):1475–1483
50. Hizaki K, Yamamoto H, Taniguchi H et al (2011) Epigenetic inactivation of calcium-sensing receptor in colorectal carcinogenesis. *Mod Pathol* 24:876–884
51. Momen-Heravi F, Masugi Y, Qian ZR et al (2017) Tumor expression of calcium sensing receptor and colorectal cancer survival: results from the nurses' health study and health professionals follow-up study. *Int J Cancer* 141:2471–2479
52. Ogino S, Nishihara R, VanderWeele TJ et al (2016) Review article: the role of molecular pathological epidemiology in the study of neoplastic and non-neoplastic diseases in the era of precision medicine. *Epidemiology* 27:602–611

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Affiliations

NaNa Keum<sup>1,2</sup>  · Li Liu<sup>1,3,4</sup> · Tsuyoshi Hamada<sup>3</sup> · Zhi Rong Qian<sup>3</sup> · Jonathan A. Nowak<sup>5</sup> · Yin Cao<sup>6</sup> · Annacarolina da Silva<sup>3</sup> · Keisuke Kosumi<sup>3</sup> · Mingyang Song<sup>1,7,8</sup> · Daniel Nevo<sup>9,10</sup> · Molin Wang<sup>9,10</sup> · Andrew T. Chan<sup>7,8,11</sup> · Jeffrey A. Meyerhardt<sup>3</sup> · Charles S. Fuchs<sup>12,13,14</sup> · Kana Wu<sup>1</sup> · Shuji Ogino<sup>1,3,5,10,15</sup> · Reiko Nishihara<sup>1,3,5,9,10</sup> · Xuehong Zhang<sup>11</sup>

✉ NaNa Keum  
nak212@mail.harvard.edu

<sup>1</sup> Department of Nutrition, Harvard T.H. Chan School of Public Health, Building 2, 3rd Floor, 665 Huntington Avenue, Boston, MA 02115, USA

<sup>2</sup> Department of Food Science and Biotechnology, Dongguk University, Goyang, South Korea

- <sup>3</sup> Department of Oncologic Pathology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA, USA
- <sup>4</sup> Department of Epidemiology and Biostatistics, and the Ministry of Education Key Lab of Environment and Health, School of Public Health, Huazhong University of Science and Technology, Wuhan, Hubei, People's Republic of China
- <sup>5</sup> Program in MPE Molecular Pathological Epidemiology, Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA
- <sup>6</sup> Division of Public Health Sciences, Department of Surgery, Washington University School of Medicine, St. Louis, MO, USA
- <sup>7</sup> Clinical and Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA
- <sup>8</sup> Division of Gastroenterology, Massachusetts General Hospital, Boston, MA, USA
- <sup>9</sup> Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA
- <sup>10</sup> Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA
- <sup>11</sup> Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA
- <sup>12</sup> Yale Cancer Center, New Haven, CT, USA
- <sup>13</sup> Department of Medicine, Yale School of Medicine, New Haven, CT, USA
- <sup>14</sup> Smilow Cancer Hospital, New Haven, CT, USA
- <sup>15</sup> Broad Institute of MIT and Harvard, Cambridge, MA, USA