

# A dietary pattern that is associated with C-peptide and risk of colorectal cancer in women

Teresa T. Fung · Frank B. Hu · Matthias Schulze · Michael Pollak · Tianying Wu · Charles S. Fuchs · Edward Giovannucci

## Abstract

**Purpose** Higher serum C-peptide concentrations have shown to be associated with an increased risk of colorectal cancer (CRC). Therefore, we used diet information to identify food groups that correlated with fasting serum concentrations of C-peptide and assess the association of this dietary pattern and CRC risk.

**Methods** Major food contributors to fasting C-peptide concentrations were identified with stepwise linear regression in a subsample ( $n = 833$ ) of women from a

large cohort. We then summed the consumption frequency of the major food contributors to form a C-peptide dietary pattern for the entire cohort ( $n = 66,714$ ). Risk for CRC was computed using Cox proportional hazard model with the C-peptide dietary pattern score as the predictor.

**Results** In up to 20 years of follow-up, we ascertained 985 cases of CRC and 758 colon cancer. After adjusting for confounders, the C-peptide dietary pattern, characterized by higher meat, fish, and sweetened beverage intake, but lower coffee, high fat dairy, and whole grains intake, showed direct association with CRC risk (RR comparing extreme quintiles = 1.29, 95 % CI = 1.05–1.58,  $p$  trend = 0.048). The same comparison was slightly stronger for colon cancer (RR = 1.35, 95 % CI = 1.07–1.70,  $p$  trend = 0.009). In stratified analysis, there was no association between the C-peptide dietary pattern and colon cancer among lean and active women. However, for overweight or sedentary women, RR for the same comparison was 1.58 (95 % CI = 1.20–2.07,  $p$  trend = 0.002) ( $p$  for interaction = 0.007).

**Conclusion** We derived a dietary pattern that correlated with C-peptide concentrations. This pattern was associated with an increase in colon cancer, especially among women who were overweight or sedentary.

---

T. T. Fung (✉)

Department of Nutrition, Simmons College, 300 The Fenway, Boston, MA 02115, USA  
e-mail: fung@simmons.edu

T. T. Fung · F. B. Hu · E. Giovannucci  
Department of Nutrition, Harvard School of Public Health, 667 Huntington Ave, Boston, MA 02115, USA

F. B. Hu · E. Giovannucci  
Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

M. Schulze  
German Institute of Human Nutrition, and Technical University of Munich, Munich, Germany

M. Pollak  
Lady Davis Institute for Medical Research, McGill University, Montreal, QC, Canada

T. Wu  
Department of Environmental Health, University of Cincinnati, Cincinnati, OH, USA

C. S. Fuchs  
Department of Medical Oncology, Dana Farber Cancer Institute, Boston, MA 02115, USA

**Keywords** Colorectal cancer · Diet · C-peptide

## Introduction

High concentrations of insulin or C-peptide (a marker of insulin) have shown a direct association with risk of colorectal cancer (CRC) in most studies [1–3]. A sedentary lifestyle and adiposity have been consistently associated with CRC, especially colon cancer, and are major determinants of insulin resistance and hyperinsulinemia. Thus, it

has been hypothesized that hyperinsulinemia accounts at least in part for the higher risk of colon cancer associated with a Western lifestyle [4]. Independent of effects on energy balance (e.g., obesity), some aspects of a Western dietary pattern may influence insulin resistance and secretion. For example, a dietary pattern characterized by higher levels of animal products and refined grains has been associated with higher C-peptide concentrations in US studies [5–7], while whole grains, fiber, fruits, vegetables, and coffee have typically been associated with lower concentrations of insulin or C-peptide [8–11].

Given the hypothesized role of hyperinsulinemia on colon cancer risk, it seems interesting to examine the potential role of the diet on colon cancer risk mediated through hyperinsulinemia. A dietary pattern may be more useful than examining single nutrients or food items, which individually may not affect insulin concentrations strongly enough to influence colon cancer risk appreciably. Glycemic index and load, a ranking system for carbohydrate sources based on their effect on blood glucose measures, and hence indirectly on insulin responses, have been inconsistently associated with colon cancer risk [12]. We recently derived a food insulin index, which directly quantifies the postprandial insulin secretion of a food and takes into account foods with a low or no carbohydrate content [13]. Although the food insulin index appears to be valid in estimating insulin responses for a given level of insulin resistance, it does not take into account the broader effect of diet on hyperinsulinemia. For example, some foods may influence the underlying insulin resistance, which could affect insulin concentrations, but would not be reflected in the dietary insulin index. Moreover, the food insulin index was computed using postprandial insulin response; therefore, it may not reflect the long term or basal insulin secretion. In fact, the dietary insulin index did not influence fasting C-peptide [13] nor colon cancer risk in recent studies [14].

An alternative and novel approach to examine the influence of the diet as a whole on insulin exposure is to derive empirically a dietary pattern associated with fasting C-peptide concentrations and then examine this pattern in relation to CRC risk. We thus used stepwise linear regression to derive a dietary pattern associated with C-peptide concentration and then examined the association between this dietary pattern and risk of CRC in women from the Nurses' Health Study.

## Methods

### Study population

The Nurses' Health Study (NHS) is a cohort study established in 1976 and consists of 121,700 female nurses aged

30–55 years living in 11 US states at the time [15]. Questionnaires are sent biennially to collect medical, lifestyle, and other health-related information. In 1980, participants completed a 61-item food frequency questionnaire (FFQ). This was expanded to 116 items in 1984, and similar FFQs were sent in 1986, 1990, 1994, 1998, and 2002. For this analysis, we used 1986 as baseline for the cancer analyses because this expanded FFQ is similar to the one used to compute the score. After excluding those with a history of cancer ( $n = 3,101$ ) (except nonmelanoma skin cancer), we included 66,714 women with follow-up from 1986 through 2006. This study was approved by the Institutional Review Board of the Brigham and Women's Hospital, Boston, MA.

### Biomarker subsample and assay

Blood was collected in 1989–1990 in a subsample of the NHS. Each willing participant was sent a blood collection kit containing instructions and needed supplies (e.g., blood tubes and needles). Each participant made arrangements for blood to be drawn, packaged the sample in an enclosed cool pack, and sent it to the laboratory by overnight courier. Almost all the samples arrived within 26 h of the blood draw. On their arrival at the laboratory, the whole-blood samples were centrifuged, aliquoted, and stored at temperatures not higher than  $-80^{\circ}\text{C}$ . The lifestyles and dietary intakes of women who returned a blood sample were in general similar to those who did not provide a blood sample. The women in this analysis were controls for previous nested case-control studies in myocardial infarct, breast cancer, colorectal cancer, and pancreatic cancer ( $n = 833$ ). Fasting plasma C-peptide was assayed using ELISA with reagents from Diagnostic Systems Laboratory (Webster, TX). The mean intra-assay coefficients of variation from the quality control samples were less than 13 % for C-peptide.

### Dietary assessment

Self-administered semi-quantitative FFQs were designed to assess average food intake over the preceding year. A standard portion size and nine possible consumption frequency categories, from “never, or <1/month” to “6+ times per day,” were given for each food. Total energy and nutrient intake were calculated by using the sum from all foods. Previous validation studies revealed reasonably good correlations between food intake and energy-adjusted nutrients assessed by the FFQ and multiple food records completed over the preceding year [16, 17].

### Case ascertainment

Incident colorectal cancer was ascertained between 1986 and 2006. In each biennial questionnaire, participants

self-report any diagnosis of colorectal cancer in the previous 2 years. We then sought permission to obtain medical records to confirm the diagnosis. Physicians unaware of dietary and lifestyle data reviewed and recorded information on histology, site, and stage of the cancer. Only adenocarcinomas of the colorectum were included, but we excluded cases associated with ulcerative colitis and familial polyposis.

#### Covariate ascertainment

Body mass index (BMI) was calculated from weight reported on each biennial questionnaire and height reported on the first questionnaire. Smoking, history of hypertension, aspirin use, multivitamin intake, menopausal status and use of postmenopausal hormone therapy, parity, and age at first birth were assessed every 2 years. Leisure-time physical activity was measured with validated questions on 10 common activities every 2 years. The information was then summed and calculated as metabolic equivalent hours (METs) [18]. Family history of colorectal cancers and information on lower bowel endoscopy screening were also collected. Deaths were identified from state vital statistics records, the National Death Index, reported by the families, and the postal system.

#### Derivation of dietary patterns

To account for laboratory drift because C-peptide for the participants was not assayed all at the same time, C-peptide was adjusted by dividing the original value with a ratio that is calculated by dividing the geometric mean of the batch by the mean of all batches [19]. Food items on the FFQ were categorized in 37 food groups as published previously [20]. Briefly, the food groups included fruits and vegetables (8 groups), meat (3 groups), fish, poultry, dairy (2 groups), potatoes (2 groups), eggs, whole grains, refined grains, alcoholic beverages (3 groups), desserts and sweets, snacks, coffee, tea, nuts, cream soups, other beverages (2 groups), condiments, pizza, salad dressings, butter, margarine, and mayonnaise-type spreads. Intake of each food group was standardized. To reduce within-subject variation, we used average intake of the 1986 and 1990 FFQs. We then used stepwise linear regression to derive linear functions of predictors (i.e., standardized food groups) that explain the most variation in C-peptide. Reduced rank regression (RRR) has been applied in studies to determine patterns from food intake data that maximize the explained variation in an intermediate set of biomarkers hypothesized to be related to the health outcome. However, in the special case of only one biomarker variable, multiple linear regression is identical to RRR [21].

Food groups with  $p$  value  $<0.1$  for the respective coefficient were retained in the model. The retained food

groups form the C-peptide dietary pattern. To calculate C-peptide pattern score for all women in the cohort, we first standardized intake of the food groups identified by stepwise linear regression, and then we summed the retained standardized food group items for each FFQ without regression weights as the C-peptide pattern score. Because the highest loading foods contribute most to the pattern, small differences in relative weights can be neglected without losing too much information. This simplification would generate pattern variables that are intuitive to interpret. A C-peptide pattern score was calculated for each FFQ for each woman.

#### Statistical analysis

We used Cox proportional hazard models to assess the association between this C-peptide dietary pattern score and risk of colorectal cancer during follow-up. To reduce random within-person variation and to best represent long-term dietary intake, we calculated cumulative averages of these scores from our repeated FFQs [22]. For example, the pattern scores in 1986 were used to predict colorectal cancer risk between 1986 and 1990, the average score from 1986 and 1990 was used to predict colorectal cancer from 1990 to 1994, and so forth.

In multivariate analyses, we adjusted for age, physical activity (quintiles), BMI (5 categories), energy intake (quintiles), alcohol intake (4 categories), multivitamin use (yes/no), history of colorectal polyps (yes/no), family history of colorectal cancer (yes/no), history of lower bowel endoscopy (yes/no), aspirin use (yes/no), and pack years of smoking (5 categories) with updated information at each 2-year questionnaire cycle. Test of trend was conducted by fitting continuous terms for the C-peptide food pattern score. We also examined associations separately by tumor location (colon, proximal colon and distal colon, rectum). Because any influence of dietary predictors of C-peptide would likely be strongly influenced by the insulin sensitivity status of the individual, we examined the difference of risk of colon cancer between lean and active women ( $n = 20,955$ ), and overweight or inactive women ( $n = 45,759$ ) by repeating our analysis for the C-peptide dietary pattern score by combined physical activity and BMI categories.

## Results

Women with higher serum C-peptide concentrations tended to be less active but were heavier (Table 1). We derived the C-peptide dietary pattern using stepwise linear regression ( $n = 833$ ). The C-peptide dietary pattern is characterized by higher meat, fish, and sweetened beverage

**Table 1** Lifestyle characteristics of participants according to quintiles of serum C-peptide levels at blood drawn,  $n = 833$ 

Median (ng/mL)	Q1	Q2	Q3	Q4	Q5	Spearman correlation coefficient ( $p$ value)
	0.75	1.04	1.31	1.70	2.67	
Age at blood drawn	56	57	58	59	58	0.13 (0.0002)
BMI	22.5	23.9	24.5	25.6	29.0	0.49 (0.0001)
Current smokers (%)	7	15	7	9	13	n/a <sup>a</sup>
Physical activity (METs/wk)	18	16	19	18	14	-0.08 (0.03)
Energy intake (kcal)	1,758	1,760	1,746	1,739	1,771	0.02 (0.95)
Alcohol (g)	7	6	6	5	5	-0.08 (0.02)
C-peptide pattern food groups <sup>b</sup>						
C-peptide food pattern score (median)	0.75	1.04	1.31	1.70	2.67	-
Red meat (servings/day)	0.50	0.55	0.50	0.53	0.60	0.07 (0.03)
Fish (servings/day)	0.34	0.32	0.34	0.35	0.34	0.04 (0.21)
High fat dairy (servings/day)	0.9	0.9	0.9	0.8	0.9	0.04 (0.27)
Coffee (servings/day)	2.5	2.3	2.3	2.2	1.9	-0.11 (0.002)
Whole grains (servings/day)	1.4	1.5	1.5	1.3	1.2	-0.05 (0.15)
Sugar-sweetened beverages (servings/day)	0.2	0.2	0.2	0.2	0.3	0.09 (0.01)
Cream soups (servings/day)	0.04	0.04	0.05	0.05	0.06	0.11 (0.02)

<sup>a</sup> Not applicable. Correlation coefficient cannot be computed because smoking status is not a continuous variable

<sup>b</sup> Average intake between 1986 and 1990 FFQ

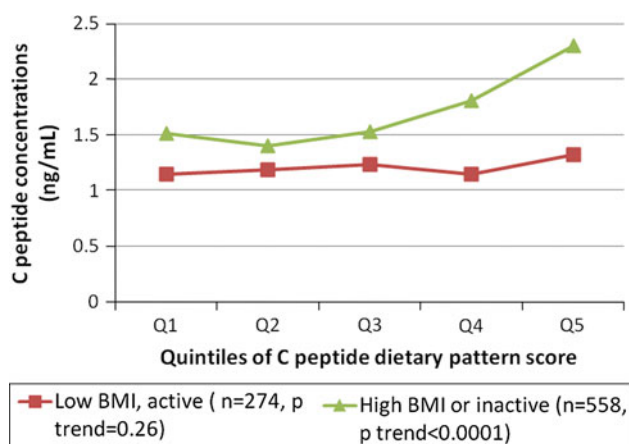
intake, but lower coffee, high fat dairy, and whole grains intake. The correlation between the C-peptide dietary pattern and measured C-peptide was 0.23 ( $p < 0.0001$ ) in the original 833 women. We validated the association between the C-peptide dietary pattern and serum C-peptide concentration with 721 women who were cases for the controls used in deriving the dietary patterns. Among lean and active cases, the difference in C-peptide concentrations

between the highest and lowest dietary score quintile was only 4 %. On the other hand, sedentary or overweight cases at the 5th quintile had C-peptide concentrations that were 38 % higher than cases in the first quintile. Women with higher C-peptide dietary pattern scores tended to have higher energy, folate, and calcium intake, were less likely to be current smokers, and had a higher BMI (Table 2). As shown in Fig. 1, actual C-peptide concentrations increased

**Table 2** Age-standardized baseline (1986) diet and lifestyle characteristics of participants according to quintiles of C-peptide dietary pattern scores ( $n = 66,714$ )

	Q1	Q2	Q3	Q4	Q5
C-peptide dietary pattern score (median)	-3.0	-1.3	-0.1	1.1	3.0
BMI	24.6	25.0	25.2	25.6	26.2
Current smokers (%)	25	21	20	19	20
Physical activity (METs/wk)	15	14	14	14	13
Energy intake (kcal)	1,757	1,663	1,678	1,738	1,997
Alcohol (g)	6	6	7	6	6
Folate <sup>a</sup> (mcg/d)	414	415	412	402	380
Calcium <sup>a</sup> (mg/d)	1,171	1,132	1,094	1,056	965
C-peptide pattern food groups					
Red meat (servings/day)	0.42	0.47	0.53	0.62	0.83
Fish (servings/day)	0.24	0.28	0.33	0.39	0.49
High fat dairy (servings/day)	1.77	1.07	0.93	0.85	0.86
Coffee (servings/day)	3.88	2.90	2.29	1.83	1.49
Whole grains (servings/day)	1.87	1.24	0.99	0.86	0.75
Sugar-sweetened beverages (servings/day)	0.08	0.10	0.14	0.20	0.60
Cream soups (servings/day)	0.02	0.03	0.04	0.05	0.10

<sup>a</sup> Adjusted for energy intake



**Fig. 1** Standardized C-peptide levels (ng/mL), adjusted by age and energy intake, by quintiles of C-peptide dietary pattern score, stratified by BMI and physical activity ( $n = 832$ ). At each quartiles of C-peptide score, C-peptide concentrations were significantly higher in women with high BMI or inactive than women with low BMI and active. (Filled square) Low BMI and active ( $n = 274$ ,  $p$  trend = 0.26). (Filled triangle) High BMI or inactive ( $n = 558$ ,  $p$  trend < 0.0001)

with increasing quintile of dietary C-peptide score among sedentary or overweight women. C-peptide concentrations for women at the 5th quintile of the dietary pattern score were 53 % higher than those at the 1st quintile. On the other hand, among women who were had a BMI < 25 kg/m<sup>2</sup> and who were relatively physically active (above the median, >10.2 METs at mid-point of follow-up), actual C-peptide concentration did not increase appreciably across increasing quintiles of the dietary pattern score. C-peptide concentrations only differed by 16 % between extreme quintiles among these women.

In 20 years of follow-up, we ascertained 985 cases of total colorectal cancer, of which 758 cases were colon

cancer and 211 were rectal cancer. The remaining cases were not classifiable at anatomical site. Among colon cancer, there were 467 proximal tumors and 291 distal tumors.

After adjusting for potential confounders, we observed a direct association between the C-peptide dietary pattern and colorectal cancer (RR comparing extreme quintiles = 1.29, 95 % CI = 1.05–1.58,  $p$  trend = 0.048) (Table 3). The association was slightly stronger for colon cancer (RR for same comparison = 1.35, 95 % CI = 1.07–1.70.  $p$  trend = 0.009) and particularly strong for tumors at the proximal colon (RR for the same comparison = 1.47, 95 % CI = 1.09–1.99,  $p$  trend = 0.006). We observed similar results when we restricted the analysis to women without a history of diabetes (data not shown).

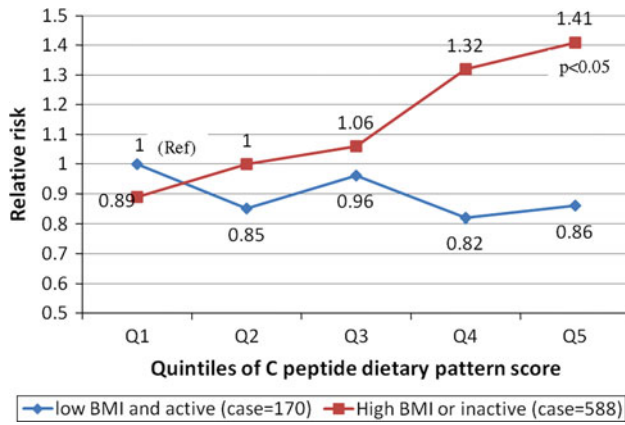
Because C-peptide is a marker of insulin secretion, we explored whether insulin resistance may modify the association between the C-peptide dietary pattern and colon cancer by modeling the association with joint classification of C-peptide dietary pattern and combined categories of BMI and physical activity. Risk for colon cancer was not elevated with higher C-peptide dietary pattern score among lean and active women (Fig. 2). However, overweight or sedentary women at the top quintile of the C-peptide dietary pattern score had an RR of 1.41 (95 % CI = 1.01–1.97), when compared with women at the bottom quintile and who were also lean and active. In stratified analysis by these two groups, we did not observe an association between the C-peptide dietary pattern and colon cancer (RR comparing extreme quintiles = 0.84, 95 % CI = 0.50–1.41,  $p$  trend = 0.23). However, among women who were overweight or sedentary, RR for the same comparison was 1.58 (95 % CI = 1.20–2.07,  $p$  trend = 0.002) ( $p$  for interaction = 0.007).

**Table 3** Relative risks (95 % CI) for risk of colorectal cancer by quintiles of C-peptide dietary pattern score ( $n = 66,714$ )

	Q1	Q2	Q3	Q4	Q5	$p$ trend
<b>Total colorectal cancer</b>						
Number of cases	189	194	200	200	202	
Age and energy adjusted	1	1.03 (0.84, 1.26)	1.09 (0.89, 1.33)	1.17 (0.96, 1.43)	1.30 (1.06, 1.59)	0.04
Multivariate adjusted <sup>a</sup>	1	1.03 (0.84, 1.26)	1.09 (0.89, 1.33)	1.18 (0.96, 1.44)	1.29 (1.05, 1.58)	0.048
<b>Colon cancer</b>						
Number of cases	143	146	152	161	156	
Age and energy adjusted	1	1.03 (0.81, 1.30)	1.10 (0.87, 1.38)	1.27 (1.01, 1.59)	1.36 (1.08, 1.72)	0.007
Multivariate adjusted <sup>a</sup>	1	1.03 (0.81, 1.29)	1.10 (0.87, 1.38)	1.27 (1.01, 1.60)	1.35 (1.07, 1.70)	0.009
<b>Rectal cancer</b>						
Number of cases	43	46	43	37	42	
Age and energy adjusted	1	1.09 (0.71, 1.65)	1.01 (0.66, 1.54)	0.91 (0.58, 1.41)	1.08 (0.70, 1.67)	0.54
Multivariate adjusted <sup>a</sup>	1	1.08 (0.71, 1.64)	1.01 (0.66, 1.55)	0.91 (0.58, 1.42)	1.07 (0.69, 1.65)	0.53

<sup>a</sup> Adjusted for age, energy intake, pack years of smoking, multivitamin use, family history, aspirin use, history of polyps, colonoscopy screening, alcohol intake, physical activity, BMI





**Fig. 2** Multivariate RR (95 % CI) for C-peptide dietary pattern score and risk of colon cancer jointly classified by BMI, physical activity ( $n = 66,714$ ), adjusted for age, energy intake, pack years of smoking, multivitamin use, family history, aspirin use, history of polyps, colonoscopy screening, alcohol intake. Low BMI and active = BMI less than 25 and physical activity at or above median,  $n = 20,955$ . High BMI or inactive = BMI above or equal 25, or physical activity below median,  $n = 45,759$ . (Filled square) High BMI or inactive ( $n = 45,759$ , case = 588). (Filled diagonal) Low BMI and active ( $n = 20,955$ , case = 170)

## Discussion

Our study was motivated by the relatively consistent finding between high insulin or C-peptide concentration and risk of colorectal or colon cancer [1–3, 23, 24], with relative risks ranging from 1.37 [1] to over 3 [23] comparing highest to lowest categories of serum levels. Strong biologic evidence supports a role of hyperinsulinemia and CRC risk [25]. Excess body weight and physical inactivity are key factors influencing insulin resistance and are also associated with colon cancer risk. Dietary factors are likely associated with colon cancer risk, but identifying the relevant factors has been challenging. Diet pattern could have a major role in influencing hyperinsulinemia by influencing insulin resistance and secretion, even independently of effects on energy balance. We thereby used an empirical approach to identify a dietary pattern that influenced C-peptide concentration and then examined this pattern in relation to CRC and colon cancer risk. In searching the literature, we did not find any previous diet-related cancer study that used C-peptide concentration to derive dietary patterns.

We identified a dietary pattern that moderately predicted C-peptide concentration. Although our approach was empirical, our findings were largely coherent with known factors. For example, lower whole grains coupled with higher sweetened beverages were characteristics of the C-peptide dietary pattern, and a plant-based diet that is higher in fiber has been linked to lower concentrations of C-peptide [8, 9]. In addition, higher coffee consumption is

also associated with lower C-peptide concentrations [11]. As predicted, the influence of the dietary pattern was strongly modified by BMI and physical activity level. For women who were not overweight and had moderate to high levels of physical activity, the C-peptide concentrations were relatively low irrespective of diet. For overweight or inactive women, C-peptide concentrations were only modestly elevated if they had a low dietary C-peptide score, but they were markedly elevated if they had higher C-peptide scores (Fig. 1).

In this analysis, we observed a direct association between the C-peptide dietary pattern score and CRC risk. The association was stronger for colon cancer, especially proximal colon cancer, and among women with high BMI or who were relatively physically inactive, who are at risk for insulin resistance. These results are consistent with studies of C-peptide or insulin concentrations and BMI and physical activity that typically show more consistent associations with colon than with rectal cancer [26]. The influence of the C-peptide dietary pattern only in women who were overweight or physically inactive is compatible with the premise that the dietary pattern influences risk predominantly in those who are at risk for insulin resistance. As shown in Fig. 1, the combined influence of dietary pattern and broad categories of BMI and physical activity yield substantial differences in C-peptide concentration.

The specific dietary factors that influence hyperinsulinemia may differ among diverse populations. We hypothesize that a common underlying etiologic factor related to colon carcinogenesis is hyperinsulinemia, but the relative importance of body weight, physical activity, dietary pattern, and genetic factors may vary across diverse populations. Thus, the approach that we used here requires confirmation in other populations.

Some of the major strengths of our cohort were the large sample size, validated dietary questionnaire, repeated assessments to dampen measurement error, and a large number with C-peptide measurement, which provided stable and verifiable patterns. Although not strictly a random sample of the cohort, those with measured C-peptide were largely randomly selected controls in nested case-control studies, so they should be broadly representative of the study population.

The strength of association between a biomarker-related dietary patterns and CRC not only depends on the correlation between the pattern and biomarkers, but also the strength of association between the biomarker and CRC. Since these are based on a single measurement, the true association if causal is likely to be stronger. In NHS, there was a suggestion of a direct association for C-peptide and colon cancer risk although it did not reach statistical significance [3]. Therefore, the magnitude of risk we observed

with the C-peptide dietary pattern for colon cancer appeared compatible with expectations. The main limitation is that this is an observational study, so we cannot entirely exclude confounding factors, though we do have detailed assessments of the known risk factors for colorectal cancer. In addition, dietary and lifestyle information was collected by self-report; therefore, some degree of measurement error is inevitable.

In conclusion, we derived a dietary pattern that is modestly associated with C-peptide concentration. This pattern that features higher intake of red meat and sweetened beverages and lower intake of whole grains and coffee is associated with an increase risk of colon cancer, especially in women who were overweight or sedentary. Because the C-peptide dietary pattern derived is study population-specific, the association between these food groups and C-peptide needs to be confirmed in other populations.

**Acknowledgments** We would like to thank the participants and staff of the Nurses' Health Study for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. This work is funded by the National Institute of Health research grants CA87969, HL60712, and CA95589.

## References

- Jenab M, Riboli E, Cleveland R et al (2007) Serum C-peptide, IGFBP-1 and IGFBP-2 and risk of colon and rectal cancers in the European prospective investigation into cancer and nutrition. *Int J Cancer* 121:368–376
- Otani T, Iwasaki M, Sasazuki S, Inoue M, Tsugane S, Group JPHC-bPS (2007) Plasma C-peptide, insulin-like growth factor-I, insulin-like growth factor binding proteins and risk of colorectal cancer in a nested case-control study: the Japan public health center-based prospective study. *Int J Cancer*. 120: 2007–2012
- Wei EK, Ma J, Pollak MN, Fuchs CS, Hankinson SE, Giovannucci EL (2005) A prospective study of C-peptide, insulin-like growth factor-I, insulin-like growth factor binding protein-1, and the risk of colorectal cancer in women. *Cancer Epidemiol Biomarkers Prev* 14:850–855
- Giovannucci EL (1995) Insulin and colon cancer. *Cancer Causes Control* 164–179:16
- Fung TT, Rimm E, Spiegelman D et al (2001) Association between dietary patterns and plasma biomarkers of obesity and cardiovascular disease risk. *Am J Clin Nutr* 73:61–67
- Giovannucci EL, Rimm E, Liu Y, Willett WC (2004) Height, predictors of C-peptide and cancer risk in men. *Int J Epidemiol* 33:217–225
- Kerver JM, Yang EJ, Bianchi L, Song WO (2003) Dietary patterns associated with risk factors for cardiovascular disease in healthy US adults. *Am J Clin Nutr* 78:1103–1110
- Bohe G, Murphy G, Rogers CJ et al (2010) Serum adiponectin, leptin, C-peptide, homocysteine, and colorectal adenoma recurrence in the polyp prevention trial. *Cancer Epidemiol Biomarkers Prev* 19:1441–1452
- Jensen MK, Koh-Banerjee P, Franz M, Sampson L, Gronbaek H, Rimm E (2006) Whole grains, bran, and germ in relation to homocysteine and markers of glycemic control, lipids, and inflammation 1. *Am J Clin Nutr* 83:2275–2283
- Loopstra-Masters Rc, Liese AD, Haffner SM, Wagenknecht LE, Hanley AJ (2011) Associations between the intake of caffeinated and decaffeinated coffee and measures of insulin sensitivity and beta cell function. *Diabetologia* 54:320–328
- Wu T, Willett WC, Hankinson SE, Giovannucci EL (2005) Caffeinated coffee, decaffeinated coffee, and caffeine in relation to plasma C-peptide levels, a marker of insulin secretion, in US women. *Diabetes Care* 28:1390–1396
- Gnagnarella P, Gandini S, La Vecchia C, Maisonneuve P (2008) Glycemic index, glycemic load, and cancer risk: a meta-analysis. *Am J Clin Nutr* 87:1793–1801
- Nimptsch K, Jc Brand-Miller, Franz M, Sampson L, Willett WC, Giovannucci EL (2011) Dietary insulin index and insulin load in relation to biomarkers of glycemic control, plasma lipids, and inflammation markers. *Am J Clin Nutr* 94:182–190
- Bao Y, Nimptsch K, Meyerhardt JA et al (2010) Dietary insulin load, dietary insulin index, and colorectal cancer. *Cancer Epidemiol Biomarkers Prev* 19:3020–3026
- Colditz GA, Martin P, Stampfer MJ et al (1986) Validation of questionnaire information on risk factors and disease outcomes in a prospective cohort of women. *Am J Epidemiol* 123:894–900
- Willett WC (1998) *Nutritional epidemiology*. Oxford University Press, New York
- Feskanih D, Rimm E, Giovannucci EL et al (1993) Reproducibility and validity of food intake measurements from a semi-quantitative food frequency questionnaire. *J Am Dietetic Assoc* 93:790–796
- Ainsworth BE, Haskell WL, Whitt MC et al (2000) Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sport Exerc* 32:S498–S504
- Rosner BA, Cook N, Portman R, Daniels S, Falkner B (2008) Determination of blood pressure percentiles in normal-weight children: some methodological issues. *Am J Epidemiol* 167:653–666
- Hu FB, Rimm E, Smith-Warner SA et al (1999) Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. *Am J Clin Nutr* 69:243–249
- McNaughton SA, Mishra GD, Brunner EJ (2008) Dietary patterns, insulin resistance, and incidence of type 2 diabetes in the Whitehall II study. *Diabetes Care* 31:1343–1348
- Hu FB, Stampfer MJ, Rimm E et al (1999) Dietary fat and coronary heart disease: a comparison of approaches for adjusting for total energy intake and modeling repeated dietary measurements. *Am J Epidemiol* 149:531–540
- Kaaks R, Toniolo P, Akhmedkhanov A et al (2000) Serum C-peptide, insulin-like growth factor (IGF)-I, IGF-binding proteins, and colorectal cancer risk in women. *Int J Cancer* 92:1592–1600
- Le Merchand L, Wang H, Rinaldi S et al (2010) Associations of plasma C-peptide and IGFBP-1 levels with risk of colorectal adenoma in a multiethnic population. *Cancer Epidemiol Biomarkers Prev* 19:1471–1477
- Giovannucci EL, Harlan DM, Archer MC et al (2010) Diabetes and cancer: a consensus report. *CA Cancer J Clin* 60:207–221
- Ning Y, Wang L, Giovannucci EL (2010) A quantitative analysis of body mass index and colorectal cancer: findings from 56 observational studies. *Obesity Rev* 11:19–30